



A non-local regularization strategy for image deconvolution

Max Mignotte*

DIRO, Univ. de Montréal, C.P. 6128, Succ. Centre-ville, Montréal, Canada H3C 3J7

ARTICLE INFO

Article history:

Received 6 September 2007
Received in revised form 26 July 2008
Available online 15 August 2008

Communicated by J.A. Robinson

Keywords:

Image deconvolution or restoration
Non-local regularization
Penalized likelihood
L-curve estimation

ABSTRACT

In this paper, we propose an inhomogeneous restoration (deconvolution) model under the Bayesian framework exploiting a non-parametric adaptive prior distribution derived from the appealing and natural image model recently proposed by Buades et al. [Buades, A., Coll, B., Morel, J.-M., 2005. A review of image denoising algorithms, with a new one. *SIAM Multiscale Model. Simul. (SIAM Interdisc. J.)*, 4(2), 490–530] for pure denoising applications. This prior expresses that acceptable restored solutions are likely the images exhibiting a high degree of redundancy. In other words, this prior will favor solutions (i.e., restored images) with similar pixel neighborhood configurations. In order to render this restoration unsupervised, we have adapted the L-curve approach (originally defined for Tikhonov-type regularizations), for estimating our regularization parameter. The experiments herein reported illustrate the potential of this approach and demonstrate that this regularized restoration strategy performs competitively compared to the best existing state-of-the-art methods employing classical local priors (or regularization terms) in benchmark tests.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

In the regularized restoration framework, the prior (or regularization) term allows us to both statistically incorporate knowledge concerning the types of restored images *a priori* defined as acceptable solutions and to stabilize (computationally speaking) the solution of this ill-conditioning inverse problem. That is why the design of efficient image prior models, and especially their ability to summarize the intrinsic properties of an original image to be recovered are crucial in the final image quality and signal-to-noise improvement (ISNR) restoration result.

Over the last two decades, there have been considerable efforts to find an efficient regularization term (or a prior distribution) capable of modeling all the intrinsic properties of a natural image, particularly its edge and textural information. To this end, several edge-preserving local regularization strategies there have been proposed in the spatial domain (Chantas et al., 2006; Foi et al., 2006; Mignotte, 2006; Neelamani et al., 2004; Banham and Katsaggelos, 1996) (e.g., *via* non-stationary, compound Markov or MRF model with robust potential functions) or in the frequential domain (Guerrero-Colon and Portilla, 2006; Bioucas-Dias et al., 2006; Foi et al., 2006; Figueiredo and Nowak, 2005, 2003; Bioucas-Dias, 2006; Neelamani et al., 2004; Banham and Katsaggelos, 1996) (e.g., *via* thresholding).

Buades et al. (2005) have recently proposed a natural and elegant extension of the image bilateral filtering paradigm. The basic idea behind the so-called *Non-Local means* (NL-means) denoising concept is simple. For a given pixel i , its new (denoised) intensity value is computed as a weighted average of grey level values within a *search window*. The weight of the pixel j in this weighted average is proportional to the similarity (according to the euclidean distance) between the neighborhood configurations of pixels i and j . In this procedure, the denoising process is due to the regularity assumption that self-similarities of neighborhoods exist in a real image¹ and that one (or several) neighborhood configuration(s) can efficiently predict the central value of the pixel, as shown by Efros and Leung (1999) for texture synthesis with a (somewhat) similar non-parametric sampling strategy.

In this paper, the idea proposed by Buades et al. (2005) is herein used to derive an efficient image prior distribution. This prior expresses that acceptable restored images are likely the solutions exhibiting similar neighborhood configurations (i.e., images owing a high degree of redundancy or exhibiting numerous similar patterns¹). Comparisons with classical deconvolution and restoration approaches using local regularization strategy (in the spatial or frequential domain) are given in order to illustrate the potential of this approach and its pros and cons for some degradation models.

¹ The repetitive character of a textured image (sometimes called the *texton*) is obvious and one can observe that flat zones present numerous similar configurations lying in the same object. Straight or curved edges also generally possess a complete line of pixels with similar configurations (mostly along the contour, see Buades et al. (2005) for relevant examples).

