Semi-Huber quadratic function and comparative study of some MRFs for Bayesian image restoration

J. I. de la Rosa
ismaelrv@ieee.org

Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Antiguo Camino a la Bufa No. 1, Col. Centro. C. P. 98000, Zacatecas, Zac., México

J. Villa-Hernández

Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Antiguo Camino a la Bufa No. 1, Col. Centro. C. P. 98000, Zacatecas, Zac., México

E. González-Ramírez

Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Antiguo Camino a la Bufa No. 1, Col. Centro. C. P. 98000, Zacatecas, Zac., México

E. de la Rosa M.

Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Antiguo Camino a la Bufa No. 1, Col. Centro. C. P. 98000, Zacatecas, Zac., México

O. Gutiérrez

Instituto Tecnológico Superior de Fresnillo, Av. Tecnológico No. 2000 Col. Solidaridad, C.P. 99010 Fresnillo, Zac., México

C. Olvera-Olvera

Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Antiguo Camino a la Bufa No. 1, Col. Centro. C. P. 98000, Zacatecas, Zac., México

R. Castañeda-Miranda

Unidad Académica de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas, Antiguo Camino a la Bufa No. 1, Col. Centro. C. P. 98000, Zacatecas, Zac., México

G. Fleury

École Supérieure d'Électricité (SUPELEC), Plateau de Moulon, 3 rue Joliot Curie, 91192 Gif-sur-Yvette Cedex, France

The present work introduces an alternative method to deal with digital image restoration into a Bayesian framework, particularly, the use of a new half-quadratic function is proposed which performance is satisfactory compared with respect to some other functions in existing literature. The bayesian methodology is based on the prior knowledge of some information that allows an efficient modelling of the image acquisition process. The edge preservation of objects into the image while smoothing noise is necessary in an adequate model. Thus, we use a convexity criteria given by a semi-Huber function to obtain adequate weighting of the cost functions (half-quadratic) to be minimized. The principal objective when using Bayesian methods based on the Markov Random Fields (MRF) in the context of image processing is to eliminate those effects caused by the excessive smoothness on the reconstruction process of image which are rich in contours or edges. A comparison between the new introduced scheme and other three existing schemes, for the cases of noise filtering and image deblurring, is presented. This collection of implemented methods is inspired of course on the use of MRFs such as the semi-Huber, the generalized Gaussian, the Welch, and Tukey potential functions with granularity control. The obtained results showed a satisfactory performance and the effectiveness of the proposed estimator with respect to other three estimators.
[DOI: http://dx.doi.org/10.2971/jeos.2013.13072]

Keywords: Digital image processing, image filtering, image deblurring, Markov random fields (MRF), half-quadratic functions

1 INTRODUCTION

The use of powerful methods proposed in the seventies under the name of Bayesian estimation [1]–[3], are nowadays essential at least in the cases of image filtering, segmentation and restoration (e.g. image deblurring) [4]. The basic idea of these methods is to construct a Maximum a posteriori (MAP) estimate of the modes or so called estimator of true images by using Markov Random Fields (MRF) in a Bayesian framework. The idea is based on a robust scheme which could be adapted to reject outliers, tackling situations where noise is present in different forms during the acquisition process [5]–[12]. Some image analysis and processing tasks involve the filtering or image recovery $\hat{x}$ (restoration) after it passes by a degradation process giving the observed image y (see Eqs. (1) and (3)). Such degradation covers the noise perturbation, blurring effects due to focus of the adquisition systems lenses or to the

bandwidth capacity, and other artifacts that may be added to the correct data. The image restoration or recuperation approaches of an image to its original condition given a degraded image, pass by reverting the effects caused by a distortion function. In fact, the degradation characteristics given by $F(x)$ and n in Eq. (1) are crucial information and they must be known or estimated during the inversion procedure. Typically, $F(x)$ is related to a point spread function H which can be linked with the probability distribution of the noise contamination n. In the case of MAP filters, usually the additive Gaussian noise is considered. A global image formulation model could be:

$$y = F(x) + n, \tag{1}$$

where, $F(x)$ is a functional that could take for instance, two forms: $F(x) = x$ and $F(x) = Hx$, being H a linear opera-

tor which models the image degradation. All variables presented along the text are, $x = (x_s)_{s \in S}$, where S is a set of pixels: which represents a realization of a Markov random field X (or image to be estimated), y: represents the observed image with additive noise n and/or distorted by H, and $\hat{x}$: is the estimator of x with respect to data y. There is another source of information which imposes a key rule in the image processing context, this is the spatial information that represents the likelihood or correlation between the intensity values of a neighborhood of pixels well specified. The modellig, when using MRF, takes into account such spatial interaction and it was introduced and formalized by Besag [1], where the powerfulness of these statistical tools is shown (as well as in pioneering works [2, 3, 13, 14]). Combining both kinds of information in an statistical framework, the restoration is led by an estimation procedure given the maximum a posteriori of the true images when the distortion functionals are known. The algorithms implemented in this work were developed considering a degraded signal, where the resulting non-linear recursive filters show excellent characteristics to preserve all the details contained in the image, and on the other hand, they smooth the noise components. Particularly four estimation schemes are implemented using, the semi-Huber potential function which is proposed as an extension of some previous works [15, 16] (see in Section 3), and the generalized MRF introduced in the work of Bouman [13], the Welch and Tukey potential functions as used in the works of Rivera [17]–[19] (see in Section 4).

Section 2 describes the general definition of MAP estimator, gives an introduction to MRFs and describes briefly alternative similar tools called Hidden Markov Fields (HMF). The potential functions compared in this paper must be obtained or proposed to conduct adequately the inversion process. Such functions are described in Section 3 and 4 where the convexity is the key to formulate an adequate criterion to be minimized. In Sections 5 and 6, the MAP estimators resulting from different MRF structures and some illustrative results are briefly discussed. Finally, in Section 7 some partial conclusions and comments are given.

2 MAP estimation, MRFs and Hidden Markov Fields (HMF)

The problem of image estimation (e.g. filtering or restoration) into a Bayesian framework deals with the solution of an inverse problem, where the estimation process is carried out in a whole stochastic environment. The most popular estimators used nowadays are:

Maximum Likelihood (ML) estimator:

$$\hat{x}_{\mathbf{ML}} = \arg \max_{x \in \mathbf{X}} p(y|x), \qquad (2)$$

where $p(\cdot)$ is the probability function of y given x. This estimator is a classical approach in estimation theory, but under certain circumstances, in image processing restoration, it results in an ill-posed problem or produces excessive noise and also causes smooth of edges. The regularization of the ML estimator leads to a Bayesian approach, where it is important to

exploit all known information or so called prior information about any process under study, which gives a better statistical estimator called

Maximum A Posteriori (MAP) estimator:

$$
\begin{aligned}
\hat{x}_{\mathbf{MAP}} &= \arg \max_{x \in \mathbf{X}} p(x|y) \\
&= \arg \max_{x \in \mathbf{X}} \left(\log p(y|x) + \log g(x) \right),
\end{aligned}
\qquad (3)
$$

in this case, the estimator is regularized by using a Markov random field function $g(x)$ which models all prior information as a whole probability distribution, where $\mathbf{X}$ is the set of possible values of x capable to maximize the posterior law $p(x|y)$, and $p(y|x)$ is the likelihood probability function from y given x.

The Markov random fields (MRF) can be represented in a general way by using the following equation:

$$g(x) = \frac{1}{Z} \exp \left(- \sum_{c \in \mathbf{C}} V_c(x_c) \right), \qquad (4)$$

where Z is a normalization constant, $\mathbf{C}$ is a set of "cliques" c, and $V_c(x_c)$ is a potential function given over the local group of points c. Generally, the "cliques" correspond to the sets of neighborhoods of pixels if $\forall k, r \in c$, k and r are neighbors, and one can construct a neighborhood system called ∂k; for the 8 closest neighbors $\partial k = \{r : |k - r| < 2\}$. A theorem introduced by Hammersley-Clifford [1, 3] probes the equivalence between the Gibbs distribution and the MRFs. The Markov random fields have the capacity to represent various image sources.

There is a variety of cost functions also known as potential functions that can be used into the MRF [13], [17]–[19]. Each potential function characterizes the interactions between pixels in the same local group. As an example, the following family represents convex functions:

$$\rho(\Delta) = |\Delta|^p \qquad (5)$$

where $\Delta = \lambda[x_k - x_r]$, λ is a constant parameter to be chosen, and p takes constant values such as $p \geq 1$ accordingly to the theorem 2 in [13].

