

# A Comparative Study of Some Markov Random Fields and Different Criteria Optimization in Image Restoration

José I. De la Rosa, Jesús Villa, Ma. Auxiliadora Araiza,  
Efrén González and Enrique De la Rosa  
*Laboratorio de Procesamiento Digital de Señales  
Facultad de Ingeniería Eléctrica, Universidad Autónoma de Zacatecas  
Mexico*

## 1. Introduction

The present chapter illustrates the use of some recent alternative methods to deal with digital image filtering and restoration. This collection of methods is inspired on the use of Markov Random Fields (MRF), which introduces prior knowledge of information that will allow, more efficiently, modeling the image acquisition process. The methods based on the MRF are analyzed and proposed into a Bayesian framework and their principal objective is to eliminate noise and some effects caused by excessive smoothness on the reconstruction process of images which are rich in contours or edges. In order to preserve object edges into the image, the use of certain convexity criteria into the MRF is proposed obtaining adequate weighting of cost functions in cases where discontinuities are remarked and, even better, for cases where such discontinuities are very smooth.

Some image analysis and processing tasks involve the filtering or image recovery  $\hat{x}$  (e.g. restoration) after  $x$  passes by a degradation process giving the observed image  $y$  (see equation (1)). Such degradation covers the noise perturbations, blurring effects due to focus of the acquisition system lenses or to the bandwidth capacity, and other artifacts that may be added to the correct data. The use of Bayesian methods proposed in the seventies (Besag, 1974; 1986; Geman and Geman, 1984), are nowadays essential at least in the cases of image segmentation and image restoration (Andrews and Hunt, 1977). 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 MRF into a Bayesian framework. The general approach 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 (Bertaux et al., 2004; Cai et al, 2008; Chan et al., 2006; Durand and Nikolova, 2006a;b; Nikolova, 2006; 2005; Nikolova and Ng, 2005; Pan and Reeves, 2006).

The restoration or recuperation of an image to its original condition given a degraded image, passes thus by reverting the effects caused by noise or / and a distortion function. In fact, the degradation characteristic is a crucial source of information and its mathematical

function must be known or estimated during the inversion procedure. Typically, this is a point spread function (PSF) which can be linked with the probability distribution of the noise contamination or likelihood function. In the case of MAP filters, usually the additive Gaussian noise is considered, however in some applications this noise is non-Gaussian (Bertaux et al., 2004; Cai et al, 2008) or unknown noise probability distribution function (pdf),  $p(e) = p(y|x)$ , with some partial knowledge. In the latter case, we propose to use the data itself to obtain a non-parametric entropy estimate (EE) of the log-likelihood pdf,  $\hat{p}_{n,h}(e)$  (De la Rosa and Fleury, 2002; De la Rosa et al., 2003; De la Rosa, 2004). Then the log-likelihood will be optimized (e.g. minimized) together with a log-MRF to obtain the MAP image estimation. A variety of applications in signal processing and instrumentation, are based in statistical modeling analysis. One of the most used is the linear regression model (simple or multi-variable in the case of images)

$$y_{i,j} = x_{i,j}^T \theta_{i,j} + e_{i,j}, \text{ with } e \sim p(e), \quad (1)$$

where  $y$  represents the response (observed data, acquired data or explained variables), to  $x$  explicative variables for  $i = 1, \dots, N$  and  $j = 1, \dots, M$ , and a system response parameterized by  $\theta$  (being  $\theta$  a linear operator which models the image degradation (or PSF)) which is associated to data  $(y, x)$ . In some applications  $\theta$  are functional parameters which will be estimated by an identification procedure if  $x$  are known, but if  $\theta$  are known and  $x$  are unknown, the estimation is made about  $x$ , or the estimation can be made for both cases (e.g. blind deconvolution). The  $e$  variables are independent random processes identically distributed accordingly to  $p(e)$ . A natural extension of the linear regression model is the non-linear regression model based on a parameterized function  $f(x, \theta)$  (see equation (2)). This function is nonlinear with respect to the parameters, and its use is also considered because it has been shown in a large variety of signal processing and control applications that the modeling could be more realistic when using nonlinear functions. The perturbations affecting the analyzed system are also modeled as stochastic processes. Then, for this case:

$$y_{i,j} = f(x, \theta)_{i,j} + e_{i,j}, \text{ with } e \sim p(e). \quad (2)$$

Another source of information which imposes a key rule in the image processing context, is the contextual or spatial information that represents the correlation between the intensity values of a neighborhood of pixels well specified. The modelling, when using MRF, takes into account such spatial interaction and it was introduced and formalized in (Besag, 1974), and extended in some pioneering works (Besag, 1986; Bouman and Sauer, 1993; Geman and Geman, 1984; Sauer and Bouman, 1992). Combining both kinds of information into a statistical framework, the restoration is led by an estimation procedure given the MAP of the true images when the distortion functionals are known. The algorithms implemented in this chapter 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. Some existent optimization methods are commented, since the optimization (e.g. maximization or minimization) task is central during the reconstruction and its performance depends on the structure of the cost functionals. Finally, a comparison between some existent Bayesian schemes and new proposed for noise filtering and image deblurring is presented. Particularly, five MAP estimation schemes are implemented where the first two schemes use two MRF; the semi-Huber (SH) potential

function which is proposed as an extension of some previous works (De la Rosa and Fleury, 2006; De la Rosa et al., 2007), and the generalized Gaussian MRF (GGMRF) introduced in (Bouman and Sauer, 1993) in both cases, Gaussian noise is assumed. The other three MAP schemes correspond to new MAP entropy estimators (MAPEE) for image filtering, where unknown noise pdf assumptions are made and the generalized Gaussian MRF is also used.

The rest of this chapter is organized as follows. Section 2 describes the general definition of a MRF and the proposal of the MAP estimator. The potential functions must be proposed to conduct adequately the inversion process, some functions are described in section 3 where the convexity level is the key to formulate an adequate criterion to be maximized or equivalently minimized. In sections 4 and 5 are deduced some MAP estimators and commented some optimization techniques resulting from different MRF structures for Gaussian noise assumptions, and MAPEE estimators for unknown noise pdf. Section 6 introduces the Kernel structures used for MAPEE estimators, while in section 7 results for Gaussian and non-Gaussian restoration are discussed. Finally, some conclusions and comments are given in section 8.