* Tel.: +1 514 343 5747; fax: +1 514 343 5834.
E-mail address: mignotte@iro.umontreal.ca
URL: <http://www.iro.umontreal.ca/~mignotte/>

2. Proposed approach

We herein use the classical *penalized likelihood* framework leading, in the context of image restoration, to the following cost function $E(x)$ to be optimized

$$\hat{x} = \arg \min_x \underbrace{\{\|y - h * x\|^2 + \gamma \Omega(x)\}}_{E(x)} \quad (1)$$

where y and x represent, respectively, the noisy (degraded by an additive and white Gaussian noise with variance σ^2) and observed blurred image and the undistorted true image. h is the Point Spread Function (PSF) of the imaging system² and $*$ is the linear convolution operator. This energy function ($E(x)$) contains two terms, the first expresses the fidelity to the available data y and the second encodes the expected property of the true undegraded image. γ is the regularization parameter controlling the contribution of the two terms. Let us note that this last parameter is crucial as it determines the overall quality of the final estimate. More precisely, γ must be small in order to achieve a small residual error $\|y - h * x\|^2$. However, a large γ is required to obtain a restored image x_γ with small prior error $\Omega(x)$. The estimation of this parameter will be discussed in Section 3.

2.1. Non-local prior

In this work, we consider a prior error $\Omega(x)$, related to a prior distribution of the form

$$P_x(x) \propto \exp\{-\gamma \Omega(x)\} \propto \exp\left\{-\gamma \underbrace{\|x - \mathcal{N}\mathcal{L}\mathcal{M}_{|y|}(x)\|_q^q}_{\|\rho(x)\|_q^q}\right\} \quad (2)$$

where $q > 1$ is a shape factor and $\mathcal{N}\mathcal{L}\mathcal{M}_{|y|}(x)$ designates the NL-means (non-local means) filter (Buades et al., 2005) applied on x (the undegraded image to be estimated).

More precisely, when given a discrete noisy and blurred image $y = \{y(i) | i \in I\}$, the estimated value $\mathcal{N}\mathcal{L}\mathcal{M}_{|y|}(x(i))$, for a pixel i of x is computed as a weighted average of all the pixels in the image by the following non-local filtering process:

$$\mathcal{N}\mathcal{L}\mathcal{M}_{|y|}(x(i)) = \sum_{j \in I} w(i, j) x(j)$$

In this filter, the family of weights $\{w(i, j)\}_j$ depend on the similarity (according to the euclidean distance) between the neighborhoods of pixels i and j , and must satisfy (Buades et al., 2005) the usual conditions $0 \leq w(i, j) \leq 1$ and $\sum_j w(i, j) = 1$ with

$$w(i, j) = \frac{1}{Z(i)} \exp\left\{-\frac{\|y(\mathcal{N}_i) - y(\mathcal{N}_j)\|_2^2}{h^2}\right\}$$

$y(\mathcal{N}_i)$ denotes the intensity grey level vector constituted by the set of grey level values located in the square neighborhood (\mathcal{N}_i) of fixed size S and centered at pixel of coordinate i . $Z(i)$ is the normalizing constant given by

$$Z(i) = \sum_j \exp\left\{-\frac{\|y(\mathcal{N}_i) - y(\mathcal{N}_j)\|_2^2}{h^2}\right\}$$

where the parameter h acts as a degree of filtering.

Thus $\mathcal{N}\mathcal{L}\mathcal{M}_{|y|}(x)$ designates the NL-means filter applied on x whose the non-local similarity graph (i.e., the family of weights $\{w(i, j)\}_j$) is previously computed from y the noisy and blurred observed image.

The proposed distribution (Eq. (2)) expresses that, for our deconvolution or restoration problem, an acceptable *a priori* solution is likely the set of estimated images invariant by NL-means denoising, i.e., an already denoised image exhibiting numerous similarity of neighborhood configuration (exhibiting an high degree of redundancy) as any natural images. The non-local similarity graph (i.e., the set of weights used in the NL-means filter) is computed from the only observable, i.e., y .

Eventually, the final result of restoration can be used to re-estimate a better non-local similarity graph and to refine, in a second pass, the restoration result. This regularization prior is described as *non-local* since pixels belonging to the whole image are used for the estimation of each new (denoised) pixel (in practice, in order to reduce the computation time, the seek of the neighborhood is limited to a window search around the pixel to be estimated).

2.2. Search of the solution

In our application, this search is performed by a steepest descent procedure which moves the estimates iteratively in the negative gradient direction $\hat{x}^{[n+1]} = \hat{x}^{[n]} - \alpha_n \nabla E(x)$, with the following iterative procedure:

$$\hat{x}^{[n+1]} = \hat{x}^{[n]} + \alpha_n (h^\# * (y - h * \hat{x}^{[n]})) - \gamma \cdot q \rho'(x^{[n]}) |\rho(x^{[n]})|^{q-1} \quad (3)$$

where $h^\#(i, j) = h(-i, -j)$ (the coordinates (i, j) represent the discrete pixel locations and for h symmetric, we have $h^\# = h$). In this form of notation, the multiplication between ρ and ρ' is done point-by-point (or pixel by pixel) and the superscript indicates the iteration number. We recall that q is the shape factor of our prior error term (see Eq. (2)). α_n is the gradient step size which changes adaptively at each iteration according to the following equations (Sullivan and Chang, 1991):