On the other hand, if one considers an equivalent problem of segmentation, the reconstruction modeling problem can be seen as Hidden Markov Fields (HMF), in this case $x = (x_s)_{s \in S}$ and $y = (y_s)_{s \in S}$ are considered as a pairwise Markov random fields where x is a realization of a hidden field X given the realization y of the observed field Y [20]. The statistical assumptions and theoretic framework are more general and well developed. The Markovianity of X is not necessary, but the Markovianity of X conditionally to Y must be assured. In the classical assumptions of Markovianity for the hidden X, the law holds according to Eq. (4). However, for the pairwise Markovianity (X, Y) a more general model law is given by

$$g(x, y) = \frac{1}{Z} \exp \left(- \sum_{c \in \mathbf{C}} V_c(x_c, y_c) \right), \qquad (6)$$

in this case the Markov fields (X, Y) are coupled assuming that the posterior law $p(x|y)$ is Markovian. In recent works,

these coupled Markov random fields have been extended in at least two directions:

1) Conditional Markov Fields (CMF) where it is directly assumed the Markovianity of $p(x|y)$, even if the pairwise (X, Y) is not Markovian [21]–[23],

2) Triplet Markov Fields (TMF), here a third finite field $u = (u_s)_{s \in S}$ is introduced assuming that the triplet field (X, U, Y) is Markovian [24]–[26]. It is possible to estimate x from y as shown for image segmentation tasks [24, 27].

These extensions of HMF are promising for segmentation, and could also be an alternative for the equivalent problems of image filtering and restoration.

3 Semi-Huber (SH) proposed potential function

The principal apport of this work is the proposition of the semi-Huber potential function for image restoration, which performance is comparable with respect to half-quadratic functional performance. In order to assure completely the robustness into the edge preserving image filtering, diminishing at the same time the convergence speed, the Huber–like norm or semi–Huber (SH) potential function is proposed as a half-quadratic (HQ) function. Such functional has been used in one dimensional robust estimation as described in [15] for the case of non-linear regression. This function is adjusted in this work in two dimensions according to the following equation:

$$\log g(x) = -\lambda \left(\sum_{\{k,r\} \in \mathbf{C}} b_{kr} \rho_1(x) \right) + \text{ct}, \quad (7)$$

where

$$\rho_1(x) = \frac{\Delta_0^2}{2} \left(\sqrt{1 + \frac{4\varphi_1(x)}{\Delta_0^2}} - 1 \right),$$

and where $\Delta_0 > 0$ is a constant value, b_{kr} is a constant that depends on the distance between the r and k pixels, ct is a constant term, and $\varphi_1(x) = e^2$ where $e = (x_k - x_r)$. The potential function $\rho_1(x)$ must fulfill the following conditions

$$\rho_1(x) \geq 0, \quad \forall x \quad \text{with } \rho_1(0) = 0,$$
$$\psi(x) \equiv \partial \rho_1(x) / \partial x, \quad \text{exists,}$$
$$\rho_1(x) = \rho_1(-x), \quad \text{is symmetric,}$$
$$w(x) \equiv \frac{\psi(x)}{2x}, \quad \text{exists,}$$
$$\lim_{x \to +\infty} w(x) = \mu, \quad 0 \leq \mu < +\infty,$$
$$\lim_{x \to +0} w(x) = M, \quad 0 < M < +\infty. \quad (8)$$

Figure 1(a) shows the behavior of the semi-Huber proposed function for different Δ_0 values, in the range of $x \in [-8, 8]$. Notice that there is not necessary a scale parameter and that the potential function meets all requirements imposed by conditions (8).

4 Generalized Gaussian MRF and other half-quadratic functions

In some works [28]–[32] a variety of new potential functions were introduced, such proposed functions are semi-quadratic functionals or half-quadratic and they characterize certain convexity into the regularization term [33, 34] (e.g. extension of penalization) which permits to build efficient and robust estimators in the sense of data preservation which is linked to the original or source image. Also, the necessary time of computation decreases with respect to other proposed schemes as shown by M. Nikolova [6]–[11], [35] and Labat [36, 37]. On the other hand, a way to obtain the posterior distribution of images has been proposed in previous works from A. Gibbs [38], in this case, it is necessary to use sophisticated stochastic simulation techniques based on the Markov Chain Monte Carlo (MCMC) methods [39, 40]. If it is possible to obtain the posterior distribution of any image, then, it is also possible to sample from such posterior distribution and obtain the MAP estimator, or other estimators such as the median estimator. The MAP and the median estimators search the principal mode of the posterior distribution.

In the present paper some potential functions are compared. The proposed semi-Huber is compared with respect to the generalized Gaussian MRF introduced in [13, 14], the Welch, and Tukey potential functions with granularity control. These two last functions were proposed and used in the works of Rivera [17]–[19] providing excellent performance.

4.1 Generalized Gaussian MRF (GGMRF)

If one considers to generalize the Gaussian MRF (when $p = q = 2$ one has a GMRF, see Eq. (18)) as proposed in [13], then the generalized potential functions can be limited such as

$$\rho_2(\Delta) = |\Delta|^p, \quad \text{for } 1 \leq p \leq 2 \quad (9)$$

obtaining the GGMRF

$$\log g(x) = -\lambda^p \left(\sum_{k \in \mathbf{S}} a_k x_k^p + \sum_{\{k,r\} \in \mathbf{C}} b_{kr} |x_k - x_r|^p \right) + \text{ct}, \quad (10)$$

where theoretically $a_k > 0$ and $b_{kr} > 0$, k is the site or pixel of interest and $\mathbf{S}$ is the set of sites into the whole MRF, and r corresponds to the local neighbors. In practice it is recommended to take $a_k = 0$ thus, the unicity of $\hat{x}_{\mathbf{MAP}}$, can be assured given that the likelihood term is quadratic $q = 2$, then

$$\log g(x) = -\lambda^p \left(\sum_{\{k,r\} \in \mathbf{C}} b_{kr} |x_k - x_r|^p \right) + \text{ct}, \quad (11)$$

and from Eq. (3), $\log p(y|x)$ is strictly convex and so $\hat{x}_{\mathbf{MAP}}$ is continuous in y, and in p. The choice of the power p is capital, since it constrains the convergence speed of the local or global estimator, and the quality of the restored image. Small values for p allows abrupt discontinuities modeling while large values smooth them. Figure 1(b) shows the behavior of the generalized Gaussian function for different p threshold values, in the range of $x \in [-8, 8]$. The proposition of such function avoids the use of a scale parameter and at the same time

