RESEARCH   **Open Access**

Regularized supervised Bayesian approach for image deconvolution with regularization parameter estimation

Bouchra Laaziri[*], Said Raghay and Abdelilah Hakim

Abstract

Image deconvolution consists in restoring a blurred and noisy image knowing its point spread function (PSF). This inverse problem is ill-posed and needs prior information to obtain a satisfactory solution. Bayesian inference approach with appropriate prior on the image, in particular with a Gaussian prior, has been used successfully. Supervised Bayesian approach with maximum a posteriori (MAP) estimation, a method that has been considered recently, is unstable and suffers from serious ringing artifacts in many applications. To overcome these drawbacks, we propose a regularized version where we minimize an energy functional combined by the mean square error with H^1 regularization term, and we consider the generalized cross validation (GCV) method, a widely used and very successful predictive approach, for choosing the smoothing parameter. Theoretically, we study the convergence behavior of the method and we give numerical tests to show its effectiveness.

Keywords: Image deconvolution, Supervised Bayesian approach, MAP estimation, Regularization, GCV method

1 Introduction

Images are indispensable in science and everyday life. Mirroring the capabilities of our human visual system, it is natural to display observations of the world in graphical form. Images are obtained in areas ranging from everyday photography to astronomy, medical imaging, remote sensing, and microscopy. In each case, there is an underlying object we wish to observe, which is the *original* or *true* image. Indeed, this *true* image is the ideal representation of the observed scene.

Yet, the observation process is never ideal: there is uncertainty in the measurements, such as blur, noise, and other types of degradation. Image restoration aims to recover an estimate of the true image from the degraded observations. The key to being able to solve this problem is proper incorporation of prior knowledge about the original image into the restoration process.

Classical image restoration seeks an estimate of the true image $f(x, y)$ when the point spread function (PSF) $h(x, y)$ is known a priori, so it is assumed that the observed image $g(x, y)$ is the output of a linear spatially invariant system

$$g(x, y) = h(x, y) * f(x, y) + \epsilon(x, y), \tag{1}$$

where * represents the convolution operation and $\epsilon(x, y)$ the errors. Therefore, it becomes a non-blind deconvolution problem.

The discretized version of model (1) is given by

$$g = Hf + \epsilon, \tag{2}$$

where H represents the 2D convolution matrix obtained from the PSF of the imaging system. For more details on this discretization refer to the paper [1].

Image deconvolution is an ill-posed problem. The solution may not depend continuously on the data, may not be unique, or it may not even exist, this means that, in practice, the data g alone is not sufficient to define a unique and satisfactory solution. Practical approaches as

*Correspondence: bouchra.laaziri@gmail.com
Laboratory of Applied Mathematics and Computer Science, Faculty of Science and Techniques, Cadi Ayyad University, Marrakesh, Morocco

regularization theory and Bayesian inversion have been successful for this task [2–8].

The earliest classical methods for non-blind image deconvolution include the Wiener filter [9, 10] and the Richardson-Lucy algorithm [11, 12]. The Wiener filter works in the frequency domain, attempting to minimize the impact of deconvolved noise at frequencies. It has widespread use as the frequency spectrum of most visual images is fairly well behaved and may be estimated easily. It is simple and effective for some special images. However, it is not stable, has serious ringing artifacts in many applications, and its restoration quality is still limited (results often too blurred in the case of gaussian blurred images in which the kernel is dense).

The Richardson-Lucy algorithm maximizes the likelihood that the resulting image, when convolved with the PSF, is an instance of the blurred image, assuming Poisson noise statistics. A key limitation of this iterative method is the considerable computation it takes to obtain a (visually) stable solution due to the low convergence speed of this type of algorithm.

Also, increasing the number of iterations not only slows down the computational process but also magnifies noise and introduces waves near sharp edges.

Another popular regularization method for denoising and deblurring is the total variation (TV) method. It was introduced to image denoising by Rudin, Osher, and Fatemi [13] and then applied to deconvolution by Rudin and Osher [14]. It preserves edges well but has sometimes undesirable staircase effect, namely, the transformation of smooth regions into piecewise constant regions (stairs), which implied that the finer details in the original image may not be recovered satisfactorily.

Based on the spirit of TV regularization and bilateral filtering, Elad et al. [15, 16] proposed a bilateral TV (BTV) as a new regularization term that is computationally cheap to implement and preserves edges. This term is more robust to noise and can preserve more details, but it tends to remove texture, create flat intensity regions, and new contours that lead to cartoon-like images.

As well as this regularization theory, the supervised Bayesian approach is a method that has been used successfully [17–22]. In order to increase stability and overcome some limitations of this method (ringing effect), we propose a novel regularized version for image deconvolution. We consider H^1 regularization term incorporating the non-negative constraint of the restored image and the well-known generalized cross validation (GCV) method for choosing suitable regularization parameters. Experiments show that the proposed approach is superior to the Wiener filter, the Richardson-Lucy algorithm, the TV method, and the BTV approach, widely used in literature, in terms of restoration quality and stability.

The remaining of this article is organized as follows. In Section 2, we review the supervised Bayesian approach and then propose our deconvolution variational model. We give the experimental tests and discuss the obtained results in Section 3. The Section 4 concludes the paper.

2 Methods
2.1 Bayesian approach
From this point, the main objective is to infer on f given the forward model (2). In the classical reconstruction problems with the known blur kernel h, we mean to use the Bayes rule:

$$p(f|g) = \frac{p(g|f)p(f)}{p(g)} \propto p(g|f)p(f) \qquad (3)$$

to obtain what is called the *posterior* law $p(f|g)$ from the likelihood $p(g|f)$ and the prior $p(f)$, and we can infer on f using this law.

2.1.1 Bayesian estimation with simple prior
The Bayesian inference approach is based on the posterior law:

$$p(f|g, \theta_1, \theta_2) = \frac{p(g|f, \theta_1)p(f|\theta_2)}{p(g|\theta_1, \theta_2)} \propto p(g|f, \theta_1)p(f|\theta_2) \qquad (4)$$

where the term $p(g|f, \theta_1)$ is the likelihood, $p(f|\theta_2)$ is the prior model, $\theta = (\theta_1, \theta_2)$ are their corresponding parameters (often called the hyper-parameters of the problem) and $p(g|\theta_1, \theta_2)$ is called the evidence of the model. This relation is showed in the following scheme:
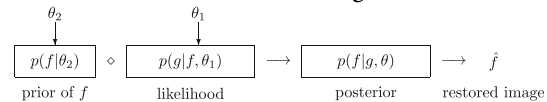
2.1.2 Assigning the likelihood $p(g|f)$ and the prior $p(f)$
We consider the forward model (2) and suppose that we have some prior knowledge about the error term ϵ. In fact, if we can assign a probability law $p(\epsilon)$, then we can deduce the likelihood term $p(g|f)$.

To account for possible non-stationarity of the noise, we propose to use

$$p(\epsilon|v_\epsilon) = \mathcal{N}(\epsilon|0, V_\epsilon) \text{ with } V_\epsilon = Dg(v_\epsilon), \qquad (5)$$

where $v_\epsilon = [v_{\epsilon_1}, \cdots, v_{\epsilon_M}]^T$ contains the unknown variances of the non-stationary noise and Dg is a diagonal matrix whose entries are the M elements of vector v_ϵ.