## 2. MAP estimation and Markov random fields

The problem of image estimation (e.g. 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 estimator (ML):** A classical procedure to estimate  $x$  when  $\theta$  is known (from equations (1) and (2)), is based in a cost function or criterion  $\mathcal{J}(x)$  which varies in function  $\psi(\cdot)$  of residuals or noise  $e(x)$ , where:

$$e_{i,j}(x) = y_{i,j} - x_{i,j}^T \theta_{i,j} \quad \text{or} \quad e_{i,j}(x) = y_{i,j} - f(x, \theta)_{i,j}, \quad (3)$$

and so

$$\mathcal{J}(x) = \sum_{i=1}^N \sum_{j=1}^M \psi(e_{i,j}(x)). \quad (4)$$

This is, for example, the case of the maximum likelihood (ML) estimator:

$$\begin{aligned} \hat{x}_{\text{ML}} &= \arg \max_{x \in \mathbb{X}} p(y|x) = \arg \min_{x \in \mathbb{X}} (-\log p(y|x)), \\ &= \arg \min_{x \in \mathbb{X}} \left( - \sum_{i=1}^N \sum_{j=1}^M \log p(e_{i,j}(x)) \right). \end{aligned} \quad (5)$$

The ML estimator is optimal when all information about the distribution  $p(e)$  is accessible, which is the case for Gaussian noise. When the knowledge about  $p(e)$  is imprecise or wrong, the estimator  $\hat{x}_{\text{ML}}$  is possibly suboptimal. Moreover, 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 gives both a more effective approach to reduce noise and smoothness (in fact, it depends on the mathematical structure of the regularization term). In a probabilistic framework, the regularization of the ML estimator leads to the 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:**

$$\begin{aligned}\hat{x}_{\text{MAP}} &= \arg \max_{x \in \mathbb{X}} p(x|y), \\ &= \arg \max_{x \in \mathbb{X}} (\log p(y|x) + \log g(x)), \\ &= \arg \min_{x \in \mathbb{X}} (-\log p(y|x) - \log g(x)).\end{aligned}\quad (6)$$

In this case, the estimator is regularized by using a MRF function  $g(x)$  which models all prior information as a whole probability distribution, where  $\mathbb{X}$  is the set of pixels  $x$  capable to maximize  $p(x|y)$ , and  $p(y|x)$  is the likelihood function from  $y$  given  $x$ .

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

$$g(x) = \frac{1}{Z} \exp \left( - \sum_{c \in \mathbb{C}} V_c(x) \right), \quad (7)$$

where  $Z$  is a normalization constant,  $\mathbb{C}$  is a set of "cliques"  $c$  or local neighborhood of pixels, and  $V_c(x)$  is a weighting function given over the local neighborhood. Generally, the "cliques" correspond to the possible set of neighborhoods of pixels. A theorem introduced by Hammersley-Clifford (Besag, 1974; Geman and Geman, 1984) probes the equivalence between the Gibbs distribution and the MRFs. The Markov random fields have the capacity to represent various images sources. The main partial disadvantage using MRFs is that the estimation procedure deals with local minimization schemes (generally it grows the computation time), while the proposal of global minimization schemes has been recently introduced by modifying the local structure of the MRFs in order to establish a better trade-off between quality of reconstruction and computation time. There is a variety of MRF models which depend on the cost functions also known as potential functions that can be used. Each potential function characterizes the interactions between pixels in the same local group. As an example, the following family represents convex functions:

$$\rho(\Delta) = |\Delta|^p \quad (8)$$

where  $\Delta = \lambda[x_s - x_r]$ ,  $\lambda$  is a constant parameter to be chosen, and  $p$  takes constant values such as  $p \geq 1$  accordingly to the theorem 2 in (Bouman and Sauer, 1993).

### 3. MRFs and convexity of potential functions

In recent works (Champagnat and Idier, 2004; Ciuciu and Idier, 2002; Ciuciu et al., 2001; Giovannelli et al., 2002; Idier, 2001; Rivera and Marroquin, 2003; 2004; Rivera, 2005) some new potential functions were introduced. Such proposed functions are semi-quadratic functionals or half-quadratic and they characterize certain convexity into the regularization term (Geman and Reinolds, 1992; Geman and Yang, 1995)(eg. 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 computation time decreases with respect to other proposed schemes as shown in (Chan et al., 2006; Durand and Nikolova, 2006a;b; Nikolova, 2006; 2005; Nikolova and Ng, 2005; Nikolova and Chan, 2007) and

(Labat and Idier, 2006; Labat, 2006) in the case of non-convex and convex potential functions. On the other hand, in previous works from (Gibbs, 2000) it has been proposed a way to obtain the posterior distribution of the images; in this case, it is necessary to use sophisticated stochastic simulation techniques based on the Markov Chain Monte Carlo (MCMC) methods (Neal, 1993; Robert and Casella, 2004). If it is possible to obtain the posterior distribution of any image, thus, it is also possible to sample from such posterior distribution and obtain the MAP estimator, or other estimators such as the median estimator which is sensible to be as well as the MAP estimator under certain circumstances imposed by the noise structure, the MAP and the median estimators search the principal mode of the posterior distribution.

In the present chapter two already reported potential functions are revised and their performance is compared. The generalized Gaussian MRF introduced in (Bouman and Sauer, 1993; Sauer and Bouman, 1992) is compared with respect to semi-Huber functional used in problems of one dimensional estimation in (De la Rosa and Fleury, 2006).

### 3.1 Generalized Gaussian MRF (GGMRF)

If one considers to generalize the Gaussian MRF (when  $p = q = 2$  one has a GMRF, see equation (15)) as proposed by Bouman, where the generalized potential functions correspond to

$$\rho_1(\Delta) = |\Delta|^p, \quad 1 < p < 2 \tag{9}$$

obtaining the GGMRF

$$\log g(x) = -\lambda^p \left( \sum_{s \in S} a_s x_s^p + \sum_{\{s,r\} \in C} b_{sr} |x_s - x_r|^p \right) + ct, \tag{10}$$

where  $ct$  is a constant term, and theoretically  $a_s > 0$  and  $b_{sr} > 0$ ,  $s$  is the site or pixel of interest and  $S$  is the set of sites into the whole MRF, and  $r$  corresponds to the local neighbors. In practice and for Gaussian noise assumptions, it is recommended to take  $a_s = 0$ , assuring the unicity of  $\hat{x}_{MAP}$ , thus the GGMRF can be written as

$$\log g(x) = -\lambda^p \left( \sum_{\{s,r\} \in C} b_{sr} |x_s - x_r|^p \right) + ct, \tag{11}$$

and from equation (2),  $\log p(y|x)$  is strictly convex and so  $\hat{x}_{MAP}$  is continuous in  $y$ , and in  $p$ . On the other hand, the MAPEE approach proposed here takes into account that the noise pdf is unknown or non-Gaussian. For this particular case, it is recommended to take  $a_s = 1$ , since the likelihood term is not given in terms of quadratic  $q = 2$  functional, and in order to relax the possible non-convexity problem it has been used

$$\log g(x) = -\lambda^p \left( \sum_{s \in S} a_s x_s^2 + \sum_{\{s,r\} \in C} b_{sr} |x_s - x_r|^p \right) + ct, \tag{12}$$

also ensuring that  $\log p(y|x)$  is convex and so  $\hat{x}_{\text{MAPEE}}$  is continuous in  $y$ , and in  $p$ . The choice of  $p$  is capital, since it constrains the convergence speed of the local or global estimator, and the quality of the restored image.