J. Europ. Opt. Soc. Rap. Public. **8**, 13072 (2013)

J. I. de la Rosa, *et al.*
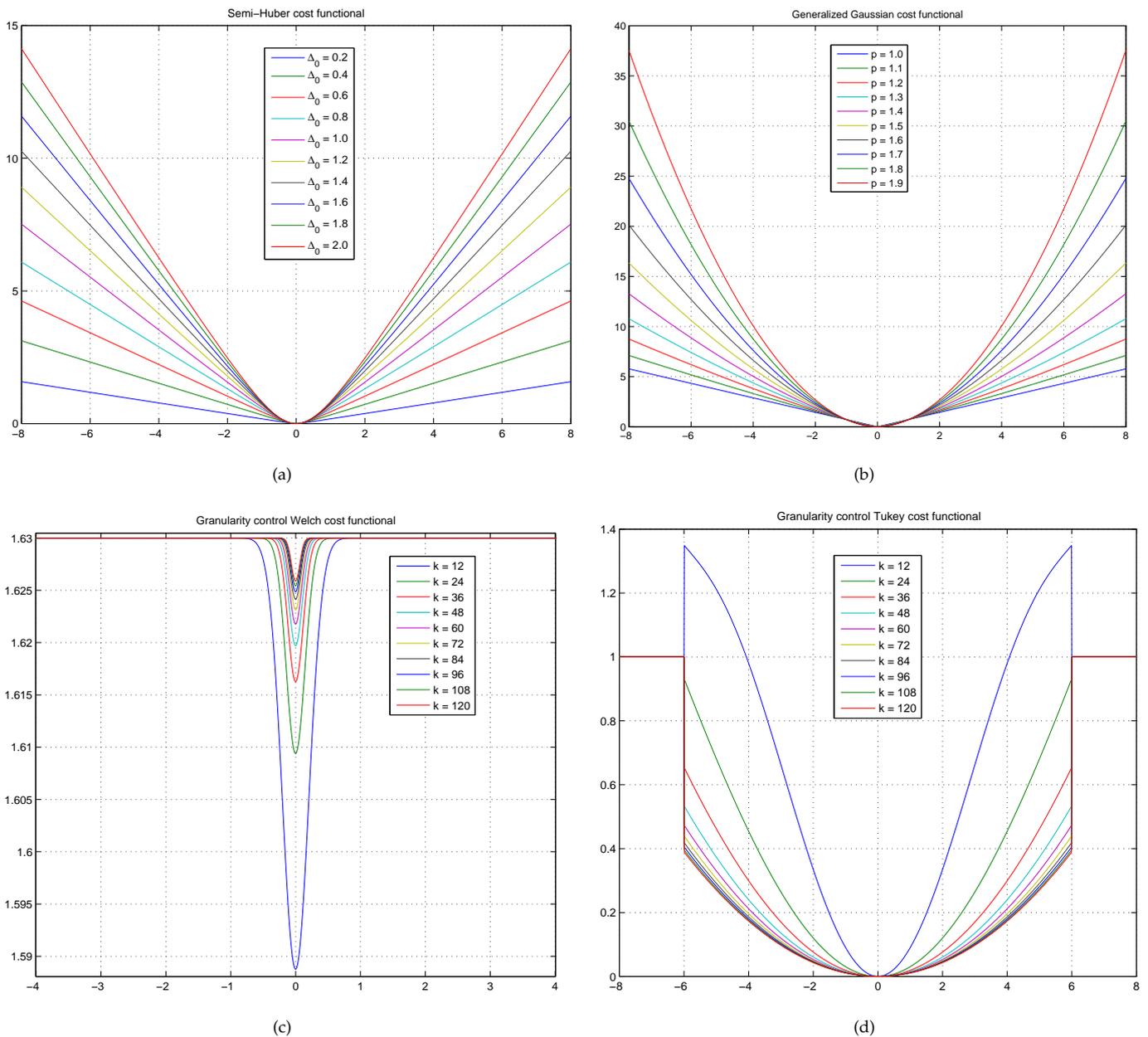
FIG. 1 The four convex potential functions used: (a) The Semi-Huber potential function for Δ_0 with different values; (b) The Generalized Gaussian potential function for p different values, while $q = 2$; (c) The granularity control Welch potential function for k_g different values, while $\mu = 0.01$; (d) The granularity control Tukey potential function for k_g different values, while $\mu = 0.01$.

the potential function meets all requirements imposed by conditions (8).

4.2 Welch potential function

Known as a hard redescender potential function with granularity control given by μ, and proposed in [17]

$$
\begin{aligned}
&\log g(x) \\
&= -\lambda \left(\mu \sum_{\{k,r\}\in\mathbf{C}} b_{kr}\varphi_1(x) + (1-\mu) \sum_{\{k,r\}\in\mathbf{C}} b_{kr}\rho_3(x) \right) \\
&\quad + \mathrm{ct},
\end{aligned}
\tag{12}
$$

where k is a positive scale parameter and

$$
\rho_3(x) = 1 - \frac{1}{2k_g}\exp(-k_g\varphi_1(x))
$$

This function is also half-quadratic such as the Tukey function presented in the following subsection. Figure 1(c) shows the behavior of the Welch function with granularity control for different k_g threshold values, in the range of $x \in [-8,8]$, $\mu = 0.01$. Also, this potential function fulfills all requirements imposed by conditions (8).

4.3 Tukey potential function

This is another hard redescender potential function, proposed in [17] that fulfills all requirements imposed by conditions (8)

$$
\begin{aligned}
&\log g(x) \\
&= -\lambda \left(\mu \sum_{\{k,r\}\in\mathbf{C}} b_{kr}\varphi_1(x) + (1-\mu) \sum_{\{k,r\}\in\mathbf{C}} b_{kr}\rho_4(x) \right) \\
&\quad + \mathrm{ct},
\end{aligned}
\tag{13}
$$

where

$$\rho_4(x) = \begin{cases} 1 - (1 - (2e/k_g)^2)^3, & \text{for } |e/k_g| < 1/2, \\ 1, & \text{otherwise.} \end{cases}$$

and where k_g is also a scale parameter. On the other hand, $\varphi_1(x)$ can be the quadratic function which together with μ induces the granularity control. Figure 1(d) shows the behavior of the Tukey function with granularity control for different k_g values, in the range of $x \in [-8, 8]$, $\mu = 0.01$.

5 MAP estimators and practical convergence

5.1 Image filtering

In this section some estimators are deduced, the single problem of filtering noise to restore an observed signal y leads to establish the estimators. The observation equation could be

$$y = x + n, \text{ where } n \sim \mathcal{N}(0, I\sigma_n^2)$$

where I is the identity matrix, and the general MAP estimator for this case is deduced from the next minimization process

$$\widehat{x}_{\mathbf{MAP1}} = \arg \min_{x \in \mathbf{X}} \left(- \log p(y|x) - \log g(x) \right). \quad (14)$$

Thus, the MAP estimators for this particular problem under hypothesis of centered Gaussian noise with variance σ_n^2 are given using the four previous potential functions. The first MAP estimator can be obtained when using our proposed semi-Huber (SH) MRF function. The global estimator can be described by the equation

$$\widehat{x}_{\mathbf{MAP1}}$$
$$= \arg \min_{x \in \mathbf{X}} \left\{ \sum_{k \in \mathbf{S}} |y_k - x_k|^2 + \lambda \left(\sum_{\{k,r\} \in \mathbf{C}} b_{k-r} \rho_1(x) \right) \right\}, \quad (15)$$

and, the local estimator leads to the following expression for the first local MAP estimator

$$\widehat{x}_1 = \arg \min_{x \in \mathbf{X}} \left\{ |y_k - x_k|^2 + \lambda \left(\sum_{r \in \partial \mathbf{k}} b_{k-r} \rho_1(x) \right) \right\}. \quad (16)$$

On the other hand, the minimization problem leads to consider various methods [6, 9]–[11, 35, 41]:

- global iterative techniques such as: the descendent gradient [35], conjugate gradient [17] (for recent propositions one can consult the work of Labat [36, 37]), Gauss-Seidel, Over-relaxed methods, etc.

- local minimization techniques: minimization at each pixel x_k (which generally needs more time, but from our point of view are more precise), where also some of the above methods can be used.

In this work the local techniques were used (the expectation maximization (EM) could also be implemented, or the complete half-quadratic scheme as proposed by Geman and Reinolds [33], and Geman and Yang [34]), since all hyper-parameters included into the potential functions were chosen heuristically or according to values proposed in references. Only, the step of minimization with respect to x was implemented to probe convergence of estimators. The second MAP estimator can be obtained when using the GGMRF function. Its global estimator is described by the following equation

$$\widehat{x}_{\mathbf{MAP2}}$$
$$= \arg \min_{x \in \mathbf{X}} \left\{ \sum_{k \in \mathbf{S}} |y_k - x_k|^q + \sigma^q \lambda^p \sum_{\{k,r\} \in \mathbf{C}} b_{k-r} |x_k - x_r|^p \right\}, \quad (17)$$

where $b_{kr} = b_{k-r} = b_{r-k}$ assuming homogeneity of the MRFs. In this case, the local MAP estimator for the GGMRF is given by

$$\widehat{x}_2 = \arg \min_{x \in \mathbf{X}} \left\{ |y_k - x_k|^q + \sigma^q \lambda^p \sum_{r \in \partial \mathbf{k}} |x_k - x_r|^p \right\}, \quad (18)$$

where, according to the value of parameters p and q, the performance of such estimator varies. For example, if $p = q = 2$, one has the Gaussian condition of the potential function where the obtained estimator is similar to the least-squares one (L_2 norm) since the likelihood function is quadratic, with an additional quadratic term of penalization which degrades (e.g. large smoothing) the estimated image

$$\widehat{x}_k = \frac{y_k + (\sigma\lambda)^2 \sum_{r \in \partial \mathbf{k}} b_{r-k} x_r}{1 + (\sigma\lambda)^2 \sum_{r \in \partial \mathbf{k}} b_{r-k}}. \quad (19)$$