$$\alpha_n = \frac{\|q_n\|^2}{\|h * q_n\|^2} \quad \text{with } q_n = h^\# * (y - h * x^{[n]}) \quad (4)$$

where, in this notation, pixels are organized in q_n and in $h * q_n$ in lexicographic order as one large column-vector and $\rho'(x_s)$ can be easily found and written by

$$\rho'(x_s) = \left(1 - \frac{1}{Z_s}\right) \cdot \text{sign}(x_s - \mathcal{N}\mathcal{L}\mathcal{M}_{|y|}(x_s)) \quad (5)$$

where Z_s is the normalizing constant (Buades et al., 2005) of the NL-means filter for each site (or pixel of coordinate) s . Customarily, $Z_s \gg 1$, we have herein considered that $\rho'(x_s) \approx \text{sign}(x_s - \mathcal{N}\mathcal{L}\mathcal{M}_{|y|}(x_s))$.

3. Regularization parameter estimation

A crucial element in this penalized likelihood framework as expressed by Eq. (1), is the proper choice of the regularization parameter γ . If γ is selected as small, the recovered image is dominated by high-frequency noise components (the solution is the so-called *under-regularized*). If γ is too large, the effect of the prior will dominate the solution and important information in the data will be lost (leading to a well-known *over-regularized* estimated image).

Several methods have been presented for estimating the regularization parameter and Thompson et al., 1991 review some of them in the context of a particular simple class of (energy-based) restoration models in which $\Omega(\cdot)$ in Eq. (1) is quadratic in x . It is the case of the so-called total predicted mean squared-error (TPME), the generalized cross-validation (GCV) or the equivalent degrees of freedom criterions (EDF) whose reliable estimation is facilitated with the help of eigen-analysis and spectral representation (Thompson et al., 1991). These simple methods cannot be applied in our energy-based restoration models involving a

² We shall assume throughout this paper that the degradation model (PSF and variance of the white Gaussian noise) is known. It might be given analytically or given numerically based on previous estimations or calibration experiments.

non-quadratic regularization functional. In the (more general) case of a non-quadratic regularizer, several computationally demanding stochastic methods have been proposed to compute the regularization parameter (or the so-called hyper-parameter) in the Maximum Likelihood (ML) sense. Amongst these, we can cite the stochastic descent algorithm proposed by Younes (1989) (which cannot be applied in the case of blur degradation) or the mean field approximation method of Zhou et al. (1997), or a gradient descent algorithm using a Markov Chain Monte Carlo sampling technique (to sample from prior and posterior distributions) as proposed by Jalobeanu et al. (2002). We can also cite, in the Stein's risk sense, the hyper-parameter stochastic estimation method proposed by Batu and etin (2008) which requires a stochastic estimator of the trace of a large matrix.

Another technique, called the L-curve (Hansen, 1992) has been proposed in the case of the Tikhonov regularization (i.e., when $\Omega(x_\gamma) = \|Lx_\gamma\|^2$ where L is either an identity matrix, a gradient or a Laplacian operator for respectively a zero, first- or second-order regularization method). The L-curve is simply a logarithmic plot of $\mathcal{L}(y; x_\gamma) = \|y - h * x_\gamma\|^2$ (called the residual norm or the likelihood energy term) versus the logarithmic plot of $\Omega(x_\gamma)$ (the prior or regularization energy term) as γ , the regularization parameter, varied. The name L-curve derives from the characteristic shape of this curve. When γ is very large, the curve is essentially an horizontal line; the restored image is excessively smooth (*over-regularized*) and the final estimate is dominated by *regularization errors*. When γ is very small, the curve is essentially a vertical line; in this case, we have a ML restoration (without prior) and thus the error between the undegraded and recovered (under-regularized) image is dominated by *perturbation errors*. The transition between these two regions of under and over-regularization, separated by the “corner” point, has been proposed in (Hansen, 1992) as an optimal value of the regularization parameter where regularization and perturbation errors (or bias and variance of the final estimate) are approximately balanced. The L-curve estimation technique remains computationally very demanding since it requires the computation of many repeated MAP restored images for a wide range of values of the regularization parameter. Once these points have been computed, a certain interpolation method is used to obtain a rational function (approximating the L-curve). The curvature of this rational function is then computed to determine the best value of the regularization parameter. In the case of our *non-local* regularization framework, this estimation technique is much simpler (deterministic and faster) to compute since the obtained curve ($\log \mathcal{L}(y; x_\gamma)$, $\log \Omega(x_\gamma)$) has a V-shape (see Fig. 1). More precisely, our modeling leads to a curve where part of the curve to the right of the corner is not horizontal but a nearly vertical diagonal line. This characteristic results from the specific nature of our non-local prior which contrary to Tikhonov-type regularizations, does not smooth the image but gradually increases the bias (or the regularization error) of the estimate (by creating high frequency components contained in some textons) when γ gradually increases. In our case, the corner can easily be searched by a simple steepest descent procedure which starts with $\gamma^{[0]} = \sigma/12$ and $\gamma^{[1]} = \sigma/8$ (for example) and then moves γ iteratively in the negative gradient direction with a fixed step size by the following iterative procedure:

$$\hat{\gamma}^{[q+2]} = \hat{\gamma}^{[q]} - \beta \text{sign} \left(\frac{1}{\frac{\log \Omega(x_{\gamma^{[q+1]}}) - \log \Omega(x_{\gamma^{[q]}})}{\log \mathcal{L}(y; x_{\gamma^{[q+1]}}) - \log \mathcal{L}(y; x_{\gamma^{[q]}})}} \right) \quad (6)$$

$$\gamma^{[q+1]} = \gamma^{[q]}$$

(with $\mathcal{L}(y; x_\gamma) = \|y - h * x_\gamma\|^2$) where the term under the fraction is the numerical approximation (first-order) of the derivative of the curve at $\gamma^{[q]}$ and $\text{sign}(\cdot)$ is the sign function. We recall that $x_{\gamma^{[q]}}$

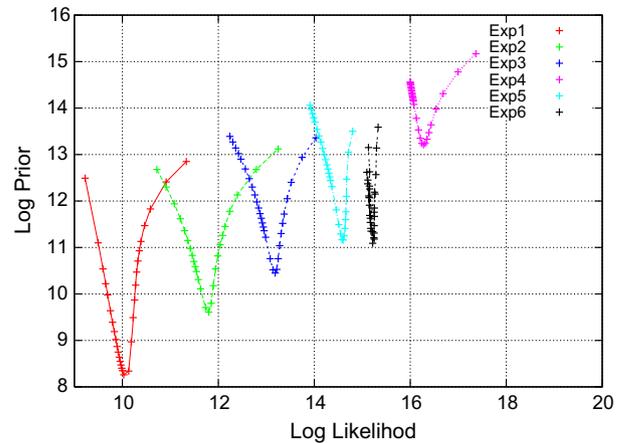


Fig. 1. Examples of L-curves obtained in our non-local restoration model for the different degradation models considered in Table 1. Each point of each curve is obtained for a value of the regularization parameter γ ($\gamma \in [0.0-10.0]$).

designates the restored image at convergence (we stop the iterative procedure (3) after stability of the restored image), for regularization parameter equals to $\gamma^{[q]}$ obtained at iteration q . We have taken $\beta = 0.1$ and we stop the procedure when the numerical approximation of the derivative changes of sign. Some authors (Lie, 2005; Xu, 1998; Ng and Allebach, 2006) have recently observed that the L-curve estimation technique always slightly over-smoothes the solution (i.e., it selects a slightly large γ). In order to take this characteristic (that we have observed in our application) into account, we have decided to weight the final estimated $\hat{\gamma}$ by 1/2.5.

4. Experimental results

4.1. Set-up

In all experiments, we have considered the NL-means algorithm with the following parameters: the size of the search window and the neighborhood (S) is set to 7×7 . The decay of the weights in the similarity measure is set to $h = 10\sigma$ (as proposed in Buades et al. (2005)) where σ is the standard deviation of the Gaussian noise and we have considered a classical Euclidean distance (and not a Gaussian weighted Euclidean distance as proposed in (Buades et al., 2005)). We precompute the set of weights in order to decrease the computational requirement of the non-local denoising procedure, herein applied to each iteration of the steepest descent iteration and take a shape factor $q = 1.5$. We start the iterative gradient descent (Eq. (3)) with $x^{[0]} = y$ and we stop the iterative procedure after stability of the restored image.

We now present a set of experimental results and comparisons illustrating the performance of the proposed approach. For the first four experiments, we have replicated the scenarios used in the evaluation of state-of-the-art methods described in (Guerrero-Colon and Portilla, 2006; Bioucas-Dias et al., 2006; Chantas et al., 2006; Foi et al., 2006; Mignotte, 2006; Figueiredo and Nowak, 2005, 2003; Bioucas-Dias, 2006; Neelamani et al., 2004; Banham and Katsaggelos, 1996; Jalobeanu et al., 2001; Liu and Moulin, 1998), with which we compare the proposed approach. We have also replicated the degradation model described in (May et al., 1998; Molina et al., 2000). In these cases, two edge-preserving methods were implemented and tested; respectively; (1) The use of a compound Gauss-Markov random fields, with an Ising model representing the upper level and a line process to model the abrupt transitions (and acting as an activator or inhibitor of the relation between two neighbor pixels). In this model, the solution is