### 3.2 Semi-Huber potential function

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 (De la Rosa and Fleury, 2006) for the case of non-linear regression. This function is adjusted in two dimensions according to the following equation:

$$\log g(x) = -\lambda \left( \sum_{\{s,r\} \in \mathbb{C}} b_{sr} \rho_2(x) \right) + ct, \tag{13}$$

where

$$\rho_2(x) = \frac{\Delta_0^2}{2} \left( \sqrt{1 + \frac{4\varphi_1(x)}{\Delta_0^2}} - 1 \right),$$

and where  $\Delta_0 > 0$  is a constant value,  $b_{sr}$  is a constant that depends on the distance between the  $r$  and  $s$  pixels, and  $\varphi_1(x) = e^2$  where  $e = (x_s - x_r)$ . The potential function  $\rho_2(x)$  fulfills the following conditions

$$\begin{aligned} \rho_2(x) &\geq 0, \forall x \quad \text{with } \rho_2(0) = 0, \\ \psi(x) &\equiv \partial \rho_2(x) / \partial x, \text{ exists,} \\ \rho_2(x) &= \rho_2(-x), \text{ is symmetric,} \\ w(x) &\equiv \frac{\psi(x)}{2x}, \text{ exists,} \\ \lim_{x \rightarrow +\infty} w(x) &= \mu, 0 \leq \mu < +\infty, \\ \lim_{x \rightarrow +0} w(x) &= M, 0 < M < +\infty. \end{aligned} \tag{14}$$

## 4. MAP estimators for Gaussian noise and optimization

In this section some estimators are deduced. The single problem of filtering noise to restore an observed signal  $y$  leads to establish the estimators.

### 4.1 Image filtering and optimization

The observation equation could be

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

where  $I$  is the identity matrix. For this particular problem when using the GGMRF and under hypothesis of Gaussian noise with variance  $\sigma_n^2$ , the MAP estimator is given by,

$$\hat{x}_{\text{MAP1}} = \arg \min_{x \in \mathbb{X}} \left\{ \sum_{s \in \mathbb{S}} |y_s - x_s|^q + \sigma^q \lambda^p \sum_{\{s,r\} \in \mathbb{C}} b_{s-r} |x_s - x_r|^p \right\}, \tag{15}$$

the minimization in equation (15) or optimization problem leads to consider various methods (Allain et al., 2006; Chan et al., 2006; Nikolova, 2006; 2005; Nikolova and Ng, 2005; Nikolova and Chan, 2007):

- global iterative techniques such as: the descendent gradient (Nikolova and Chan, 2007), conjugate gradient (Rivera and Marroquin, 2003) (for recent propositions one can consult the work (Labat and Idier, 2006; Labat, 2006)), Gauss-Seidel, Over-relaxed methods, etc.
- local minimization techniques: minimization at each pixel  $x_s$  (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) was not implemented, nor complete half-quadratic scheme as proposed in (Geman and Reynolds, 1992) and in (Geman and Yang, 1995), 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. For example, the local MAP estimator for the GGMRF is given by

$$\hat{x}_{s1} = \arg \min_{x \in \mathbb{X}} \left\{ |y_s - x_s|^q + \sigma^q \lambda^p \sum_{r \in \partial s} b_{r-s} |x_s - x_r|^p \right\}, \quad (16)$$

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 the estimated image

$$\hat{x}_s = \frac{y_s + (\sigma\lambda)^2 \sum_{r \in \partial s} b_{r-s} x_r}{1 + (\sigma\lambda)^2 \sum_{r \in \partial s} b_{r-s}}. \quad (17)$$

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, is difficult to implement in a precise way

$$\hat{x}_s = \text{median}(y_s, x_{r_1}, \dots, x_{r_l}). \quad (18)$$

This criterion is not differentiable at zero and this fact causes instability in the minimization procedure. For intermediate values of  $p$  and  $q$  these estimators become sub-optimal, and the iterated methods can be used to minimize the obtained criterions. Such iterative methods are the sub-gradient, or the Levenberg-Marquardt method of MATLAB 7, the last method was used as optimization technique in this work. For cases where  $q \neq p$ , for example  $q = 2$  and  $1 < p < 2$ , some studies and different propositions of a priori functions have been proposed in (Durand and Nikolova, 2006a;b; Nikolova, 2006; 2005; Nikolova and Ng, 2005; Nikolova and Chan, 2007), particularly in (Durand and Nikolova, 2006a;b; Nikolova, 2006; Nikolova and Ng, 2005) where non-convex regularized least-squares schemes are deduced and its convergence is analyzed.

The local or global condition of the estimator depends thus on

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

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

Moreover, when the noise is Gaussian, the value of  $q = 2$  is a good choice, but if the noise is non-Gaussian, and if the structure of noise is known, thus the likelihood term changes giving a particular estimator such as that proposed in (Bertaux et al., 2004), or in (Cai et al, 2008), giving some properties of robustness to the minimization procedure. If the structure of noise is supposed unknown, still one could reconsider the GGMRF with values for  $q \in (1, 2]$ , or consider also another type of MRFs as those described in some works of Idier (Champagnat and Idier, 2004; Ciuciu and Idier, 2002; Ciuciu et al., 2001; Giovannelli et al., 2002) and Nikolova (Nikolova, 2006; 2005; Nikolova and Ng, 2005; Nikolova and Chan, 2007).

A second MAP estimator can be obtained when using the SH potential function, the global estimator can be described by the equation

$$\hat{x}_{\text{MAP2}} = \arg \min_{x \in \mathbb{X}} \left\{ \sum_{s \in \mathbb{S}} |y_s - x_s|^2 + \lambda \left( \sum_{\{s,r\} \in \mathbb{C}} b_{s-r} \rho_2(x) \right) \right\}. \quad (19)$$

As in the previous MAP estimator, it has been proposed to implement the local estimator which leads to a similar expression for the first local MAP estimator (see equation (16)), that is

$$\hat{x}_{s2} = \arg \min_{x \in \mathbb{X}} \left\{ |y_s - x_s|^2 + \lambda \left( \sum_{r \in \partial s} b_{s-r} \rho_2(x) \right) \right\}. \quad (20)$$

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 (Champagnat and Idier, 2004; Giovannelli et al., 2002; Idier, 2001; Nikolova and Ng, 2005), another is the minimization procedure which in any sense will regulate the convergence speed as proposed in (Allain et al., 2006; Geman and Reynolds, 1992; Geman and Yang, 1995; Labat and Idier, 2006; Labat, 2006; Nikolova, 2006; 2005; Rivera and Marroquin, 2003), etc.