On the other hand, in the case of $p = q = 1$, the criterion is absolute (L_1 norm), and the estimator converges to the median estimator, which in practice, it is difficult to implement in a precise way

$$\widehat{x}_k = \text{median} (y_k, x_{r_1}, \dots, x_{r_l}). \quad (20)$$

This criterion is not differentiable at zero and this fact causes instability in the minimization procedure. For intermediate values of p and q the estimators become sub-optimal, and the iterated methods can be used to minimize the obtained criterions. Some iterative methods are the sub–gradient, or the Levenberg–Marquardt method of MATLAB 7, the last was used in this work. For cases where $q \neq p$, for example $q = 2$ and $1 < p < 2$, some studies and different prior functions have been proposed in [7]–[11, 35], particularly in [7]–[11] where non-convex regularized least-squares schemes are deduced and its convergence is analyzed (where $0 < p < 1$) with very good times of convergences as presented in [42]. The local or global condition of the estimator depends thus on:

1) if one has values of $1 < p < 2$: the estimator $\widehat{x}_{min.loc} \rightarrow \widehat{x}_{min.glob}$, which means that a local minimum would coincide with a global minimum,

2) moreover, if $p \neq q$, the criterion is not homogeneous, but: $\widehat{x}(\alpha y, \lambda) = \alpha \widehat{x}(y, \alpha^{1-q/p}\lambda)$, assuring the convergence and existence of the estimator which is continuous with respect to p.

Moreover, since the noise is Gaussian, the value of $q = 2$ is a good choice. However, if the noise is non-Gaussian, and if the structure of noise is known, then the likelihood term changes to give a particular estimator (such as that proposed by Bertaux [5]) with some properties of robustness to the minimization procedure. If the structure of noise is assumed unknown, still one could reconsider the use of the GGMRF with values for $q \in (1, 2]$, or consider also another type of potential functions as those described in some works of Idier [28]–[31] and Nikolova [9]–[11, 35].

The third MAP estimator is obtained when using the Welch potential function with granularity control, that is,

$$\widehat{x}_{\mathbf{MAP3}} = \arg \min_{x \in \mathbf{X}} \left\{ \sum_{k \in \mathbf{S}} |y_k - x_k|^2 + \lambda \right.$$
$$\left. \cdot \left(\mu \sum_{\{k,r\} \in \mathbf{C}} b_{k-r}\varphi_1(x) + (1-\mu) \sum_{\{k,r\} \in \mathbf{C}} b_{k-r}\rho_3(x) \right) \right\}, \quad (21)$$

and its local version is

$$\widehat{x}_3 = \arg \min_{x \in \mathbf{X}} \left\{ |y_k - x_k|^2 + \lambda \right.$$
$$\left. \cdot \left(\mu \sum_{r \in \partial \mathbf{k}} b_{k-r}\varphi_1(x) + (1-\mu) \sum_{r \in \partial \mathbf{k}} b_{k-r}\rho_3(x) \right) \right\}. \quad (22)$$

And finally, the fourth MAP estimator is deduced from the Tukey potential function with granularity control, deriving the following global estimator

$$\widehat{x}_{\mathbf{MAP4}} = \arg \min_{x \in \mathbf{X}} \left\{ \sum_{k \in \mathbf{S}} |y_k - x_k|^2 + \lambda \right.$$
$$\left. \cdot \left(\mu \sum_{\{k,r\} \in \mathbf{C}} b_{k-r}\varphi_1(x) + (1-\mu) \sum_{\{k,r\} \in \mathbf{C}} b_{k-r}\rho_4(x) \right) \right\}, \quad (23)$$

where the local estimator is given by

$$\widehat{x}_4 = \arg \min_{x \in \mathbf{X}} \left\{ |y_k - x_k|^2 + \lambda \right.$$
$$\left. \cdot \left(\mu \sum_{r \in \partial \mathbf{k}} b_{k-r}\varphi_1(x) + (1-\mu) \sum_{r \in \partial \mathbf{k}} b_{k-r}\rho_4(x) \right) \right\}. \quad (24)$$

The use of a prior distribution function based on the logarithm, with any degree of convexity and quasi-homogeneous, permits to consider a variety of possible choices of potential functions. Maybe, the most important challenges that must be well solved are: the adequate selection of hyper-parameters from potential functions, where different versions of the EM algorithms try to tackle this problem [11, 28, 31, 32], another is the minimization procedure which in any sense will regulate the convergence speed as proposed in [9, 10, 17, 33, 34, 36, 37, 41].

5.2 Image deconvolution

On the other hand, for the problem of image deblurring to restore an observed signal y, the observation equation used is given by

$$y = Hx + n, \text{ with } n \sim \mathcal{N}(0, I\sigma_n^2) \quad (25)$$

for the four MAP estimators the likelihood term changes, such that,

$$\widehat{x}_{\mathbf{MAP_m}} = \arg \min_{x \in \mathbf{X}} \left\{ \sum_{\mathbf{S}} |y - Hx|^2 - \log g(x) \right\}, \quad (26)$$

where the matrix H is known and given by the following truncated Gaussian blurring function,

$$h(i, j) = \exp \left(\frac{-i^2 - j^2}{2\sigma_b^2} \right), \text{ for } -3 \leq i, j \leq 3, \quad (27)$$

as used also in [42], with $\sigma_b = 1.5$, and $m = 1, 2, 3, 4$ according to the four SH, GGMRF, Welsh and Tukey potential functions. Here, the results were improved combining ideas introduced in a similar Bayesian way by Levin [43] adding a Sparse prior (SP) for filtering and then reconstructing the image.

6 Some experiments

Results presented in this section were concerned experimenting extensively with five images: synthetic, Lena, Cameraman, Boat and fringe pattern, to probe the performance of the presented estimators.

6.1 Image filtering

Continuing with the problem of filtering noise, some estimation results are presented when images are contaminated by Gaussian noise, and there are not other type of distortions. The first experiment was made considering $\sigma_n = 5, 10, 15$. In the work by De la Rosa [16] some results were previously presented based on the analysis of a synthetic image and the standard image of "Lena", different levels of noise were added to the images: $n \sim \mathcal{N}(0, I\sigma_n^2)$, the values of σ_n are given such that the obtained degradation is perceptible and difficult to eliminate. The results were compared using different values for Δ_0 and $\lambda = 1$ (MAP1), different values for p and λ preserving $q = 2$ (MAP2), and different values for k, μ and λ (MAP3 and MAP4). Generally, with the four estimators the filtering task gives good visual results (see Figures 2, 3, 4 and 5), but the time of computation is different between them. The faster estimator is the MAP3, while the slowest is the MAP2 with $p = 1.2$ which results correspond to the Cameraman in Figure 5(d). In the case of the Welsh and Tukey functionals the tuning problems must be solved implementing in correct ways more sophisticated algorithms based, for example, on the expectation maximization method. Figure 6 shows a synthetic generated fringe image, which was used to probe performance of estimators. In this case, it is known how the noise structure that contaminates data is, but the signal-to-noise is unknown.