Table 1

Blur, noise variance and BSNR (dB) for experiments Exp1–6

	Blur	σ^2	BSNR
Exp1	9×9 uniform [CAMERAMAN 256×256]	.308	40
Exp2	$h_{ij} = (1 + i^2 + j^2)^{-1}, i, j = -7, \dots, 7$ [CAMERAMAN 256×256]	2	32
Exp3	$h_{ij} = (1 + i^2 + j^2)^{-1}, i, j = -7, \dots, 7$ [CAMERAMAN 256×256]	8	26
Exp4	$[1, 4, 6, 4, 1]^t [1, 4, 6, 4, 1]/256$ [LENA 512×512]	49	16.5
Exp5	5×5 uniform [CAMERAMAN 256×256]	33.3	20
Exp6	$\infty[1 + i^2 + j^2/16]^{-3}, i, j = -9, \dots, 9$ [CAMERAMAN 256×256]	62.5	17

estimated thanks to an extension of the classical simulated annealing. (2) The ARTUR model of Charbonnier et al. (1997) also implemented and tested in (Molina et al., 2000) in the restoration context. In these experiments, original images are CAMERAMAN (experiments 1, 2, 3, 5 and 6) of size 256×256 and LENA (experiment 4) (of size 512×512). Table 1 displays blur, noise and the resulting BSNR (the ratio between the variance of the noise and the variance of blurred image without noise) for each of the experiments.

In order to also test the proposed estimation approach, we have thus considered, for comparisons

- (1) Algorithm *Sup. N-L Rest*: the proposed “supervised” non-local restoration approach with a regularization parameter manually tuned (by trials and errors) to give the best restoration result in the ISNR sense. To this end, we have set $\gamma = 0.04, 0.18, 0.30, 1.2, 0.48, 0.53$, respectively, for Exp1–6.
- (2) Algorithm *Unsup. N-L Rest*: the non-local restoration approach combined with the regularization parameter estimation proposed in Section 3.
- (3) Algorithm *Unsup. N-L Rest2*: the Unsup. N-L Rest. algorithm whose final result of restoration is used to reestimate a better non-local similarity graph (i.e., the family of weights $\{w(i, j)\}$) followed by a non-local restoration procedure (Sup. N-L Rest) with a regularization parameter γ fixed to the value estimated by the first step. In this second step, the neighborhood size S is set to one or three pixels. This choice will be made explicit in the following (see Section 4.2).

The obtained result is compared to the existing state-of-the-art algorithms in Tables 2 and 3. The best ISNR results provided by the existing restoration algorithms and the results provided by our approach for each degradation level are indicated in bold. In order to compare the quality of the restoration result in (Molina et al.,

Table 2

ISNR (dB) for experiments Exp1–4

Methods	ISNR (dB)			
	Exp1	Exp2	Exp3	Exp4
<i>Sup. N-L Rest.</i>	7.34	6.44	4.82	3.83
<i>Unsup. N-L Rest.</i>	7.08	6.39	4.80	3.82
<i>Unsup. N-L Rest2.</i>	7.81	7.14	5.24	3.84
Guerrero-Colon and Portilla (2006)	7.33	7.45	5.55	–
Bioucas-Dias et al. (2006)	8.52	–	–	2.97
Chantas et al. (2006)	8.91	–	–	3.77
Foi et al. (2006)	8.58	8.29	6.34	4.55
Mignotte (2006)	8.23	7.58	5.70	1.63
Figueiredo and Nowak (2005)	8.16	7.46	5.24	2.84
Bioucas-Dias (2006)	8.10	7.40	5.15	2.85
Figueiredo and Nowak (2003)	7.59	6.93	4.88	2.94
Neelamani et al. (2004)	7.30	–	–	–
Banham and Katsaggelos (1996)	6.70	–	–	–
Jalobeanu et al. (2002)	–	6.75	4.85	–
Liu and Moulin (1998)	–	–	–	1.08

Table 3

ISNR/PSNR (dB) for experiments Exp5–6

Methods	ISNR (dB)	
	Exp5	Exp6
Sup. N-L Rest.	3.81	2.77
Unsup. N-L Rest.	3.80	2.64
Unsup. N-L Rest2.	4.24	2.99
Mignotte (2006)	3.50	1.9
May et al. (1998)	3.43	–
Molina et al. (2000)	–	0.17 (PSNR = 21.1)
Charbonnier et al. (1997) (in Molina et al. (2000))	–	0.16 (PSNR = 20.8)

2000) which use the peak signal-to-noise ratio (PSNR) (defined in dB between two images x and \hat{x} of size N pixels by $10\log_{10}([N \times 255^2]/\|x - \hat{x}\|^2)$), we use the following conversion $ISNR = PSNR\|y - x\|^2/(N \times 255^2)$, i.e., $ISNR = PSNR \cdot MSE(y - x)/255^2$ where $MSE(\cdot)$ represents the average mean square error of the considered degradation model.