## 4.2 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

$$y = \Theta x + e, \text{ with } e \sim \mathcal{N}(0, I\sigma_n^2). \quad (21)$$

Now, for the two previous MAP estimators the likelihood term changes, such that,

$$\hat{x}_{\text{MAP}_k} = \arg \min_{x \in \mathbb{X}} \left\{ \sum_{s \in \mathbb{S}} |y_s - \Theta x_s|^2 - \log g(x) \right\}, \quad (22)$$



where the matrix  $\Theta$  is known and given by the following truncated Gaussian blurring function,

$$\theta(i, j) = \exp\left(\frac{-i^2 - j^2}{2\sigma_b^2}\right), \text{ for } -3 \leq i, j \leq 3, \tag{23}$$

as used also in (Nikolova et al., 2010), with  $\sigma_b = 1.5$ . And  $k = 1, 2$  according to the GGMRF, and the SH potential functions. Here, the results were improved combining ideas introduced in a similar Bayesian way by Levin (Levin et al., 2007a;b) adding a Sparse prior (SP) for filtering and then reconstructing the image.

**5. MAP entropy estimators (EE) for unknown noise pdf**

Our proposition for a new MAP scheme is to use a generalized Gaussian MRF introduced in section 3.1, together with three kernel estimators used in (De la Rosa and Fleury, 2002; De la Rosa et al., 2003; De la Rosa, 2004) to obtain cost functionals or criterions based on the entropy of the approximated likelihood function (first term of equation (6))  $\hat{p}_{n,h}(e)$ . Thus,  $-\log p(y|x)$  is built on the basis of the entropy of an estimate (EE) version  $\hat{p}_{n,h}(e)$  of the distribution  $p(e)$ . A first proposition is due to Pronzato and Thierry (Pronzato and Thierry, 2000a;b; 2001). The approximation is obtained using the classical kernel estimators which uses the empirical distribution of the random vector  $e_1(x), \dots, e_n(x)$ . The next expression denote such estimators:

$$\hat{p}_{n,h}(e) = \hat{p}_{n,h}(e|e_1(x), \dots, e_n(x)) = \frac{1}{n} \sum_{i=1}^n K_h(e - e_i). \tag{24}$$

This expression assumes the hypothesis that  $p(e)$  is symmetric, two times differentiable and positive, indeed, it is assumed that  $K(z)$  is a kernel weighted function which satisfies some imposed conditions treated in (Masry, 1983) and subsequently taken back by Devroye (Devroye, 1992; 1989; Devroye and Krzyzák, 1999; Devroye, 1997), Berlinet (Berlinet and Devroye, 1994), and Loader (Loader, 1999) in some of their research work. The bandwidth  $h = h_n$  is given in function of the sample size  $n$ . This parameter could be considered as a sequence of positive numbers that must satisfy:  $h_n \rightarrow 0$  and  $nh_n \rightarrow \infty$  when  $n \rightarrow \infty$ . The strong uniform consistency of  $\hat{p}_{n,h}(e)$  and its convergence toward  $p(e)$ , depend on a convenable procedure of bandwidth selection. A simple and faster procedure which has been retained in this work is the technique proposed and developed by Terrell (Terrell, 1990; Terrell and Scott, 1985). In the two dimensional kernel cases, the previous idea has been extended in this work accordingly to the equation:

$$\begin{aligned} \hat{p}_{n,h}(e) &= \hat{p}_{n,h}(e|e_{1,1}(x), \dots, e_{n,n}(x)) \\ &= \frac{1}{n^2} \sum_{k=1}^n \sum_{l=1}^n K_h(e - e_{k,l}). \end{aligned} \tag{25}$$

If the convergence and consistence of  $\hat{p}_{n,h}(e)$  are assumed, such that  $\hat{p}_{n,h}(e) \rightarrow p(e)$ , then the entropy criterion over  $\hat{p}_{n,h}(e)$  can be approximated to  $-\log p(y|x)$ . The fact that the entropy of any probability density function is invariant by translation, leads to consider one practical artifact to build a suitable criterion. An extended criterion  $\hat{p}_{n,h}(e_E)$  is based on the residuals or noise extended vector which is given by:  $e_E = \{e_{1,1}(x), \dots, e_{n,n}(x), -e_{1,1}(x), \dots, e_{n,n}(x)\}$

and on a suitable choice of  $h$ :

$$\mathcal{J}_e(x) = H_A(\hat{p}_{n,h}(e_E)) \approx -\log p(y|x), \quad (26)$$

where  $H_A(f) = -\int_{-A_n}^{A_n} f(x) \log f(x) dx$ . Finally, if we assume that the EE is a version of the log-likelihood function into the MAP estimator, then a general version of the MAP-entropy estimator (MAPEE) which assumes unknown noise pdf can be constructed from the fact that  $-\log p(y|x)$  can be approximated by the entropy of an estimate version  $\hat{p}_{n,h}(e)$  of the distribution  $p(e)$ , that is  $H_A(\hat{p}_{n,h}(e_E))$ , thus:

$$\hat{x}_{\text{MAPEE}} = \arg \min_{x \in \mathcal{X}} \{H_A(\hat{p}_{n,h}(e_E)) - \log g(x)\}. \quad (27)$$

Now, substituting a particular  $\log g(x)$  such as the GGMRF into the equation (27) one could obtain at least three MAPEE estimators given by

$$\hat{x}_{\text{MAPEE}_m} = \arg \min_{x \in \mathcal{X}} \left\{ H_A(\hat{p}_{n,h}(e_E)) - \lambda^p \left( \sum_{s \in \mathcal{S}} a_s x_s^2 + \sum_{\{s,r\} \in \mathcal{C}} b_{sr} |x_s - x_r|^p \right) \right\}, \quad (28)$$

for  $m = 1, 2, 3$  according to the three following kernels: Normal, cosine and Hilbert.

## 6. Kernel structures

A function of the form  $K(z)$  is assumed as a fixed kernel  $K_h(z) = 1/(h^d)K(z/h)$ , where  $h$  is called the kernel bandwidth and  $h > 0$ . The fundamental problem in kernel density estimation lies in both the selection of an appropriate value for  $h$  and the selection of the kernel structure. The choice of  $K(z)$  could depend on the smoothness of  $p(e)$ . Three different kernels or nonparametric schemes are reviewed in this section to approximate  $\hat{p}_{n,h}(e_E)$ . The gaussian kernel, which has proved to give good performance when  $h$  is selected by using the over-smoothed principle introduced by Terrell (Terrell, 1990; Terrell and Scott, 1985) when the errors vector is generally of finite length  $n$ . Second kernel is the cosine based weights functions, where  $h$  is viewed in a different way. Finally, the third kernel is obtained from a recent class of Hilbert kernels (Devroye and Krzyżak, 1999). It avoids the bandwidth  $h$  selection and its performance depends on other parameters, whose selection is very easy (parameters  $d$  and  $k$  are defined in section 6.3).