Once again the obtained results coincide with the previous results for other images, but with an increase of computation time, which has a relation with the image dimensions (as

J. Europ. Opt. Soc. Rap. Public. **8,** 13072 (2013)

J. I. de la Rosa, *et al.*
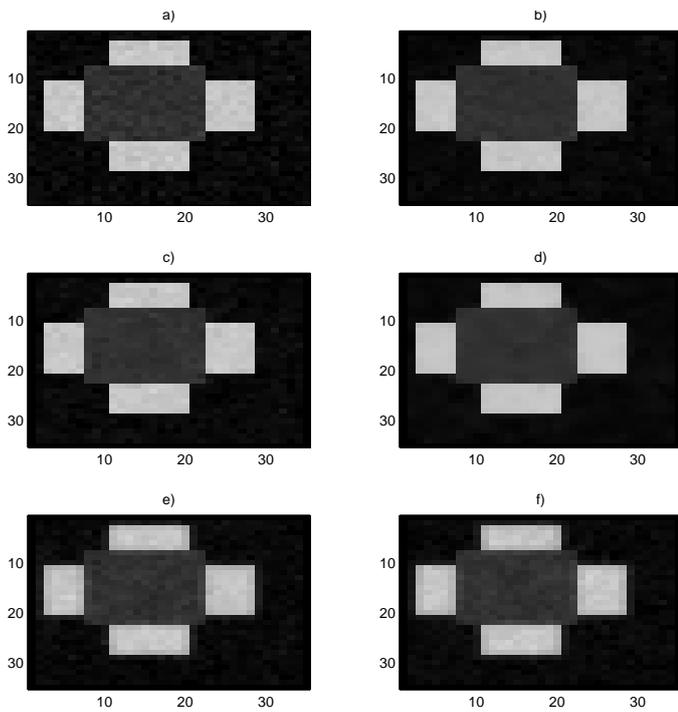
FIG. 2 a) Image with noise, $\sigma_n = 6$ (35×35), b) MAP2 estimation with $p = 1.2$ and $q = 2$ (8 s), c) MAP2 estimation with $p = 1.1$ and $q = 2$ (13 s), d) MAP1 estimation (6 s), e) MAP3 estimation (5 s) and f) MAP4 estimation (4 s).
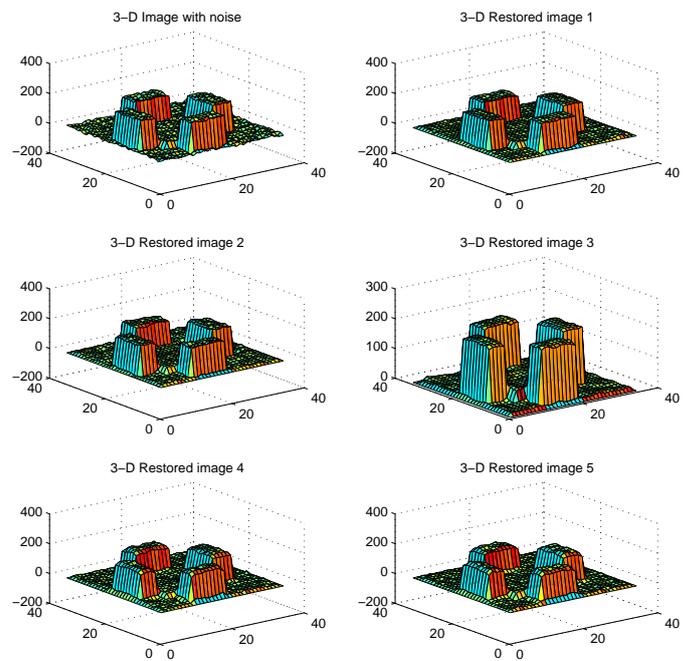
FIG. 3 a) Image with noise, $\sigma_n = 6$ (35×35) 3-D view, b) MAP2 estimation with $p = 1.2$ and $q = 2$, c) MAP2 estimation with $p = 1.1$ and $q = 2$, d) MAP1 estimation, e) MAP3 estimation and f) MAP4 estimation.
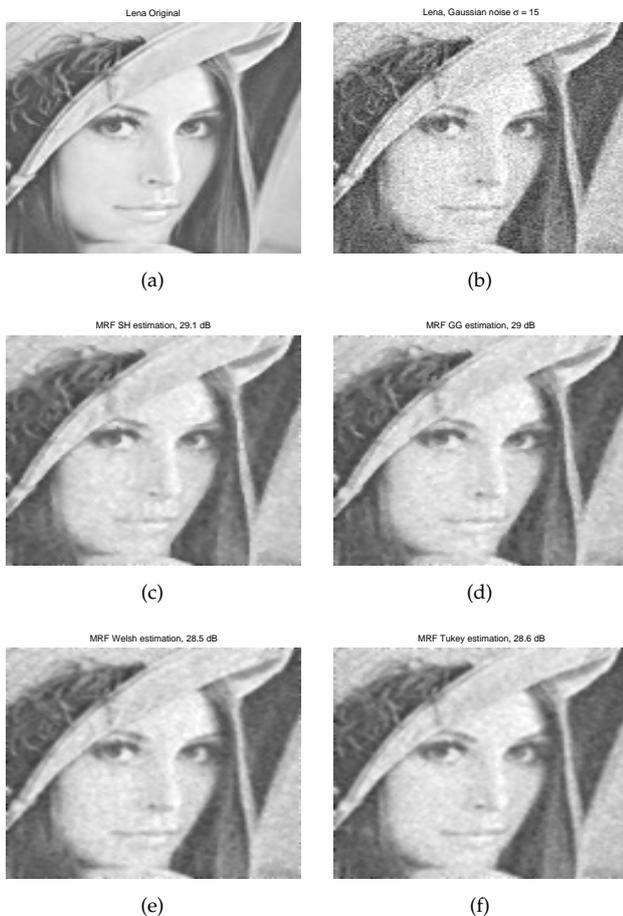
FIG. 4 Results for Lena standard image: (a) describes the original image; (b) describes the noisy image using Gaussian noise with $\sigma = 15$; (c) filtered image using MAP1 ($\Delta_0 = 20$); (d) filtered image using MAP2 ($\lambda = 30$, $p = 1.2$); (e) filtered image using MAP3 ($k_g = 2000$, $\mu = 0.025$, $\lambda = 30$); (f) filtered image using MAP4 ($k_g = 2000$, $\mu = 0.025$, $\lambda = 30$).

FIG. 5 Results for Cameraman standard image: (a) describes the original image; (b) describes the noisy image using Gaussian noise with $\sigma = 15$; (c) filtered image using MAP1 ($\Delta_0 = 20$); (d) filtered image using MAP2 ($\lambda = 30$, $p = 1.2$); (e) filtered image using MAP3 ($k_g = 2000$, $\mu = 0.025$, $\lambda = 30$); (f) filtered image using MAP4 ($k_g = 2000$, $\mu = 0.025$, $\lambda = 30$).

J. Europ. Opt. Soc. Rap. Public. **8**, 13072 (2013)

J. I. de la Rosa, *et al.*
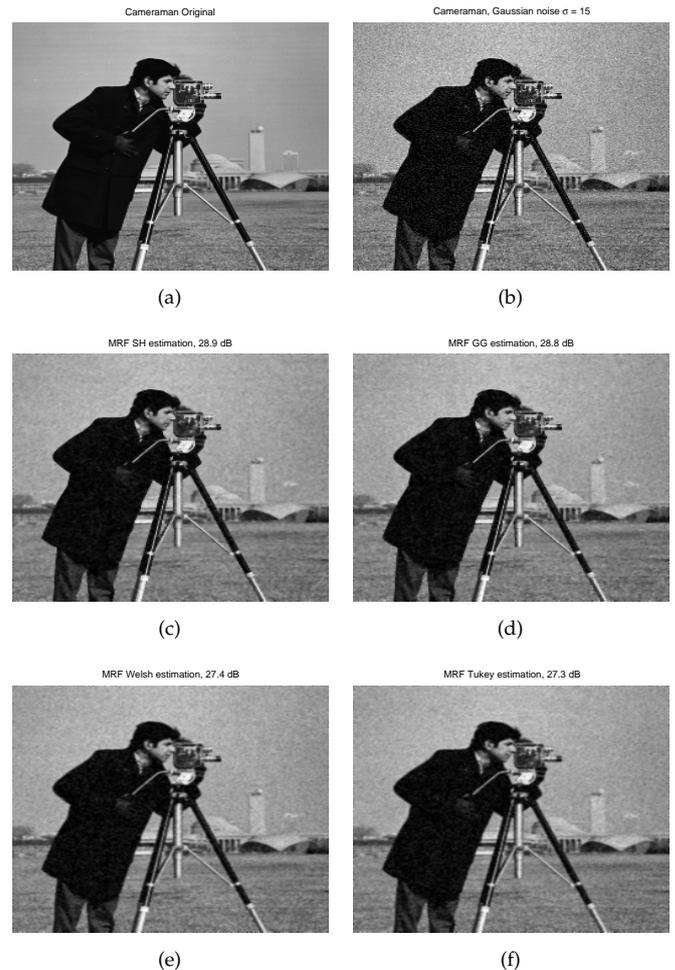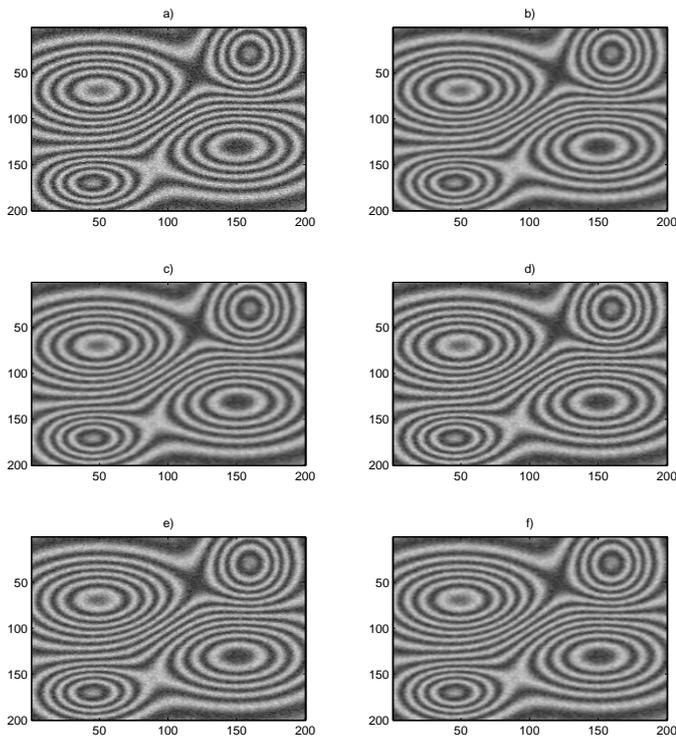
FIG. 6 Image with Gaussian noise with unknown σ (200 × 200), b) MAP2 estimation, p = 1.5 (100 s), c) MAP2 estimation p = 1.2 (120 s), d) MAP1 estimation (52 s), e) MAP3 estimation (50 s) and f) MAP4 estimation (52 s).