Table 4 shows the time in seconds and the number of iterations taken by each restoration, for each considered degradation model (cf. Table 1) (system used: AMD Athlon 64 Processor 3500+, 2.2 GHz, 4424 bogomips and running on Linux).

4.2. Discussion

In order to test the limit of our prior, we experimented with use of the original (undegraded) image to find the “optimal” weights which will then be used in our restoration procedure. Table 5 summarizes the ISNR results for the different experiments (and for different neighborhood sizes, γ being empirically set to an optimal value). The table shows that when the neighborhood size is small,

Table 4

Time in seconds for the different experiments

	Time (s) (Nb. iterations)		
	Sup. N-L Rest	Unsup. N-L Rest	Unsup. N-L Rest 2
Exp1	281 ₍₉₇₈₎	407	635
Exp2	35 ₍₁₀₁₎	129	161
Exp3	27 ₍₇₅₎	78	103
Exp4	20 ₍₇₎	72	86
Exp5	19 ₍₉₀₎	33	50
Exp6	314 ₍₈₂₎	487	765

Table 5

ISNR (dB) obtained for optimal weights estimated on the original (undegraded image) for experiments Exp1–6

Sup. N-L Rest.	ISNR (dB)					
	Exp1	Exp2	Exp3	Exp4	Exp5	Exp6
Window size $[1 \times 1]$	16.99	11.50	8.17	4.09	6.53	5.10
Window size $[3 \times 3]$	8.01	7.35	6.07	4.44	6.03	5.09

very good ISNR results are possible. Restoration results can excel when neighborhood size equals to one pixel. In the latter case, the family of weights thus encodes the similarity between the true (undegraded) luminances of each pair of pixels within a search window. However, this experience is anecdotal, since we do not know the original image (and we also need a neighborhood size sufficiently large enough to ensure a reliable estimation of the weight values against the noise). It shows that improvement could be obtained if a better similarity non-local graph was estimated (i.e., if a more reliable estimation procedure was used to assess the set of weights). In order to take the preceding remark into account, we have tested the strategy using the final result of restoration to reestimate a better non-local similarity graph which will be used in a second pass to refine the restoration result. The experiments have yielded several interesting results. If the non-local graph is reestimated with the same original neighborhood size (i.e., 7×7 pixels), this latter strategy does not emerge as an efficient means of improving the restoration result. However, this strategy turns out to be very efficient if the neighborhood size is smaller than the one used in the first step of the restoration algorithm. This remark also confirms the aforementioned experience (using the true undegraded image to find the “optimal” weights). Since the restoration result of *Sup. N-L Rest* or *Unsup. N-L Rest* are already widely denoised and closer to the true (undegraded) image than the degraded image, the neighborhood size in the second step has to be smaller. For *Unsup. N-L Rest2* we use one pixel for the

neighborhood size when the variance of the noise (σ^2) is below 15, otherwise we use three pixels for the neighborhood size.

We can observe that the proposed non-local restoration method leads to interesting and sometimes competitive restoration results for various level of blur and noise degradations in benchmark tests, especially for degradation exhibiting more noise than blur and even for blur expressed by a point spread functions exhibiting zeros in the frequential domain (such as the uniform blur and for which the distance used in the similarity graph could be altered). Figs. 2 and 3 show some restoration results for Exp1 and Exp4.

Our NL-means-based regularization term (encoding the inherent redundancy property of any textured images) seems particularly efficient for LENA image (Exp4) which contains several textures (thus exhibiting numerous similar neighborhood configurations or repetitive patches on each textured area). In this case the ISNR obtained by our restoration method shows very good result comparatively to the other algorithms. In the case of the ISNR restoration results related to Exp1–3 and Exp5–6 on the CAMERAMAN image (which is less textured and has relatively more piecewise homogeneous regions or “geometric structures”), our model remains competitive. However, a segmentation-based regularization term Mignotte (2006) or any regularization term, promoting a piecewise smooth restored image, seems more appropriate.

We have tested the influence of the variation of the parameters q (the norm of the NL-means prior) and the “size of the neighborhood” on the result of the SNR improvement measure (for all the



Fig. 2. From left to right, original image, noisy-blurred image for Exp1 (see Table 1) and restored image using the proposed restoration approach (algorithm Unsup. N-L Rest2.) ISNR = 7.79 dB (see Table 2).



Fig. 3. From left to right, original (cropped) Lena image noisy-blurred image for Exp4 (see Table 1) and restored image using the proposed approach (algorithm Unsup. N-L Rest2.) ISNR = 3.79 dB (see Table 2).