### 6.1 Normal or gaussian kernel

Among the different classical kernels (Berliner and Devroye, 1994), the gaussian kernel has been chosen since it is one of the most popular and easy to implement estimator. The following expression resumes this estimator by a sum of exponentials:

$$\hat{p}_{n,h}(z) = \frac{1}{n^2 h \sqrt{2\pi}} \sum_{k=1}^n \sum_{l=1}^n \exp\left(-\frac{(z - z_{k,l})^2}{2h^2}\right). \quad (29)$$

In such a case, and considering that a fixed kernel structure has been chosen, Terrell (Terrell, 1990) proposes to use an over-smoothed bandwidth  $h$  that corresponds to:

$$h = 3 \left( \frac{1}{2\sqrt{\pi}(35)} \right)^{\frac{1}{5}} \sigma n^{-\frac{1}{5}},$$

this bandwidth value guarantees the minimization of the Mean Integrated Squared Error (MISE),  $\sigma$  is the standard deviation of the sample  $z$ , and  $\int K(z)^2 dz = \frac{1}{(2\sqrt{\pi})}$ . Under mild conditions, the kernel density estimates based on the over-smoothing principle are consistent and for sufficiently large sample sizes, they will display all information present in the underlying errors density  $p(e)$ , we chose  $n \times n$  array of samples  $z_{k,l}$  near  $z$ .

### 6.2 Cosine based weights functions

A sequence of special weight cosine functions

$$c_q(z) = 1/(A_q) \cos^{2q}(z/2)$$

requires only  $O(q^2n)$  arithmetic operations, where  $q$  is a slowly growing function that increases without bound together with  $n$  (i.e.  $q(n) < n$  or  $q(n) > n$ ). This estimator is similar to the kernel classical estimators and evaluates a series expansion in a more efficient way. The role of  $h$  is replaced by  $q$  ( $c_q(z) \equiv K_h(z)$ ) and the efficiency is attained (according to the time of calculus), since the selection of  $q$  is generally more simple than  $h$ . Thus (25) is equivalent to

$$\hat{p}_{n,q}(z) = \frac{1}{n^2 A_q} \sum_{k=1}^n \sum_{l=1}^n \left( \frac{1 + \cos(z - z_{k,l})}{2} \right)^q, \tag{30}$$

where the  $A_q$  value could be approximated by

$$A_q \approx \frac{2\sqrt{\pi}}{\sqrt{q}}.$$

Sufficient consistency and convergence conditions are given in (Eğecioğlu and Srinivasan, 2000) for the one dimensional case.

### 6.3 The Hilbert kernel

Finally, other kernel density estimates called the Hilbert kernel estimate is used. The equation  $K_h(z) = 1/(h^d)K(z/h)$  is considered equivalent to  $K(u) = 1/\|u\|^d$ , where the smoothing factor  $h$  is canceled obtaining:

$$\hat{p}_n(z) = \frac{1}{n^2} \sum_{k=1}^n \sum_{l=1}^n \frac{1}{\|z - z_{k,l}\|^d}. \tag{31}$$

The Hilbert estimates are viewed as a universally consistent density estimate whose expected performance ( $L_1, L_\infty$ , pointwise) is monotone in  $n$  (at least in theory) for all densities. The consistency of this class of estimators is proved in Devroye and Krzyżak (1999)(see theorem 2). The Hilbert density estimate of order  $k$  ( $k > 0$ ) is a redefined subclass that avoids the

infinite peaks produced during estimation, in one dimensional case and using the value of  $k = 2$  the kernel estimate is given by:

$$\hat{p}_n(z) = \sqrt{\frac{4}{V_d^2 \pi n(n-1) \log n} \sum_{1 \leq i < j \leq n} \frac{1}{\text{Den}_{i,j}}}, \quad (32)$$

where  $\text{Den}_{i,j} = \|z - z_i\|^{2d} + \|z - z_j\|^{2d}$  and  $V_d$  is the volume of the unit ball in  $\mathbb{R}^d$ . This last expression is also called Cauchy density estimate, due to its similarity to the multivariate Cauchy density,  $\|\cdot\|$  denotes the  $L_2$  metric on  $\mathbb{R}^d$ . Finally, it is assumed that  $\hat{p}_n(z) \rightarrow p(z)$  at least in probability for almost all  $z$ . For a suitable choice of  $A_n$  and alternatively of  $h_n$ ,  $q$ , or  $d$  and  $k$ , these estimators could be "blind asymptotically efficient". The asymptotic properties and the strong consistency of the truncated entropy estimators were analyzed in (Pronzato and Thierry, 2000a;b; 2001). More over, in some other papers, the power of these nonparametric tools have been largely used for different signal processing problems (Pronzato et al., 2004), (Wolsztynski et al., 2004a) – (Wolsztynski et al., 2005c).

## 7. Denoising and deblurring experiments

Results presented in this section were obtained experimenting extensively with four images: synthetic, Lena, Cameraman, and Boat (shown in Figure 1), to probe the performance of the presented estimators.

### 7.1 Image filtering for Gaussian noise

Some estimation results are presented when images are contaminated by Gaussian noise, and there are no other type of distortions (all  $\theta_{i,j} = 1$ ). The first experiment was made considering that  $\sigma_n = 5, 10, 15$ . In the work (De la Rosa et al., 2003) 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:  $e \sim \mathcal{N}(0, I\sigma_n^2)$ , the values of  $\sigma_n$  are given such that the obtained degradation is perceptible and difficult to eliminate. The obtained results were compared using different values for  $p$  and  $\lambda$  preserving  $q = 2$  (MAP1), and different values for  $\Delta_0$  and with  $\lambda = 1$  (MAP2). Generally, with the two MAP estimators the filtering task gives good visual results preserving contours and smoothing noise (see Figures 2 and 3), but the computation time is different between them, the fastest estimator in general is the MAP2 (see image 2(d)), while the slowest is the MAP1 with  $p = 1.2$ . Table 1 shows the performance of the MAP estimators for the problem of filtering Gaussian noise for different images, where an objective evaluation is obtained 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 MAP2, gives better or similar performance with respect to MAP1. Some interesting applications of robust estimation are particularly focused in phase recovery from fringe patterns as presented in the work Villa et al. (2005), phase unwrapping, and some other problems in optical instrumentation. In this sense, some filtering results were thus obtained using the presented MAP estimators with similar results to those reported here. On the other hand, the use of half-quadratic potential functions permits flexibility for the computation time (Durand and Nikolova, 2006a;b; Nikolova and Ng, 2005). But still is a challenge to tune correctly the hyper-parameters to obtain a better performance in the sense