shown in Table 1). Some interesting applications of robust estimation are particulary focused on phase recovery from fringe patterns as presented in work [44], phase unwrapping, and some other problems in optical instrumentation. In this sense some filtering results were thus obtained using the presented MAP estimators.

Finally, Table 1 shows the performance of the four MAP estimators for the problem of filtering Gaussian noise, where an objective evaluation is conducted accordingly to the peak signal to noise ratio (PSNR). Also the computation times in MATLAB are shown in Table 1. Such comparative evaluation shows that our proposed approach MAP1, gives better or similar performance with respect to MAP2, MAP3, and MAP4. On the other hand, the use of half-quadratic potential functions permits flexibility on the computation times [7, 8, 11], but still it is a challenge to tune correctly the hyper-parameters to obtain a better performance in the sense of restoration. Perhaps the most simple potential function to tune is the semi-Huber (MAP1). Also making the correct hypothesis over the noise could help to improve the performance of the estimator. This could be directly reflected by proposing a more adapted likelihood function, as proposed by Bertaux [5] and some other recent works [9] (in cases of non-Gaussian noise), where a connection with variational and partial differential equations is illustrated evoking the famous work of Perona and Malik [45], and some recent related works.

6.2 Image deconvolution

Now, for the problem of image deconvolution some estimation results are presented when images are contaminated by Gaussian noise, and Gaussian distortion (with $\sigma_b = 1.5$) Blur-
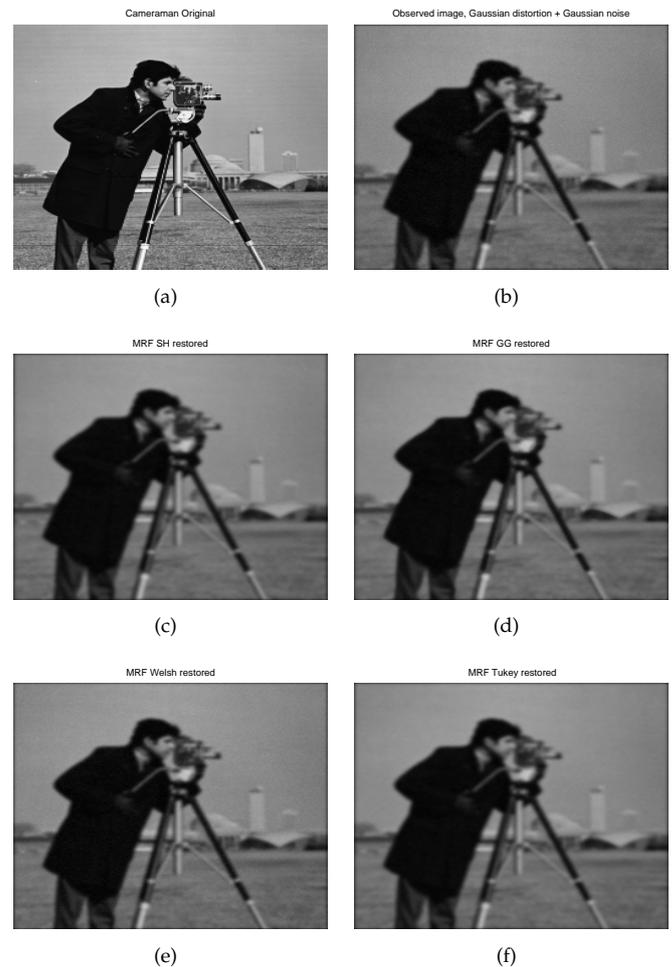
FIG. 7 Results for Cameraman standard image: (a) describes the original image; (b) describes the distorted image using Gaussian noise with $\sigma = 3$; (c) restored image using MAP1 ($\Delta_0 = 20$); (d) restored image using MAP2 ($\lambda = 30$, $p = 1.2$); (e) restored image using MAP3 ($k_g = 2000$, $\mu = 0.0015$, $\lambda = 10$); (f) restored image using MAP4 ($k_g = 2000$, $\mu = 0.0015$, $\lambda = 10$).

ring the image. This second experiment was made considering thus $\sigma_n = 3, 5, 7$. The results were compared using different values for Δ_0 and with $\lambda = 1$ (MAP1), different values for p and λ preserving $q = 2$ (MAP2), and different values for k, μ and λ (MAP3 and MAP4). Figures 7 and 9 show a comparison of results obtained for the Cameraman and Boat images accordingly to the four MAP estimators. One can notice that preserving values of hyper-parameters near those used for the filtering case, the estimators smooth the noise but does not made good deblurring or recuperation of the image (the PSNR for this case is depicted in Table 2). One must change the hyper-parameters values searching a trade of between the granularity of the noise and the sharpness of the image to make a good deblurring task (decompositing in two steps the image reconstruction). Figure 8 and 10 show the obtained results on the Cameraman and Boat images using a combination of proposed estimators together with a Sparse prior (SP) deconvolution technique introduced in [43]. The improvement in the restoration is visible; here the recuperation was obtained in two steps; first the noise was smoothed and then, the deblurring was obtained using SP deconvolution technique (our approach can be used in both steps, tuning the hyper-parameters two times). On the hands, in Table 2 the performance of the four MAP estimators and the SP de-

	$\sigma = 15$	MAP1	MAP2	MAP3	MAP4
Synthetic	PSNR	24.7	24.6	24.7	24.6
35×35	PSNR filt.	28.8	27.6	25.5	25.6
	Time (s)	5.9	7.6	4.6	5.2
Lena	PSNR	24.5	24.5	24.6	24.6
120×120	PSNR filt	29.1	28.9	26.0	27.5
	Time (s)	75.3	80.6	59.9	59.1
Cameraman	PSNR	24.6	24.7	24.6	24.6
256×256	PSNR filt	28.9	28.8	27.3	27.4
	Time (s)	341.7	355.9	243.8	243.6
Boat	PSNR	24.6	24.6	24.6	24.6
512×512	PSNR filt	29.4	29.5	27.3	28.8
	Time (s)	1243.3	1545.5	1014.2	1051.4

TABLE 1 Results obtained in evaluating the filtering capacity of the different MAP estimators using four images.