From this, we can define the expression of the likelihood:

$$p(g|f, v_\epsilon) = \mathcal{N}(g|Hf, V_\epsilon). \qquad (6)$$

To be able to estimate the unknown variances, we assign an Inverse Gamma conjugate prior on v_{ϵ_i}:

$$p(v_\epsilon) = \prod_i p(v_{\epsilon_i}) \text{ with } p(v_{\epsilon_i}) = \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_i}, \beta_{\epsilon_i}), \forall i,$$

$$(7)$$

where α_{ϵ_i} and β_{ϵ_i} are two positive parameters. The shape parameter α_{ϵ_i} controls the height of the probability density function of the Inverse Gamma distribution, and the scale parameter β_{ϵ_i} controls the spread [23].

The next step is to assign a prior to the unknown *f*. Here too, different approaches can be used, we may have some prior information such as the mean and the variance of the unknown quantity. The objective is to assign a prior law $p(f|\theta_2)$ in such a way to translate our incomplete prior knowledge on *f*.

We propose

$$p(f|v_f) = \mathcal{N}(f|0, V_f) \text{ with } V_f = Dg(v_f).$$

$$(8)$$

We also assign an Inverse Gamma conjugate prior on v_{f_j}:

$$p(v_f) = \prod_j p(v_{f_j}) \text{ with } p(v_{f_j}) = \mathcal{IG}(v_{f_j}|\alpha_{f_j}, \beta_{f_j}), \forall j. \quad (9)$$

2.1.3 Supervised Bayesian approach

We consider ϵ, v_ϵ and v_f are known. We can define the expression of the likelihood as

$$p(g|f, v_\epsilon) = \mathcal{N}(g|Hf, v_\epsilon I).$$

$$(10)$$

We propose the prior for *f*:

$$p(f|v_f) = \mathcal{N}(f|0, v_f I),$$

$$(11)$$

with the supposes of the parameters and hyper-parameters, the joint posterior of all the unknowns becomes

$$p(f|g, v_\epsilon, v_f) \propto p(g|f, v_\epsilon)p(f|v_f)$$

$$(12)$$

We can obtain the estimate of *f* as below:

$$\hat{f} = \arg\max_f \left\{ p(f|g, v_\epsilon, v_f) \right\}$$

$$= \arg\max_f \left\{ \exp(-\frac{1}{2v_\epsilon} \| g - Hf \|_2^2)\exp(-\frac{1}{2v_f} \| f \|_2^2) \right\}$$

$$= \arg\min_f \left\{ \frac{1}{2v_\epsilon} \| g - Hf \|_2^2 + \frac{1}{2v_f} \| f \|_2^2 \right\}. \quad (13)$$

So the criterion to be optimized is a quadratic one:

$$J(f) = \frac{1}{2v_\epsilon} \| g - Hf \|_2^2 + \frac{1}{2v_f} \| f \|_2^2$$

$$\propto \| g - Hf \|_2^2 + \lambda_f \| f \|_2^2 \quad (14)$$

where $\lambda_f = \frac{v_\epsilon}{v_f}$. If we work it out directly, this criterion has a solution:

$$\hat{f} = (H^T H + \lambda_f I)^{-1} H^T g$$

$$(15)$$

Using the singular value decomposition (SVD) of *H* by assuming $\mathcal{H}$, the Fourier transform of *H*, to be circular bloc-circulant (CBC), it can be shown that $\hat{f}$ in (15) can be computed using the Fourier transform. The result is comparable to the Wiener filter:

$$\hat{\mathcal{F}} = \frac{\overline{\mathcal{H}}\mathcal{G}}{\overline{\mathcal{H}}\mathcal{H} + \lambda_f},$$

$$(16)$$

where $\mathcal{F}$ and $\mathcal{G}$ denote the Fourier transforms of *f* and *g*, respectively.

2.2 The proposed method

In the situation where ϵ, v_ϵ and v_f are known, we can use the MAP estimation. As mentioned, this estimation is unstable and suffers from serious ringing artifacts. To overcome these drawbacks, we propose a regularized MAP method where we minimize an energy functional combined by the mean square error with H^1 regularization term:

$$\arg\min_f \left\{ \| g - Hf \|_2^2 + \lambda_f \| f \|_2^2 + \lambda \| \nabla f \|_2^2 \right\}, \quad (17)$$

where $\nabla = (\nabla_h, \nabla_v)$ is the gradient operator combined by difference operators along horizontal and vertical directions. The last two terms are regularization terms which ask that *f* should be smooth in H^1 sense.

The staircase effect is partly due to the fact that the used norms are not biased against discontinuous nor continuous functions (e.g., TV-norm). The term $\| \nabla f \|_2^2$ has a strong bias against discontinuous functions, so this model substantially reduces the staircase effect and recover the smooth region's value in the image.

We use the Neumann boundary condition, a natural choice in image processing [24], to discretize the gradient by a finite difference scheme. This type of boundary condition requires that $\lambda_f \neq 0$ in the aim to prove the coercivity of the proposed functional in the space $H^1(\Omega)$, Ω is a bounded Lipschitz domain in $\mathbb{R}^2$, and then the existence of a solution of the minimization problem (17). Numerically, this choice avoids remarkably the apparition of a blurring effect in the resulting images instead of fixing λ_f at 0, using an empirical reduction rule to set it from large to small. We will give more details on this in the next section.

Solving the problem using Fourier properties results in a simple algorithm, that is easier to implement and takes a very short time to converge.

2.2.1 The convergence behavior

Let *E* be a Banach space, $F : E \longrightarrow \mathbb{R}$, and consider the minimization problem

$$\inf_{u \in E} F(u)$$

$$(18)$$

Theorem 1 *Let E be a reflexive Banach space and F : $E \longrightarrow \mathbb{R}$ a sequentially weakly lower semi-continuous and coercive function, then the problem* (18) *has a solution. Furthermore, if F is strictly convex, then this solution is unique.*

Proof Proving the existence of a solution of problem (18) is usually achieved by the following steps, which constitute the direct method of the calculus of variations:

- Step 1: One constructs a minimizing sequence $u_n \in E$, i.e., a sequence satisfying $\lim\limits_{n \longrightarrow +\infty} F(u_n) = \inf\limits_{u \in E} F(u)$.

- Step 2: F is coercive $\left(\lim\limits_{\|u\|_E \longrightarrow +\infty} F(u) = +\infty\right)$, one can obtain a uniform bound $\| u_n \|_E \leq C, C > 0$. E is reflexive, then by the theorem of weak sequential compactness [25] one deduces the existence of $u_0 \in E$ and of a subsequence denoted also as $(u_n)_n$ such that $u_n \underset{E}{\rightharpoonup} u_0$.

- Step 3: F is sequentially weakly lower semi-continuous (for all sequence $x_n \underset{E}{\rightharpoonup} x$ we have $\lim\limits_{x_n \rightharpoonup x} F(x_n) \geq F(x)$), so $\lim\limits_{u_n \rightharpoonup u_0} F(u_n) \geq F(u_0)$, which obviously implies that $F(u_0) = \inf\limits_{u \in E} F(u)$.