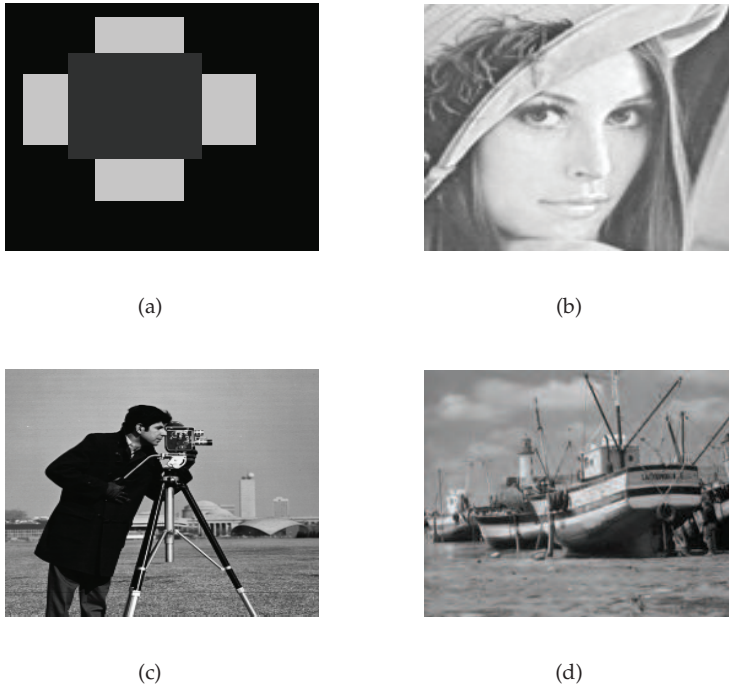


Fig. 1. The figure shows: (a) synthetical probe image; (b) the classical Lena image; (c) the classical Cameraman image; (d) the classical Boat image.

∖	$\sigma_n = 15$	MAP1	MAP2
synthetic	PSNR noise	24.6	24.7
35 × 35	PSNR filt.	27.6	28.8
	Time (sec)	7.6	5.9
Lena	PSNR noise	24.5	24.5
120 × 120	PSNR filt.	28.9	29.1
	Time (sec)	80.6	75.3
Cameraman	PSNR noise	24.7	24.6
256 × 256	PSNR filt.	28.8	28.9
	Time (sec)	355.9	341.7
Boat	PSNR noise	24.6	24.6
512 × 512	PSNR filt.	29.5	29.4
	Time (sec)	1,545.5	1,243.3

Table 1. Results obtained when evaluating the filtering capacity of the two MAP estimators.

of restoration quality, the most simple potential function to tune is the semi-Huber (MAP2). 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

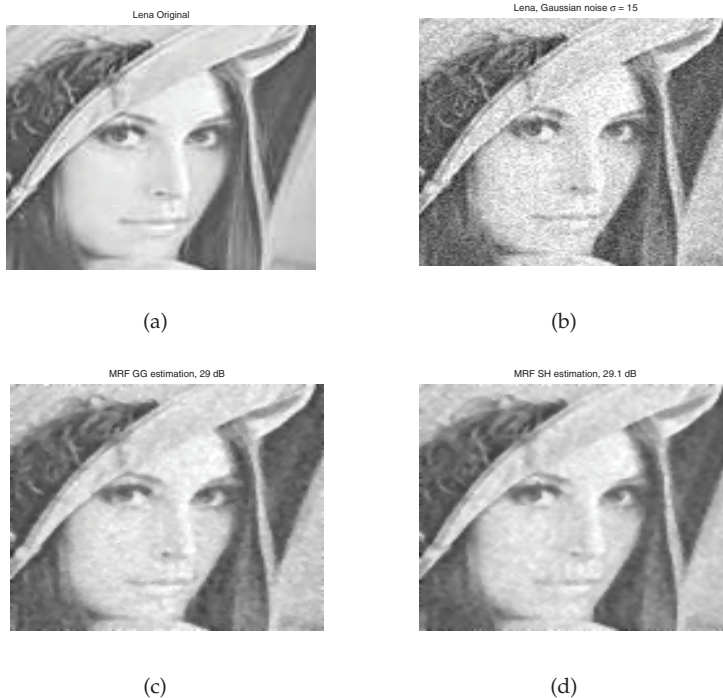


Fig. 2. Results for Lena standard image: (a) describes the original Lena image; (b) describes the noisy image using Gaussian noise with  $\sigma_n = 15$ ; (c) filtered image using MAP1 (with  $\lambda = 30$ ,  $p = 1.2$ ); (d) filtered image using MAP2 (with  $\Delta_0 = 20$ ).

function, as proposed in section 5 for cases of non-Gaussian noise or unknown pdf. Moreover, also the problem could be solved by using variational and partial differential equations as illustrated in the famous work of Perona and Malik (Perona and Malik, 1990), and some recent related works.

## 7.2 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  $\sigma_b = 1.5$ ) blurring the image. This second experiment was carried out considering that  $\sigma_n = 3, 5, 7$ . The results were also compared using different values for  $\Delta_0$  and with  $\lambda = 1$  (MAP2), different values for  $p$  and  $\lambda$  preserving  $q = 2$  (MAP1). Figure 4 shows a comparison of results obtained for the Cameraman image accordingly to the two MAP estimators. One can notice that preserving values of hyper-parameters near those for the filtering case, the estimators smooth the noise but does not made a good recuperation of edges into the image. One must change the hyper-parameters values searching a trade of between the granularity of the noise and the sharpness of the image. Figure 5 shows the results obtained on the Cameraman image when using a combination of proposed estimators together with a Sparse prior (SP) deconvolution technique introduced in (Levin et al., 2007a;b), the improvement in restoration is visible. In

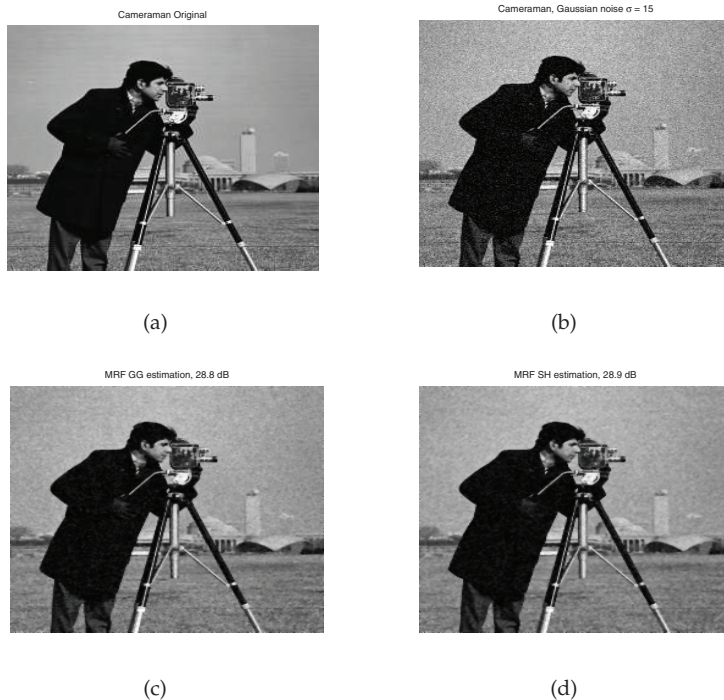


Fig. 3. Results for Cameraman standard image: (a) describes the original Cameraman image; (b) describes the noisy image using Gaussian noise with  $\sigma_n = 15$ ; (c) filtered image using MAP1 (with  $\lambda = 30$ ,  $p = 1.2$ ); (d) filtered image using MAP2 (with  $\Delta_0 = 20$ ).

the first stage only the SP deconvolution is obtained (see image 5(b)), where some ringing and granularity artifacts occur due to noise and they cannot be correctly smoothed. Then, applying first the filtering of noise and then the SP deconvolution, the images 5(c) and 5(d) are obtained. In Table 2 it is shown the performance of the MAP estimators together with the SP deconvolution for three images giving similar results. An objective evaluation is made using the PSNR measure and also computation times in MATLAB are shown.