FIG. 8 Results for Cameraman standard image: (a) describes the original image; (b) describes restored image using only a frequency domain (FD) deconvolution technique; (c) restored image using MAP2 and FD ($\Delta_0 = 20$); (d) restored image using MAP1 and FD ($\lambda = 0.15$, $p = 1.2$); (e) restored image using MAP3 and FD ($k_g = 2000$, $\mu = 0.0015$, $\lambda = 10$); (f) restored image using MAP4 and FD ($k_g = 2000$, $\mu = 0.0015$, $\lambda = 10$).
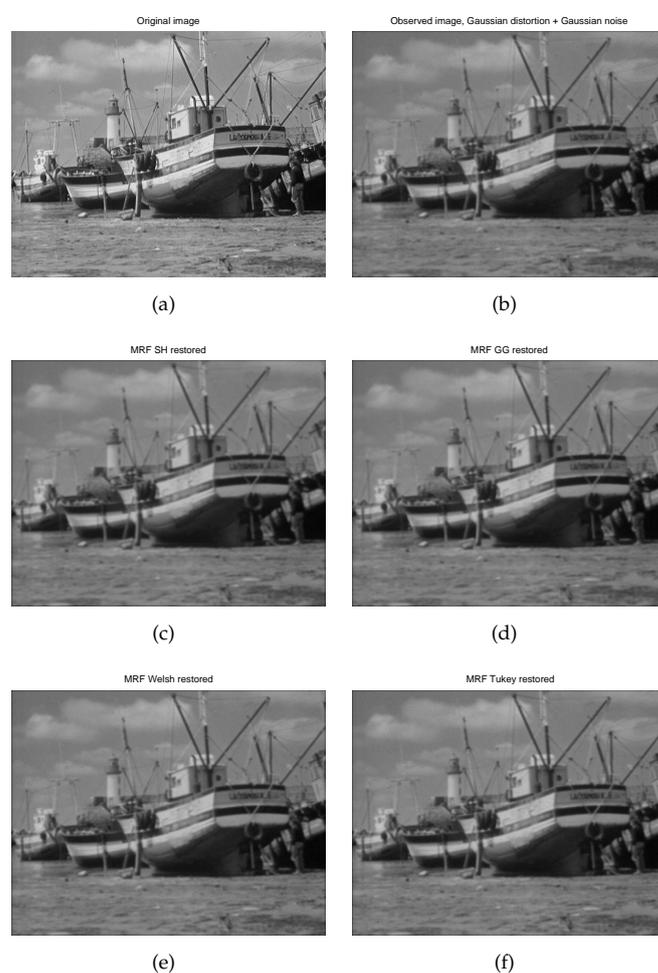
FIG. 9 Results for Boat standard image: (a) describes the original image; (b) describes the distorted image using Gaussian noise with $\sigma = 3$; (c) restored image using MAP1 ($\Delta_0 = 20$); (d) restored image using MAP2 ($\lambda = 30$, $p = 1.2$); (e) restored image using MAP3 ($k_g = 2000$, $\mu = 0.0015$, $\lambda = 10$); (f) restored image using MAP4 ($k_g = 2000$, $\mu = 0.0015$, $\lambda = 10$).

convolution is shown, where an objective evaluation is made accordingly to the PSNR and also computation times in MAT-LAB are shown. Here also the approach MAP1, gives similar performance with respect to MAP2, MAP3, and MAP4.

7 Conclusions and comments

The use of a prior distribution functions based on the logarithm, with any degree of convexity and quasi-homogeneous, permits to consider a variety of possible choices of potential functions. Maybe, the most important challenges that must be well solved are: the adequate selection of hyper-parameters

＼	$\sigma_n = 3$	MAP1	MAP2	MAP3	MAP4
Lena	PSNR	17.4	17.4	17.4	17.4
120×120	PSNR restored	17.5	17.6	17.4	17.3
	PSNR restored FD	20.8	20.8	20.8	20.4
	Time (s)	58.9	91.3	58.8	59.2
Cameraman	PSNR	19.3	19.3	19.3	19.3
256×256	PSNR restored	19.4	19.4	19.4	19.3
	PSNR restored FD	22.5	22.3	22.5	22.2
	Time (s)	257.7	408.0	256.8	256.9
Boat	PSNR	20.4	20.4	20.4	20.4
512×512	PSNR restored	20.5	20.5	20.4	20.4
	PSNR restored FD	25.6	25.8	26.0	25.9
	Time (s)	1014.9	1606.5	1011.8	1017.5

TABLE 2 Results obtained in evaluating the deconvolution capacity of the different MAP estimators.
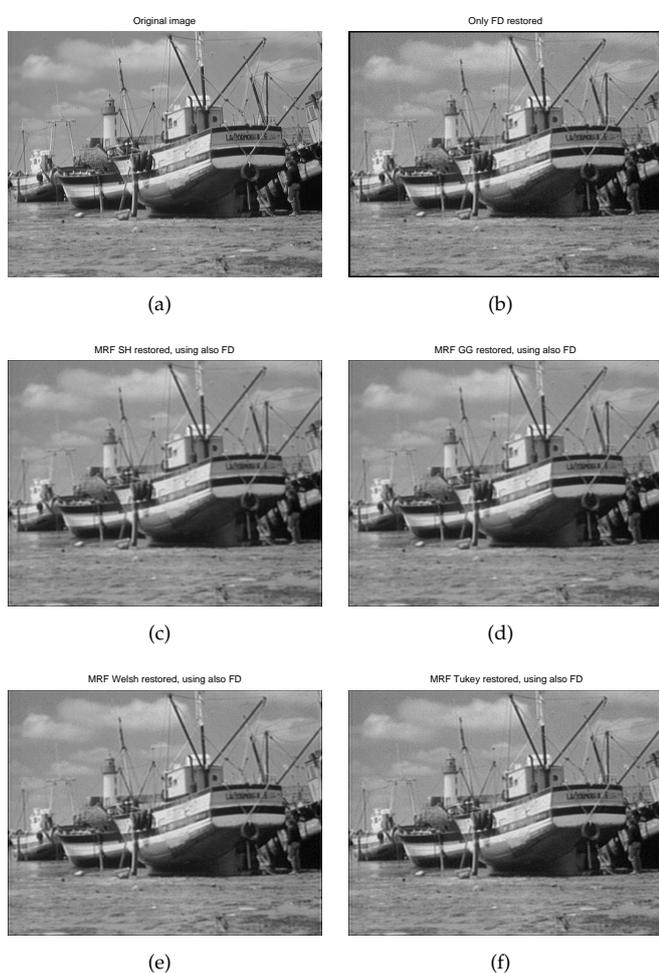
FIG. 10 Results for Boat standard image: (a) describes the original image; (b) describes restored image using only a Sparse prior (SP) deconvolution technique; (c) restored image using MAP1 and SP ($\Delta_0 = 20$); (d) restored image using MAP2 and SP ($\lambda = 0.15$, $p = 1.2$); (e) restored image using MAP3 and SP ($k_g = 2000$, $\mu = 0.0015$, $\lambda = 10$); (f) restored image using MAP4 and SP ($k_g = 2000$, $\mu = 0.0015$, $\lambda = 10$).

from potential functions, where different versions of the EM algorithms try to tackle this problem, another is the minimization procedure which in any sense will regulate the convergence speed as proposed by Allain [41], German [33, 34], Rivera [17], Labat [36, 37], Nikolova [9, 10, 35].

In the case of the semi-Huber potential function (MAP1), the tuning is less complicated and of course, the estimator manipulation is far simpler than Welsh (MAP3) and Tukey (MAP4). On the other hand, this problem can be solved as argued by Idier [28] and Rivera [17] by implementing more sophisticated algorithms with the compromise to reduce time of computation and better quality in restoration as recently exposed by Ruimin Pan [12, 28, 42]. Some advantages on the use of GGMRF as prior information into the Bayesian estimation (MAP2) are: the continuity of the estimator is assured as a function of the data values when $1 < p \leq 2$ even for Gaussian and non-Gaussian noise asumption. The edge preserving is also assured, over all when $p \to 1$ and obviously it depends on the choice between the interval $1 < p < 2$. The final objective of this work has been to contribute with a series of software tools for image analysis focused for instance to optical instrumentation tasks such as those treated in the works [44] and [18, 19] obtaining competitive results in filtering and reconstruction. More over, the extensions of HMF will be considered in future work as an alternative for solving the problems of image filtering and restoration.

Acknowledgements

References

[1] J. E. Besag, "Spatial interaction and the statistical analysis of lattice systems," J. Roy. Stat. Soc. Ser. B Met. **36**, 192–236 (1974).