F is strictly convex ($F(\lambda x + (1-\lambda)y) < \lambda F(x) + (1-\lambda)F(y)$, for all $x \neq y \in E$ and $\lambda \in]0;1[$); therefore, the minimum is unique. $\square$

So, we have

$$F : H^1(\Omega) \to \mathbb{R}$$
$$f \mapsto \| g - Hf \|_2^2 + \lambda_f \| f \|_2^2 + \lambda \| \nabla f \|_2^2 \tag{19}$$

Existence:

- $H^1(\Omega)$ is a Hilbert Banach reflexive space.
- F is coercive :

$$\begin{aligned} F(f) &\geq \lambda_f \| f \|_2^2 + \lambda \| \nabla f \|_2^2 \\ &\geq \min(\lambda_f, \lambda)(\| f \|_2^2 + \| \nabla f \|_2^2) \\ &\geq \alpha \| f \|_{H^1}^2, \end{aligned} \tag{20}$$

where $\alpha = \min(\lambda_f, \lambda) > 0$.

- F is sequentially weakly lower semi-continuous.

Let $f_n \longrightarrow f$ in H^1, then $\| f_n \|_2^2 \longrightarrow \| f \|_2^2$ and $\| \nabla f_n \|_2^2 \longrightarrow \| \nabla f \|_2^2$.

Furthermore, $g - Hf_n = g - H(f_n - f) - Hf$, as $f_n \longrightarrow f$ in H^1, $f_n - f \longrightarrow 0$ in $H^1 \Longrightarrow g - Hf_n \longrightarrow g - Hf$ in H^1, finally $\| g - Hf_n \|_2^2 \longrightarrow \| g - Hf \|_2^2$.

So, the problem admits a solution.

Uniqueness:

The function F is strictly convex, then the solution is unique.

In the Fourier domain, our model (17) is equivalent to

$$\arg\min_{\mathcal{F}} \left\{ \| \mathcal{G} - \mathcal{HF} \|_2^2 + \lambda_f \| \mathcal{F} \|_2^2 + \lambda \| D\mathcal{F} \|_2^2 \right\}, \tag{21}$$

where $D = (D_h, D_v)$ denotes the Fourier transform of $\nabla = (\nabla_h, \nabla_v)$, and

$$\| D\mathcal{F} \|_2^2 = \| D_h \mathcal{F} \|_2^2 + \| D_v \mathcal{F} \|_2^2 . \tag{22}$$

By taking the Wirtinger derivative of functional in (21) with respect to $\mathcal{F}$ and setting the result to be 0, we get the optimal condition of $\mathcal{F}$ as follows:

$$-\overline{\mathcal{H}}\mathcal{G} + |\mathcal{H}|^2 \mathcal{F} + \lambda_f \mathcal{F} + \lambda |D|^2 \mathcal{F} = 0, \tag{23}$$

where $|D|^2 = |D_h|^2 + |D_v|^2$. The above equality gives the solution of $\mathcal{F}$

$$\hat{\mathcal{F}} = \frac{\overline{\mathcal{H}}\mathcal{G}}{|\mathcal{H}|^2 + \lambda_f + \lambda|D|^2} \tag{24}$$

We use the inverse Fourier transform to get the estimation.

2.2.2 A parameter selection method for H^1 regularization

We now consider a parameter choice method. An appropriate selection of the regularization parameters λ_f and λ is important in the regularization. The well-known methods for this purpose are the L-curve [26] and the GCV method [27, 28]; here, we consider the GCV one. It is a widely used and very successful predictive method for choosing the smoothing parameter. The basic idea is that, if one datum point is dropped, then a good value of the regularization parameter should predict the missing datum value fairly well. In our case, the regularization parameter λ_f is first selected by the empirical reduction rule, then λ is chosen to minimize the GCV function

$$\text{GCV}(\lambda) = \frac{\| \left(I - HH_\lambda^{-1} H^T \right) g \|_2^2}{\left[tr \left(I - HH_\lambda^{-1} H^T \right) \right]^2}, \tag{25}$$

where $H_\lambda = \mathcal{H}^T \mathcal{H} + \lambda_f I + \lambda D^T D$. This function can be simplified using the Generalized Singular Value Decompositions (GSVD) [29] of the pair $(\mathcal{H}, D)$. Thus, there exist orthonormal matrices U, V and invertible matrix X such that

$$U^T \mathcal{H} X = C = Dg(c_1, ..., c_N), \qquad c_i \geq 0, \tag{26}$$

$$V^T D X = S = Dg(s_1, ..., s_N), \qquad s_i \geq 0, \tag{27}$$

where $N = mn$. Therefore, the GCV function when used with this regularization can be simplified to

$$GCV(\lambda) = \frac{\sum_{i=1}^N \left(s_i^2 \tilde{g}_i / \left(c_i^2 + \lambda s_i^2 + \lambda_f \right) \right)^2}{\left(\sum_{i=1}^N s_i^2 / \left(c_i^2 + \lambda s_i^2 + \lambda_f \right) \right)^2}, \tag{28}$$

with $\tilde{g} = U^T g$. For the particular case where the matrix D reduces to the identity I, the GSVD of the pair $(\mathcal{H}, I)$ reduces to the SVD of the matrix $\mathcal{H}$ and the expression of GCV is given by the following formula

$$\text{GCV}(\lambda) = \frac{\sum_{i=1}^{N} (\tilde{g}_i/(\sigma_i^2 + \lambda + \lambda_f))^2}{(\sum_{i=1}^{N} 1/(\sigma_i^2 + \lambda + \lambda_f))^2}, \qquad (29)$$

where σ_i is the ith singular value of the matrix $\mathcal{H}$.

$GCV(\lambda)$ in this case is a continuous function, so we use the MATLAB function *fminbnd*, which is based on a combination of golden section search and quadratic interpolation search, to find the value of λ at which $GCV(\lambda)$ is minimized.

Beginning with an initial image, the following sequence defines our algorithm:

Algorithm 1 MAP-H^1 in Fourier domain

Compute the regularization operator $D = (D_h, D_v)$.
1. Fourier transform of g to obtain $\mathcal{G}$.
2. Fourier transform of h to obtain $\mathcal{H}$.
3. Estimate the regularization parameter λ_f to calculate λ using (28).
4. Compute $\hat{\mathcal{F}}$ using (24).
5. Inverse Fourier transform of $\hat{\mathcal{F}}$ to obtain $\hat{f}$.

---

3 Results and discussion

This section presents a culmination of all the numerical tests we performed of the proposed approach for solving image deconvolution problem, and compares it with the Wiener filter, the Richardson-Lucy deconvolution [30], the TV approach, and the BTV method. We consider the problem (2), the blurring matrix H is determined from the PSF h [31] of the imaging system which defines how each pixel is blurred. We use three different types of blur kernel: binary blur kernel of size 21×21 and normalized elements to sum 1 (Fig. 1) and Gaussian blur kernel of size 20×20 and standard deviation 3 and a Motion blur kernel of length 15, generated by MATLAB routine fspecial('gaussian', 20, 3) and fspecial('motion', 15), respectively. This choice is for showing the effectiveness of our approach against different types of degradation.