### 7.3 Image filtering for non-Gaussian noise

Now, for the problem of filtering non-Gaussian noise, some estimation results were obtained when images are contaminated by Gamma, Beta, Uniform and impulsive noise. Here, it is assumed that there are not other type of distortions (all  $\theta_{i,j} = 1$ ). The observation equation then could be

$$y = x + e, \text{ where } e \sim \mathcal{G}(\alpha, \beta), e \sim \mathcal{B}(\alpha, \beta), \dots$$

A third experiment was made considering Gamma noise where  $\alpha = 1.5, 3.5, 7$  and  $\beta = 1, 2, 3$ , and also two factors of amplification of noise were used  $\sigma_a = 5, 10$  ( $\sigma_a \mathcal{G}(\alpha, \beta)$ ). The values of  $\alpha$  and  $\beta$  are given in such a way that the obtained degradation is perceptible and difficult to eliminate.





Fig. 4. Results for Cameraman standard image: (a) describes the original Cameraman image; (b) describes the distorted image using Gaussian noise with  $\sigma_n = 3$ ; (c) partially restored image using MAP1 ( $\lambda = 30, p = 1.2$ ); (d) partially restored image using MAP2 ( $\Delta_0 = 20$ ).

\	$\sigma_n = 3$	MAP1	MAP2
Lena 120 × 120	PSNR distorted	17.4	17.4
	PSNR rest.	17.6	17.5
	PSNR rest. SP	20.8	20.8
	Time (sec)	91.3	58.9
Cameraman 256 × 256	PSNR distorted	19.3	19.3
	PSNR rest.	19.4	19.4
	PSNR rest. SP	22.3	22.5
	Time (sec)	408.0	257.7
Boat 512 × 512	PSNR distorted	20.4	20.4
	PSNR rest.	20.5	20.5
	PSNR rest. SP	25.8	25.6
	Time (sec)	1,606.5	1,014.9

Table 2. Results obtained when evaluating the deconvolution capacity of the two MAP estimators.



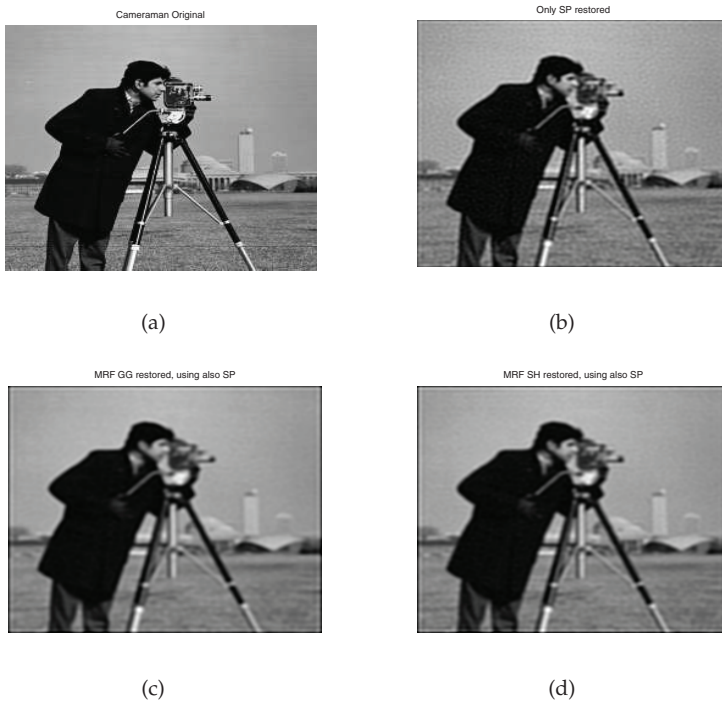


Fig. 5. Results for Cameraman standard image: (a) describes the original Cameraman image; (b) describes restored image using only a Sparse prior (SP) deconvolution technique Levin et al. (2007a); (c) restored image using MAP1 and SP ( $\lambda = 0.15, p = 1.2$ ); (d) restored image using MAP2 and SP ( $\Delta_0 = 20$ ).

Figure 6 shows the filtering of the synthetical image using the three MAPEE estimators introduced in section 5 for parameters  $\lambda = 10, p = 1.2$  of the GGMRF. Generally, with the three estimators the filtering task gives good performing results (see also Figure 7 where a three dimensional-3D view is shown), but the computation times are different between them, the fastest estimator was MAPEE3 (image 6(d)). Some objective measurements were obtained for all probe images, such as the PSNR depicted in Table 3 which agree with the visual results. Also all times of computation are shown in Table 3. On the other hand, in the filtering case of the Cameraman and the Boat images the results coincide with those obtained for synthetic and Lena images. Visually one can see in Figures 8 and 9 the robustness performance of MAPEE estimators. Some other experiments were conducted for Beta and impulsive noise. For the case of Beta noise generation we choose  $\alpha = 1.5, 2.5, 5, 7.5$  and  $\beta = 1$ , and also two factors of amplification of noise were used  $\sigma_a = 10, 20$  ( $\sigma_a \mathcal{B}(\alpha, \beta)$ ). The filtering results are promising for non-Gaussian noise, the obtained results for the Beta noise filtering also confirm the robustness of the proposed approach. Moreover, as one can see for impulsive noise with a 15% of outliers in Figures 10 and 11, the robustness is assured when using the three MAPEE estimators, where for instance, in the synthetic image case the PSNR was improved from 11.45 dB to 17.1 dB.