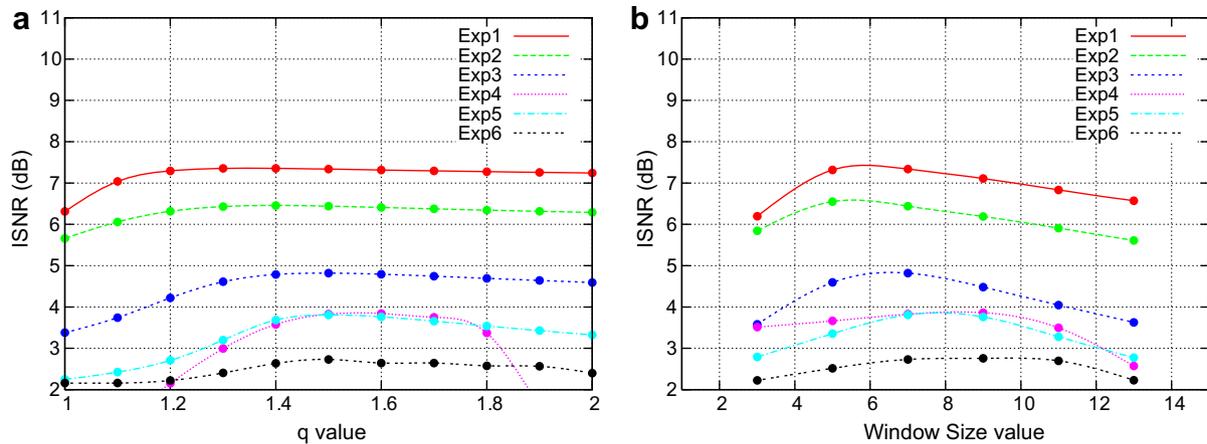


Fig. 4. Evolution of the SNR improvement, for the different experiments (algorithm Sup. N-L Rest.), along the value of the parameter: (a) q (the “neighborhood size” parameter and γ being set to their optimal value for each experiment) and (b) neighborhood size (γ and q being set to their optimal value for each experiment).

experiments and for our algorithm Sup. N-L Rest). Fig. 4 shows the evolution of the ISNR improvement along several discrete values of q ($q \in [1, \dots, 2]$) and “neighborhood size” $\in [3, \dots, 13]$ pixels. Experiments show that our proposed restoration model is not overly sensitive if $q \in [1.3, \dots, 1.8]$ and if the “neighborhood size” parameter (used in the first pass of our restoration algorithm and thus directly on the degraded image) is between 5×5 and 9×9 .

We can also observe that the estimation procedure presented for the regularization parameter is particularly well suited to this model. Starting from the initial value given in Section 3, only 3–6 iterations are necessary to converge and to produce reliable and nearly optimal value for this parameter.

5. Conclusion

In this paper, we have presented a deconvolution/restoration approach whose regularization term encodes the inherent high redundancy of any natural images. This new prior derived from the denoising algorithm proposed by Buades et al. allows to efficiently constrain a deconvolution procedure, demonstrating its ability to summarize the intrinsic redundancy property of any natural image. In this context, the L-curve based approach proposed by Hansen et al. is well suited to a robust, fast, deterministic and easy estimation of the optimal value of the regularization parameter. Finally, we believe that this adaptive regularization strategy could also be efficiently extended in order to regularize a number of inverse problems in image processing or computer vision such as tomography, superresolution, segmentation or reconstruction problems.

Acknowledgements

The author is grateful to all the anonymous reviewers for their many valuable comments and suggestions that helped to improve this paper. In particular, he acknowledges the contribution of the reviewer who suggested we reestimate the weights of the non-local graph from the restored image in order to refine the restoration result. This encouraged the author to do more tests which have come to the comments given in Section 4.2 and finally improved the restoration results.

References

Banham, M.R., Katsaggelos, A.K., 1996. Spatially adaptive wavelet-based multiscale image restoration. *IEEE Trans. Image Process.* 5 (4), 619–634.