For comparison, it is hard to determine whether one method is better than the others just by looking at the images; therefore, it is necessary to compute the peak signal-to-noise ratio (PSNR), which is defined as

$$\text{PSNR}(\hat{f}, f) = 10 \log 10 \left(\frac{255^2}{\frac{1}{st} \| \hat{f} - f \|_2^2} \right)$$
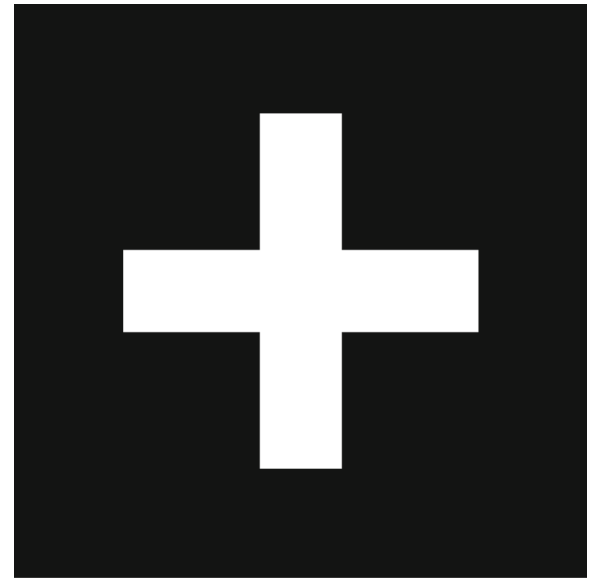

Fig. 1 First kernel of convolution. The first kernel used to blur images

where s and t are numbers of row and column of the image. Note that PSNR is a standard image quality measure that is widely used in comparing image restoration results. We also use the Structural Similarity Index Measure (SSIM), an image quality metric that assesses the visual impact of three characteristics of an image: luminance, contrast, and structure [32].

A crucial issue in solving the problem is the determination of the regularization parameters λ_f and λ. A good selection of the parameters will result in a promising deblurring result, whereas a bad choice may lead to slow convergence as well as the existence of severe artifacts in the results. Bigger values lead to smoother deblurred images and more stability of the algorithm. However, too big λ_f and λ will over smooth the image details and decrease the restoration quality. Generally, when the degradation in the blurry image is significant, the values need to be set large, to reduce the blur as much as possible. However, in the continuing iterations, the blurry effect is decreased gradually. In this case, small values are required since large values will damage the fine details in the image. By considering these effects, a direct implementation is to set λ_f from large to small according to an empirical reduction rule [33–35]

$$\lambda_f^{t+1} = \max \left(\lambda_f^t . r, \lambda_f^{\min} \right);$$

Fig. 2 Set of seven images used in the tests. This is a set of seven images used as an original image in the tests

which depends on the initial value λ_f^0, the minimal value $\lambda_f^{\min}$ and the reduction factor $r \in (0,1)$. We choose $r = 0.5$. This setting ensures the improvement of the convergence speed of the algorithm. We compute the other regularization parameter λ using the estimated value of λ_f and the GCV function (28).

A projection on the convex set C is required to impose the image constraint

$$C := \{f \mid 0 \le f(n) \le 1, \forall n\}$$

For fair comparison, in Wiener filter [36], $< |\hat{b}|^2 >$ denotes the spectral of $|\hat{b}|^2$ which can be estimated from the noise level. $< |\hat{b}|^2 > = \frac{\alpha}{\text{variance}(g)}$, we look for the optimal α that gives us the best result in each case. Furthermore, we consider the number of iterations as the stopping criterion in the Richardson-Lucy algorithm.

The output image in both methods exhibits ringing introduced by the discrete Fourier transforms used, so to reduce this effect, we use the function *edgetaper* before the processing.

In TV and BTV-based image restoration methods, the computations are too intensive to run until convergence (even with a large tolerance). Instead, we run until a specified reasonable maximum iteration.

We try to get the best result with each method and then compare it with our approach.

We use seven images in our experiments which are the standards for image processing (Fig. 2). They are of different gray-level histogram. The blurry versions are obtained by convolving the original images with the three PSFs defined previously. We show the deconvolution results in Figs. 3, 4, 5, 6, 7, 8, and 9, and PSNR and SSIM values in Tables 1, 2, and 3.

By carefully observing these results, we find that the edges and details are well recovered, and the ringing artifacts are well suppressed too in each case using our approach instead of the Wiener filter, the Richardson-Lucy algorithm, the TV method, and the BTV approach (slow convergence methods). Also, more the structure and details of the image are important more the degradation is worse and the restoration quality is less. In terms of image quality measures, the proposed MAP-H^1 has the highest PSNR and SSIM values among all, the BTV method is the

Fig. 3 Deconvolution of *Lena* image. From left to right, the first column represents degradation by kernel 1, the second column for the degradation by kernel 2, and by kernel 3 in the third column. From top to bottom, the order of the rows is as follows: blurred images using the three kernels respectively, restored images using RL algorithm [11] (kernel 1 (iterations = 100), kernel 2 (iterations = 200) and kernel 3 (iterations = 40)), restored images using Wiener filter [9], restored images using TV method [14], restored images using BTV approach [16], and finally the restored images using our proposed MAP-H^1 model