[2] J. E. Besag, "On the statistical analysis of dirty pictures," J. Roy. Stat. Soc. Ser. B Met. **48**, 259–302 (1986).

[3] S. Geman, and D. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," IEEE T. Pattern Anal. **6**, 721–741 (1984).

[4] H. C. Andrews, and B. R. Hunt, *Digital image restoration* (Prentice-Hall, Inc., New Jersey, 1977).

[5] N. Bertaux, Y. Frauel, P. Réfrégier, and B. Javidi, "Speckle removal using a maximum-likelihood technique with isoline gray-level regularization," J. Opt. Soc. Am. A **21**, 2283–2291 (2004).

[6] T. F. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," SIAM J. Appl. Math. **66**, 1632–1648 (2006).

[7] S. Durand, and M. Nikolova, "Stability of the minimizers of least squares with a non-convex regularization. Part I: Local behavior," J. Appl. Math. Optimizat. **53**, 185–208 (2006).

[8] S. Durand, and M. Nikolova, "Stability of the minimizers of least squares with a non-convex regularization. Part II: Global behavior," J. Appl. Math. Optimizat. **53**, 259–277 (2006).

[9] M. Nikolova, *Functionals for signal and image reconstruction: properties of their minimizers and applications* (Centre de Mathématiques et de Leurs Applications CNRS-UMR 8536, 2006).

[10] M. Nikolova, "Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares," Multiscale Model. Simul. **4**, 960–991 (2005).

[11] M. Nikolova, and M. Ng, "Analysis of half-quadratic minimization methods for signal and image recovery," SIAM J. Sci. Comput. **27**, 937–966 (2005).

[12] R. Pan, and S. J. Reeves, "Efficent Huber-Markov edge-preserving image restoration," IEEE T. Image Process. **15**, 3728–3735 (2006).

[13] C. Bouman, and K. Sauer, "A generalized Gaussian image model for edge-preserving MAP estimation," IEEE T. Image Process. **2**, 296–310 (1993).

[14] K. Sauer, and C. Bouman, "Bayesian estimation of transmission tomograms using segmentation based optimization," IEEE T. Nucl. Sci. **39**, 1144–1152 (1992).

[15] J. I. De la Rosa, and G. Fleury, "Bootstrap methods for a measurement estimation problem," IEEE T. Instrum. Meas. **55**, 820–827 (2006).

[16] J. I. De la Rosa, J. J. Villa, and Ma. A. Araiza, "Markovian random fields and comparison between different convex criteria optimization in image restoration," in *Proceedings to International Conference on Electronics, Communications, and Computers 2007*, 1–6 (IEEE-UDLA, Cholula, 2007).

[17] M. Rivera, and J. L. Marroquin, "Efficent half-quadratic regularization with granularity control," Image Vision Comput. **21**, 345–357 (2003).

[18] M. Rivera, and J. L. Marroquin, "Half-quadratic cost functions for phase unwrapping," Opt. Lett. **29**, 504–506 (2004).

[19] M. Rivera, "Robust phase demodulation of interferograms with open or closed fringes," J. Opt. Soc. Am. A **22**, 1170–1175 (2005).

[20] W. Pieczynski, and A.-N. Tebbache, "Pairwise Markov random fields and segmentation of textured images," Machine Graph. Vision **9**, 705–718 (2000).

[21] J. Lafferty, A. McCallum, and F. Pereira, "Conditional Random Fields: Probabilistic models for segmenting and labeling sequence data," in *Proceedings to International Conference on Machine Learning 2001*, 282–289 (International Machine Learning Society, Williamstown, 2001).

[22] S. Kumar, and H. Martial, "Discriminative random fields," Int. J. Comput. Vision **68**, 179–201 (2006).

[23] A. Quattoni, S. Wang, L.-P. Morency, M. Collins, and T, Darrell "Hidden conditional random fields," IEEE T. Pattern Anal. Mach. Intell. **29**, 1848–1853 (2007).

[24] D. Benboudjema, and W. Pieczynski, "Unsupervised statistical segmentation of nonstationary images using triplet Markov fields," IEEE T. Pattern Anal. Mach. Intell. **29**, 1367–1378 (2007).

[25] P. Zhang, M. Li, Y. Wu, L. Gan, M. Liu, F. Wang, and G. Liu, "Unsupervised multi-class segmentation of SAR images using fuzzy triplet Markov fields model," Pattern Recogn. **45**, 4018–4033 (2012).

[26] F. Wang, Y. Wu, Q. Zhang, P. Zhang, M. Li, and Y. Lu, "Unsupervised change detection on SAR images using triplet Markov field model," IEEE Geosci. Remote Sens. Lett. **10**, 697–701 (2013).

[27] D. Benboudjema, N. Othman, B. Dorizzi, and W. Pieczynski, "Segmentation d' images des yeux par champs de markov triplets: Application à la biométrie," in *Proceedings to Colloque GRETSI 2013* (GRETSI, Brest, 2013).

[28] F. Champagnat, and J. Idier, "A connection between half-quadratic criteria and EM algorithms," IEEE Signal. Proc. Lett. **11**, 709–712 (2004).

[29] P. Ciuciu, and J. Idier, "A half-quadratic block-coordinate descent method for spectral estimation," Signal Processing **82**, 941–959 (2002).

[30] P. Ciuciu, J. Idier, and J.-F. Giovannelli, "Regularized estimation of mixed spectra using circular Gibbs-Markov model," IEEE T. Signal Proces. **49**, 2202–2213 (2001).

[31] J.-F. Giovannelli, J. Idier, R. Boubertakh, and A. Herment, "Unsupervised frequency tracking beyond the Nyquist frequency using Markov chains," IEEE T. Signal Proces. **50**, 2905–2914 (2002).

[32] J. Idier, "Convex half-quadratic criteria and interacting auxiliary variables for image restoration," IEEE T. Image Process. **10**, 1001–1009 (2001).

[33] D. Geman, and G. Reinolds, "Constrained restoration and the recovery of discontinuities," IEEE T. Pattern Anal. Mach. Intell. **14**, 367–383 (1992).

[34] D. Geman, and C. Yang, "Nonlinear image recovery with half-quadratic regularization," IEEE T. Image Process. **4**, 932–946 (1995).

[35] M. Nikolova, and R. Chan, "The equivalence of half-quadratic minimization and the gradient linearization iteration," IEEE T. Image Process. **16**, 1623–1627 (2007).

[36] C. Labat, and J. Idier, *Convergence of truncated half-quadratic algorithms using conjugate gradient* (IRCCyN UMR CNRS 6597, 2006).

[37] C. Labat, *Algorithmes d'optimisation de critères pénalisés pour la restauration d'images. Application à la déconvolution de trains d'impulsions en imagerie ultrasonore*, (Ph.D. Thesis, École Centrale de Nantes, 2006) in French.

[38] A. L. Gibbs, *Convergence of Markov Chain Monte Carlo algorithms with applications to image restoration*, (Ph.D. Thesis, University of Toronto, 2000).

[39] R. M. Neal, *Probabilistic inference using Markov chain Monte Carlo methods* (Department of Computer Science, University of Toronto, Technical Report CRG-TR-93-1, 1993).

[40] C. P. Robert, and G. Casella, *Monte Carlo Statistical Methods* (Second edition, Springer Verlag, New York, 2004).

[41] M. Allain, J. Idier, and Y. Goussard, "On global and local convergence of half-quadratic algorithms," IEEE T. Image Process. **15**, 1130–1142 (2006).

[42] M. Nikolova, M. K. Ng, and C.-P. Tam, "Fast nonconvex nonsmooth minimization methods for image restoration and reconstruction," IEEE T. Image Process. **19**, 3073–3088 (2010).

[43] A. Levin, R. Fergus, F. Durand, and W. T. Freeman, "Image and depth from a conventional camera with a coded aperture," in *Proceedings to SIGGRAPH '07*, 10.1145/1275808.1276464 (ACM, San Diego, 2007).

[44] J. J. Villa, J. I. De la Rosa, G. Miramontes, and J. A. Quiroga, "Phase recovery from a single fringe pattern using an orientational vector field regularized estimator," J. Opt. Soc. Am. A **22**, 2766–2773 (2005).

[45] P. Perona, and J. Malik, "Scale-space and edge detection using anisotropic diffusion," IEEE T. Pattern Anal. Mach. Intell. **12**, 629–639 (1990).