- Batu, O., Etin, M., 2008. Hyper-parameter selection in advanced synthetic aperture radar imaging algorithms. In: *IEEE Conf. on Signal Processing and Communications Applications*, Aydin, Turkey.
- Bioucas-Dias, J., 2006. Bayesian wavelet-based image deconvolution: A GEM algorithm exploiting a class of heavy-tailed priors. *IEEE Trans. Image Process.* 15, 937–951.
- Bioucas-Dias, J., Figueiredo, M., Oliveira, J., 2006. Adaptive total-variation image deconvolution: A majorization-minimization approach. In: *Proc. EUSIPCO'2006*.
- Buades, A., Coll, B., Morel, J.-M., 2005. A review of image denoising algorithms, with a new one. *SIAM Multiscale Model. Simulat. (SIAM Interdiscip. J.)* 4 (2), 490–530.
- Chantas, G., Galatsanos, N., Likas, A., 2006. Bayesian restoration using a new nonstationary edge-preserving image prior. *IEEE Trans. Image Process.* 15 (10), 2987–2997.
- Charbonnier, P., Blanc-Feraud, L., Aubert, G., Barlaud, M., 1997. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Process.* 5 (12), 298–311.
- Efros, A.A., Leung, T.K., 1999. Texture synthesis by non-parametric sampling. In: *7th Internat. Conf. on Computer Vision, ICCV'99*, Kerkyra, Greece, pp. 1033–1038.
- Figueiredo, M.A.T., Nowak, R.D., 2003. An EM algorithm for wavelet-based image restoration. *IEEE Trans. Image Process.* 12, 906–916.
- Figueiredo, M., Nowak, R., 2005. A bound optimization approach to wavelet-based image deconvolution. In: *IEEE Internat. Conf. on Image Processing - ICIP'05*, vol. II, pp. 782–785.
- Foi, A., Dabov, K., Katkovnik, V., Egiazarian, K., 2006. Shape-adaptive DCT for denoising and image reconstruction. In: *Proc. SPIE Electronic Imaging 2006, Image Processing: Algorithms and Systems V*, vol. 6064A-18.
- Guerrero-Colon, J.A., Portilla, J., 2006. Deblurring-by-denoising using spatially adaptive gaussian scale mixtures in overcomplete pyramids. In: *IEEE Internat. Conf. on Image Processing, ICIP'06*, vol. I, pp. 625–628.
- Hansen, P., 1992. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM J. Sci. Comput.* 14 (6), 1487–1503.
- Jalobeau, A., Blanc-Feraud, L., Zerubia, J., 2002. Hyperparameter estimation for satellite image restoration using a MCMC maximum likelihood method. *Pattern Recognition* 32 (2), 341–352.
- Jalobeau, A., Kingsbury, N., Zerubia, J., 2001. Image deconvolution using hidden Markov tree modeling of complex wavelet packets. In: *IEEE Internat. Conf. on Image Processing, ICIP-2001*.
- Lie, J., 2005. Inverse problems and regularization methods. University of Bergen, Technical report. Available from: <<http://www.uib.no/People/nmaji/papers/Geilo-Eassay-Inverse-Problems-Johan.pdf>>.
- Liu, J., Moulin, P., 1998. Complexity-regularized image restoration. In: *IEEE Internat. Conf. on Image Processing, ICIP'98*, pp. 555–559.
- May, K., Stathaki, T., Constantinides, A.G., Katsaggelos, A.K., 1998. Iterative determination of local bound constraints in iterative image restoration. In: *IEEE Internat. Conf. on Image Processing, ICIP'98*, vol. 2, pp. 833–836.
- Mignotte, M., 2006. A segmentation-based regularization term for image deconvolution. *IEEE Trans. Image Process.* 15 (7), 1973–1984.
- Molina, R., Katsaggelos, A.K., Mateos, J., Hermoso, A., Segall, C.A., 2000. Restoration of severely blurred high range images using stochastic and deterministic relaxation algorithms in compound Gauss–Markov random fields. *Pattern Recognition* 33 (4), 555–571.
- Neelamani, R., Choi, H., Baraniuk, R., 2004. Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems. *IEEE Trans. Signal Process.* 52 (2), 418–433.
- Ng, D., Allebach, J.P., 2006. A subspace matching color filter design methodology for a multispectral imaging system. *IEEE Trans. Image Process.* 15 (9), 2631–2643.

- Sullivan, B.J., Chang, H.-C., 1991. A generalized landweber iteration for ill-conditioned signal restoration. In: IEEE Internat. Conf. on Acoustics, Speech, and Signal Processing, ICASSP'91, pp. 1729–1732.
- Thompson, A.M., Brown, J., Kay, J., Titterton, D.M., 1991. A study of methods of choosing the smoothing parameter in image restoration by regularization. IEEE Trans. Pattern Anal. Machine Intell. 13 (4), 326–339.
- Xu, P., 1998. Truncated SVD method for discrete ill-posed problems. Geophys. J. Internat. 135, 505–514.
- Younes, L., 1989. Parametric inference for imperfectly observed Gibbsian fields. Prob. Th. Fields, vol. 82. Springer-Verlag, pp. 625–645.
- Zhou, Z., Leahy, R.M., Qi, J., 1997. Approximate Maximum Likelihood estimation for Gibbs priors. IEEE Trans. Image Process. 6 (6), 844–861.