Fig. 4 Deconvolution of *Cameraman* image. From left to right, the first column represents degradation by kernel 1, the second column for the degradation by kernel 2, and by kernel 3 in the third column. From top to bottom, the order of the rows is as follows: blurred images using the three kernels respectively, restored images using RL algorithm [11] (kernel 1 (iterations = 90), kernel 2 (iterations = 300), and kernel 3 (iterations = 200)), restored images using Wiener filter [9], restored images using TV method [14], restored images using BTV approach [16], and finally the restored images using our proposed MAP-H^1 model
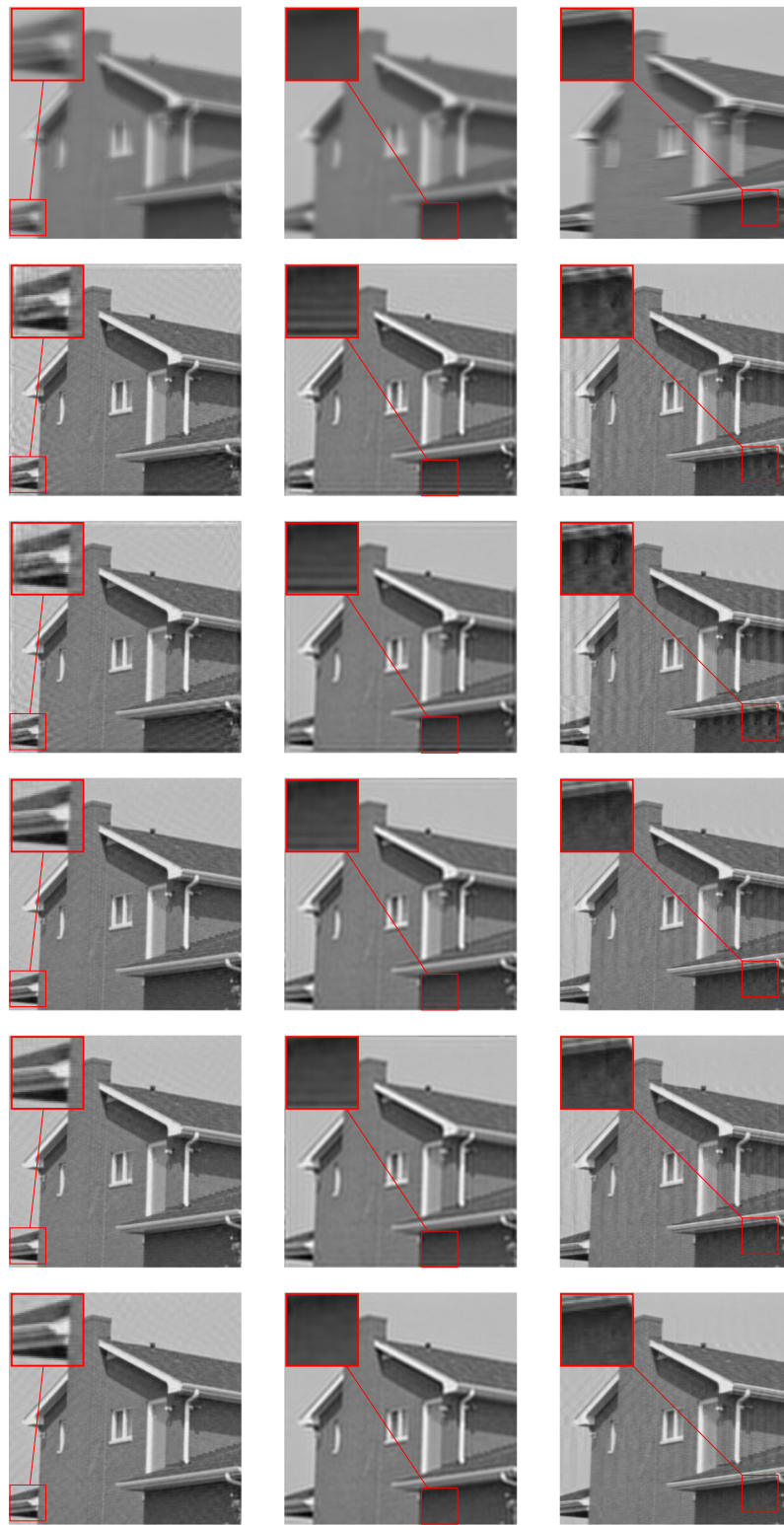
Fig. 5 Deconvolution of *House* image. From left to right, the first column represents degradation by kernel 1, the second column for the degradation by kernel 2, and by kernel 3 in the third column. From top to bottom, the order of the rows is as follows: blurred images using the three kernels respectively, restored images using RL algorithm [11] (kernel 1 (iterations = 70), kernel 2 (iterations = 140), and kernel 3 (iterations = 40)), restored images using Wiener filter [9], restored images using TV method [14], restored images using BTV approach [16], and finally the restored images using our proposed MAP-H^1 model
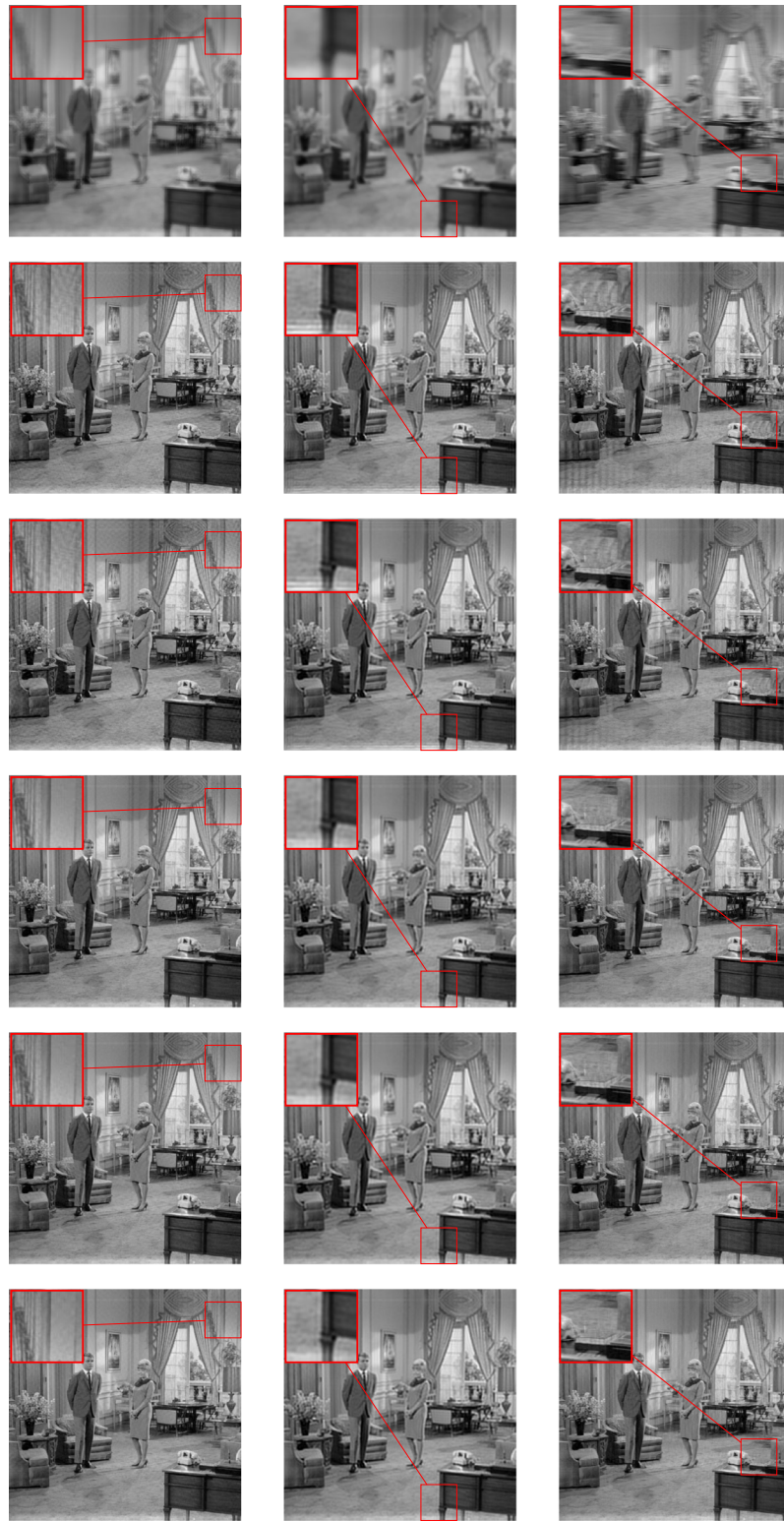
Fig. 6 Deconvolution of *Couple* image. From left to right, the first column represents degradation by kernel 1, the second column for the degradation by kernel 2, and by kernel 3 in the third column. From top to bottom, the order of the rows is as follows: blurred images using the three kernels respectively, restored images using RL algorithm [11] (kernel 1 (iterations = 150), kernel 2 (iterations = 310), and kernel 3 (iterations = 90)), restored images using Wiener filter [9], restored images using TV method [14], restored images using BTV approach [16], and finally the restored images using our proposed MAP-H^1 model
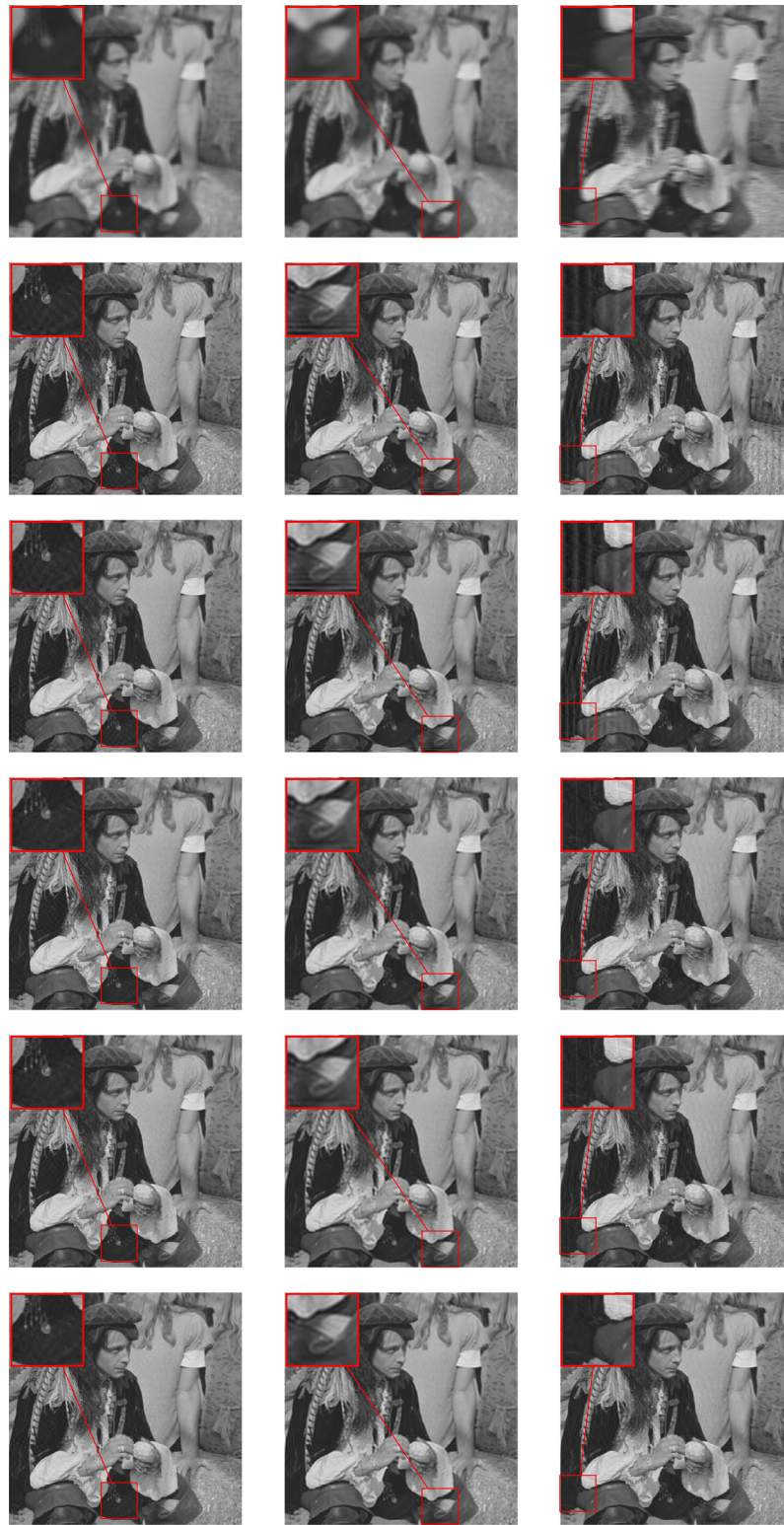
Fig. 7 Deconvolution of *Pirate* image. From left to right, the first column represents degradation by kernel 1, the second column for the degradation by kernel 2, and by kernel 3 in the third column. From top to bottom, the order of the rows is as follows: blurred images using the three kernels respectively, restored images using RL algorithm [11] (kernel 1 (iterations = 130), kernel 2 (iterations = 230), and kernel 3 (iterations = 80)), restored images using Wiener filter [9], restored images using TV method [14], restored images using BTV approach [16], and finally the restored images using our proposed MAP-H^1 model

Fig. 8 Deconvolution of *Boat* image. From left to right, the first column represents degradation by kernel 1, the second column for the degradation by kernel 2, and by kernel 3 in the third column. From top to bottom, the order of the rows is as follows: blurred images using the three kernels respectively, restored images using RL algorithm [11] (kernel 1 (iterations = 150), kernel 2 (iterations = 330), and kernel 3 (iterations = 80)), restored images using Wiener filter [9], restored images using TV method [14], restored images using BTV approach [16], and finally the restored images using our proposed MAP-H^1 model
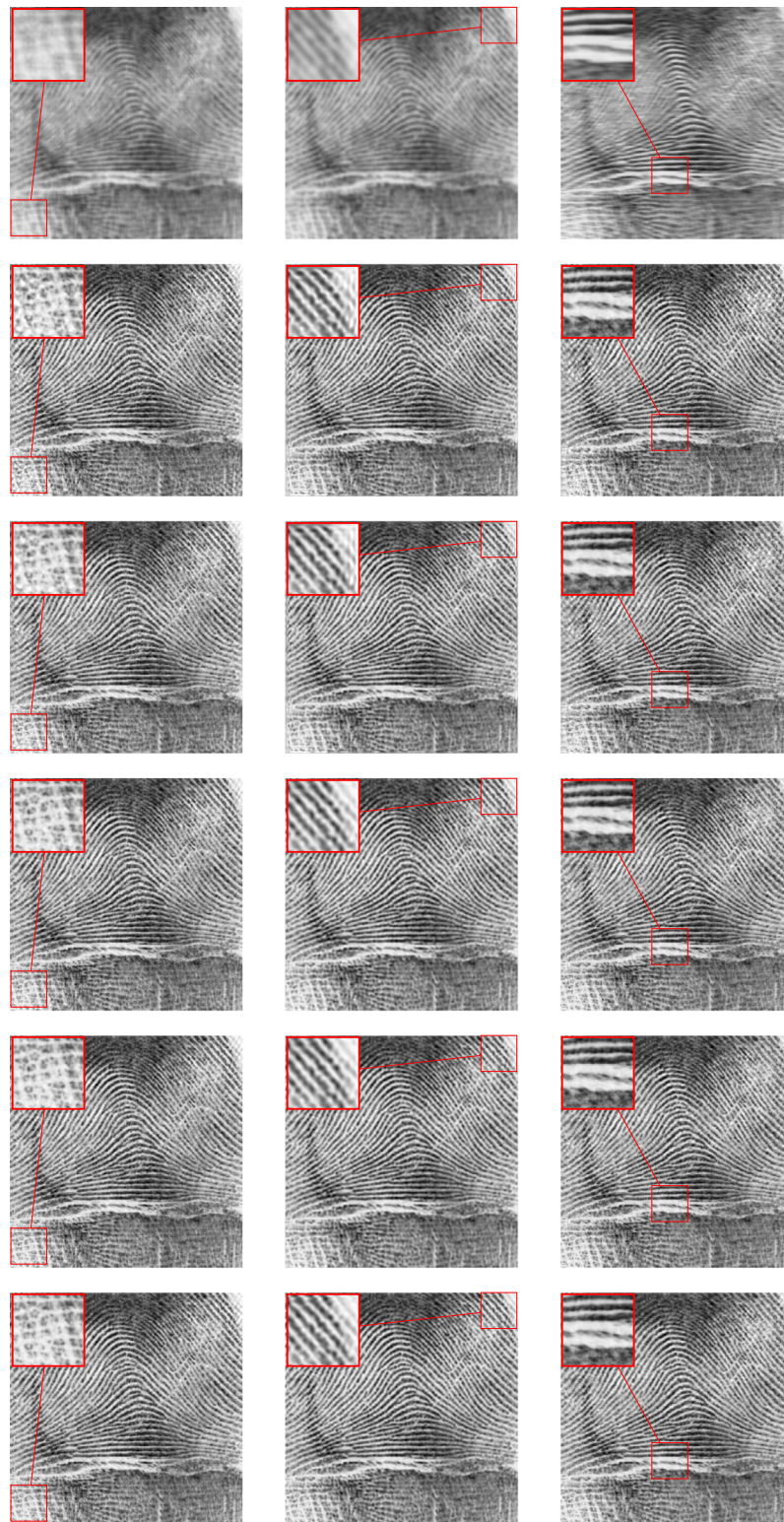
Fig. 9 Deconvolution of *Fingerprint* image. From left to right, the first column represents degradation by kernel 1, the second column for the degradation by kernel 2, and by kernel 3 in the third column. From top to bottom, the order of the rows is as follows: blurred images using the three kernels respectively, restored images using RL algorithm [11] (kernel 1 (iterations = 100), kernel 2 (iterations = 240), and kernel 3 (iterations = 90)), restored images using Wiener filter [9], restored images using TV method [14], restored images using BTV approach [16], and finally the restored images using our proposed MAP-H^1 model

Table 1 PSNR and SSIM values obtained by five methods for seven different images degraded by kernel 1

Image	Size	Measure	Blurred	Method				
				RL	Wiener	TV	BTV	MAP-H^1
Lena	256 × 256	PSNR	23.15	30.73	31.01	32.75	33.17	**33.95**
		SSIM	0.6731	0.8428	0.8647	0.9147	0.9202	**0.9381**
Cameraman	512 × 512	PSNR	24.66	34.90	35.22	36.27	36.37	**37.67**
		SSIM	0.7824	0.9251	0.9249	0.9527	0.9546	**0.9651**
House	256 × 256	PSNR	25.09	31.97	32.00	35.59	35.94	**36.13**
		SSIM	0.7379	0.8502	0.8569	0.9263	0.9292	**0.9347**
Couple	512 × 512	PSNR	24.17	33.48	33.77	34.11	34.49	**34.89**
		SSIM	0.6127	0.9139	0.9158	0.9373	0.9397	**0.9458**
Pirate	512 × 512	PSNR	24.75	33.75	33.55	33.80	34.24	**34.70**
		SSIM	0.6268	0.9114	0.9140	0.9231	0.9283	**0.9365**
Boat	512 × 512	PSNR	23.98	33.57	33.68	33.60	33.95	**34.52**
		SSIM	0.6181	0.9044	0.9051	0.9151	0.9161	**0.9276**
Fingerprint	512 × 512	PSNR	17.54	26.84	26.98	28.56	28.66	**29.27**
		SSIM	0.3292	0.9177	0.9189	0.9495	0.9507	**0.9564**

For each test setting, six results are provided: Blurred, Richardson-Lucy algorithm, Wiener filter, TV method, BTV approach, and our proposed model. *Bold* format: the best score in each line

second best, in general, following by the TV approach, the Wiener filter and finally the Richardson-Lucy algorithm.

Depending on the size of the image, the execution of the main proposed algorithm requires an average of 2 to 20 s on Intel(R) Celeron(R) CPU N2815 1.86 GHz computer, making it faster than the other methods.

4 Conclusions

In this work, we have extended the supervised Bayesian approach by adding H^1 regularization term in an energy formulation and by proposing a method for choosing the regularization parameter. Numerical results show that the proposed algorithm is stable and can suppress ringing artifacts successfully using the proposed techniques instead of the Wiener filter, the Richardson-Lucy algorithm, the TV method, and the BTV approach, robust methods used in literature, for different types of the blur kernel. Future works will be focused more on using other blur kernels and tests on real images.

Table 2 PSNR and SSIM values obtained by five methods for seven different images degraded by kernel 2

Image	Size	Measure	Blurred	Method				
				RL	Wiener	TV	BTV	MAP-H^1
Lena	256 × 256	PSNR	23.14	26.52	26.45	26.91	26.94	**27.88**
		SSIM	0.6794	0.7720	0.7824	0.7943	0.7967	**0.8410**
Cameraman	512 × 512	PSNR	24.38	31.24	30.76	30.29	30.38	**31.58**
		SSIM	0.7737	0.8959	0.8987	0.8895	0.8910	**0.9153**
House	256 × 256	PSNR	24.74	28.93	28.60	29.18	29.26	**30.73**
		SSIM	0.7264	0.7861	0.7860	0.7982	0.8011	**0.8469**
Couple	512 × 512	PSNR	23.60	26.83	26.97	26.75	26.77	**27.44**
		SSIM	0.5625	0.7341	0.7431	0.7292	0.7293	**0.7653**
Pirate	512 × 512	PSNR	24.67	28.10	28.06	28.09	28.16	**28.83**
		SSIM	0.6282	0.7778	0.7852	0.7739	0.7766	**0.8068**
Boat	512 × 512	PSNR	23.63	27.31	27.39	27.08	27.10	**27.81**
		SSIM	0.5939	0.7546	0.7623	0.7454	0.7457	**0.7774**
Fingerprint	512 × 512	PSNR	18.07	25.64	25.96	26.23	26.37	**27.40**
		SSIM	0.3938	0.8568	0.8642	0.8660	0.8683	**0.8959**

For each test setting, six results are provided: blurred, Richardson-Lucy algorithm, Wiener filter, TV method, BTV approach, and our proposed model. *Bold* format: the best score in each line

Table 3 PSNR and SSIM values obtained by five methods for seven different images degraded by kernel 3

Image	Size	Measure	Blurred	Method				
				RL	Wiener	TV	BTV	MAP-H^1
Lena	256 × 256	PSNR	21.56	28.02	28.33	31.85	32.17	**32.21**
		SSIM	0.6006	0.8049	0.8281	0.8892	0.8948	**0.9070**
Cameraman	512 × 512	PSNR	23.12	36.80	36.42	34.83	34.95	**39.97**
		SSIM	0.7824	0.9513	0.9446	0.9362	0.9386	**0.9698**
House	256 × 256	PSNR	24.79	32.03	32.01	35.74	35.92	**36.65**
		SSIM	0.7639	0.8919	0.8937	0.9367	0.9392	**0.9510**
Couple	512 × 512	PSNR	23.68	31.71	32.59	33.89	33.75	**34.43**
		SSIM	0.6191	0.9109	0.9296	0.9387	0.9367	**0.9475**
Pirate	512 × 512	PSNR	24.04	33.01	33.26	34.37	34.68	**36.06**
		SSIM	0.6217	0.9145	0.9197	0.9274	0.9308	**0.9492**
Boat	512 × 512	PSNR	23.32	31.07	31.67	33.12	33.30	**34.33**
		SSIM	0.6408	0.8927	0.9026	0.9169	0.9174	**0.9395**
Fingerprint	512 × 512	PSNR	17.57	26.67	26.78	28.55	28.55	**29.15**
		SSIM	0.3930	0.9352	0.9348	0.9558	0.9561	**0.9580**

For each test setting, six results are provided: blurred, Richardson-Lucy algorithm, Wiener filter, TV method, BTV approach, and our proposed model. *Bold* format: the best score in each line

Abbreviations
BTV: Bilateral total variation; CBC: Circular bloc-circulant; GCV: Generalized cross validation; GSVD: Generalized singular value decompositions; MAP: Maximum a posteriori; PSF: Point spread function; PSNR: Peak signal-to-noise ratio; RL: Richardson-Lucy; SSIM: Structural Similarity Index Measure; SVD: Singular value decomposition; TV: Total variation

Author's contributions
The three authors contributed equally to the realization of this work. The author(s) read and approved the final manuscript.

Authors' information
Bouchra LAAZIRI, first author, PhD student of the Laboratory of Applied Mathematics and Computer Science, Faculty of Science and Techniques, Cadi Ayyad University, Marrakesh, Morocco Said Raghay, Professor at Department of Mathematics, Laboratory of Applied Mathematics and Computer Science, Faculty of Science and Techniques, Cadi Ayyad University, Marrakesh, Morocco. Abdelilah Hakim, Professor at Department of Mathematics, Laboratory of Applied Mathematics and Computer Science, Faculty of Science and Techniques, Cadi Ayyad University, Marrakesh, Morocco.

Availability of data and materials
The set of images used to demonstrate the effectiveness of the proposed approach is composed of seven standard images used for image processing. These images are frequently found in literature and available on the following site: www.imageprocessingplace.com/root_files_V3/image_databases.htm.

Competing interests
The authors declare that they have no competing interests.

References
1. A. Mohammad-Djafari, Inverse Problems in imaging systems and the general Bayesian inversion framework. J. Iran. Assoc. Electr. Electron. Eng. **3**(2), 3–21 (2006)
2. T. F. Chan, C. K. Wong, Total variation blind deconvolution. IEEE Trans. Image Process. **7**(3), 370–375 (1998)
3. T. J. Holmes, Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach. J. Opt. Soc. Am. A. **9**(7), 1052–1061 (1992)
4. S. U. Pillai, B. Liang, Blind image deconvolution using a robust gcd approach. IEEE Trans. Image Process. **8**(2), 295–301 (1999)
5. A. Levin, Y. Weiss, F. Durand, W. T. Freeman, Understanding blind deconvolution algorithms. IEEE Trans. Pattern. Anal. Mach. Intell. **33**(12), 2354–2367 (2011)
6. H. Liao, K. MNg, Blind deconvolution using generalized cross-validation approach to regularization parameter estimation. IEEE Trans. Image Process. **20**(3), 670–680 (2011)
7. M. Rostami, O. Michailovich, Z. Wang, Image deblurring using derivative compressed sensing for optical imaging application. IEEE Trans. Image Process. **21**(7), 3139–3149 (2012)
8. F. Sroubek, P. Milanfar, Robust multichannel blind deconvolution via fast alternating minimization. IEEE Trans. Image Process. **21**(4), 1687–1700 (2012)
9. N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*. (MIT press, Cambridge, MA, 1950)
10. R. C. Gonzalez, R. E. Woods, *Digital image processing*. (Prentice hall, Upper Saddle River, 2002)
11. W. H. Richardson, Bayesian-based iterative method of image restoration. JOSA. **62**(1), 55–59 (1972)
12. L. B. Lucy, An iterative technique for the rectification of observed distributions. Astron. J. **79**(6), 745–754 (1974)
13. L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms. Phys. D nonlinear Phenom. **60**(1-4), 259–268 (1992)
14. L. I. Rudin, S. Osher, Total variation based image restoration with free local constraints. Proc. 1st Int. Conf. Image Process. IEEE. **1**, 31–35 (1994)
15. M. Elad, On the origin of the bilateral filter and ways to improve it. IEEE Trans. Image Process. **11**(10), 1141–1151 (2002)
16. S. Farsiu, M. D. Robinson, M. Elad, P. Milanfar, Fast and robust multiframe super resolution. IEEE Trans. Image Process. **13**(10), 1327–1344 (2004)
17. R. Molina, J. Mateos, A. K. Katsaggelos, Blind deconvolution using a variational approach to parameter, image, and blur estimation. IEEE Trans. Image Process. **15**(12), 3715–3727 (2006)

18.   S. U. Park, N. Dobigeon, A. O. Hero, Semi-blind sparse image reconstruction with application to MRFM. IEEE Trans. Image Process. **21**(9), 3838–3849 (2012)

19.   Z. Xu, E. Y. Lam, Maximum a posteriori blind image deconvolution with Huber–Markov random-field regularization. Opt. Lett. **34**(9), 1453–1455 (2009)

20.   S. D. Babacan, J. Wang, R. Molina, A. K. Katsaggelos, Bayesian blind deconvolution from differently exposed image pairs. IEEE Trans. Image Process. **19**(11), 2874–2888 (2010)

21.   L. Blanco, L. M. Mugnier, Marginal blind deconvolution of adaptive optics retinal images. Opt. Express. **19**(23), 7–23239 (2011)

22.   S. Yousefi, N. Kehtarnavaz, Y. Cao, Computationally tractable stochastic image modeling based on symmetric markov mesh random fields. IEEE Trans. Image Process. **22**(6), 2192–2206 (2013)

23.   A. G. Glen, On the inverse gamma as a survival distribution. J. Qual. Technol. **43**(2), 158–166 (2011)

24.   M. K. Ng, R. H. Chan, W. C. Tang, A fast algorithm for deblurring models with Neumann boundary conditions. SIAM J. Sci. Comput. **21**(3), 851–866 (1999)

25.   G. Aubert, P. Kornprobst, *Mathematical problems in image processing: partial differential equations and the calculus of variations*. (Springer Science & Business Media, New York, 2006)

26.   P. C. Hansen, D. P. O'Leary, The use of the L-curve in the regularization of discrete ill-posed problems. SIAM J. Sci. Comput. **14**(6), 1487–1503 (1993)

27.   G. H. Golub, M. Heath, G. Wahba, Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics. **21**(2), 215–223 (1979)

28.   A. H. Bentbib, M. El Guide, K. Jbilou, Matrix Krylov subspace methods for image restoration. New Trends Math. Sci. **3**(3), 136–148 (2015)

29.   G. H. Golub, C. van Loan, *Matrix computation, the third edition*. (The Johns Hopkins University Press, Baltimore, 1996)

30.   M. Bergounioux, *Introduction au traitement mathématique des images-méthodes déterministes*. (Springer, Berlin, 2015)

31.   P. C. Hansen, Regularization tools: a Matlab package for analysis and solution of discrete ill-posed problems. Numer. Algoritm. **6**(1), 1–35 (1994)

32.   Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)

33.   Y. W. Tai, P. Tan, M. S. Brown, Richardson-lucy deblurring for scenes under a projective motion path. IEEE Trans. Pattern. Anal. Mach. Intell. **33**(8), 1603–1618 (2010)

34.   M. S. Almeida, L. B. Almeida, Blind and semi-blind deblurring of natural images. IEEE Trans. Image Process. **19**(1), 36–52 (2009)

35.   E. Faramarzi, D. Rajan, M. P. Christensen, Unified blind method for multi-image super-resolution and single/multi-image blur deconvolution. IEEE Trans. Image Process. **22**(6), 2101–2114 (2013)

36.   M. Bergounioux, *Approche fréquentielle : Filtre de Wiener. Introduction au traitement mathématique des images-méthodes déterministes vol. 76*. (Springer, Berlin, 2015), pp. 87–90

Publisher's Note