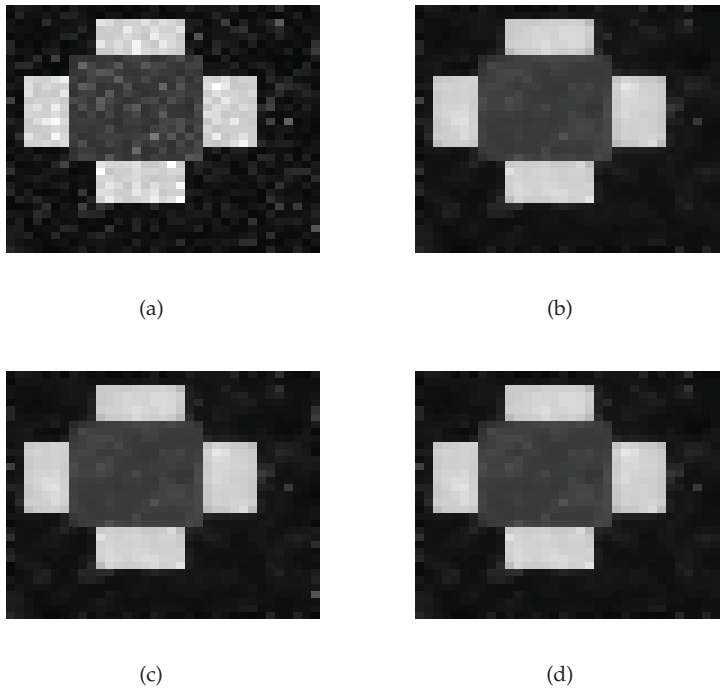


Fig. 6. Results for synthetical probe image: (a) describes the noisy synthetical image, for Gamma noise with  $\alpha = 1.5$  and  $\beta = 2$ , and  $\sigma_a = 5$ ; (b) filtered image using MAPEE1 (with Normal kernel); (c) filtered image using MAPEE2 (with cosine kernel); and, (d) filtered image using MAPEE3 (with Hilbert kernel).

\	$\alpha = 1.5, \beta = 2$	MAPEE1	MAPEE2	MAPEE3
synthetic	PSNR noise	22.2	22.4	22.6
35 × 35	PSNR filtered	24.1	24.1	24.4
	Time (sec)	10.6	12.1	9.3
Lena	PSNR noise	16.3	16.2	16.3
120 × 120	PSNR filtered	17.9	17.9	18
	Time (sec)	123.6	141.6	109.8
Camerman	PSNR noise	16.4	16.4	16.4
256 × 256	PSNR filtered	18.5	18.5	18.5
	Time (sec)	563.2	644.8	499.9
Boat	PSNR noise	16.3	16.4	16.4
512 × 512	PSNR filtered	18.7	18.7	18.8
	Time (sec)	2,340.1	2,569.8	2,063.5

Table 3. Results obtained in evaluating the filtering capacity of the three MAPEE estimators for Gamma  $\mathcal{G}(\alpha, \beta)$  noise, with  $\sigma_a = 10$ .

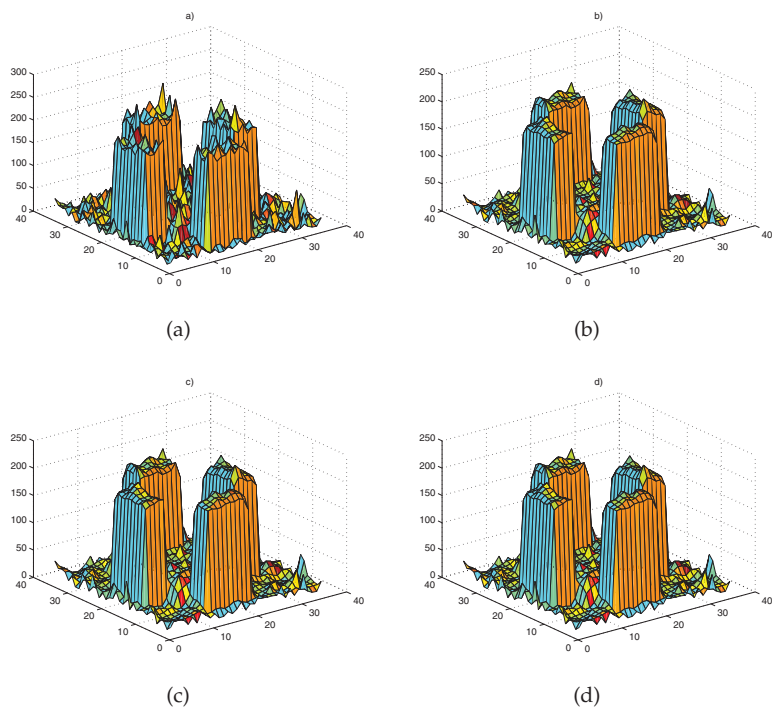


Fig. 7. Results for synthetic image (3-D view): (a) describes the noisy synthetic image, for Gamma noise; (b) filtered image using MAPEE1 (with Normal kernel); (c) filtered image using MAPEE2 (with cosine kernel); and, (d) filtered image using MAPEE3 (with Hilbert kernel).

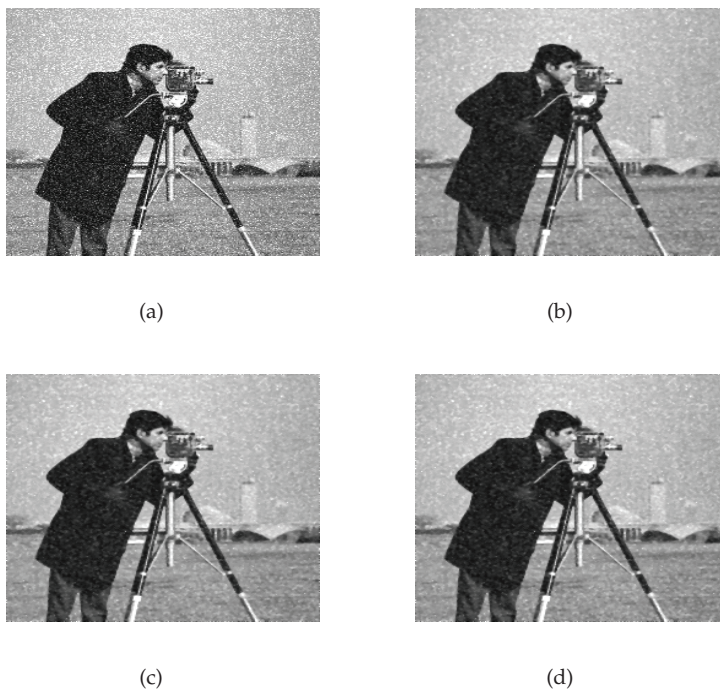


Fig. 8. Results for Cameraman image: (a) describes the noisy Cameraman image, for Gamma noise with  $\alpha = 3.5$  and  $\beta = 2$ , and  $\sigma_n = 10$ ; (b) filtered image using MAPEE1 (with Normal kernel); (c) filtered image using MAPEE2 (with cosine kernel); and, (d) filtered image using MAPEE3 (with Hilbert kernel).



(a)



(b)



(c)



(d)

Fig. 9. Results for Boat image: (a) describes the noisy Boat image, for Gamma noise; (b) filtered image using MAPEE1 (with Normal kernel); (c) filtered image using MAPEE2 (with cosine kernel); and, (d) filtered image using MAPEE3 (with Hilbert kernel).

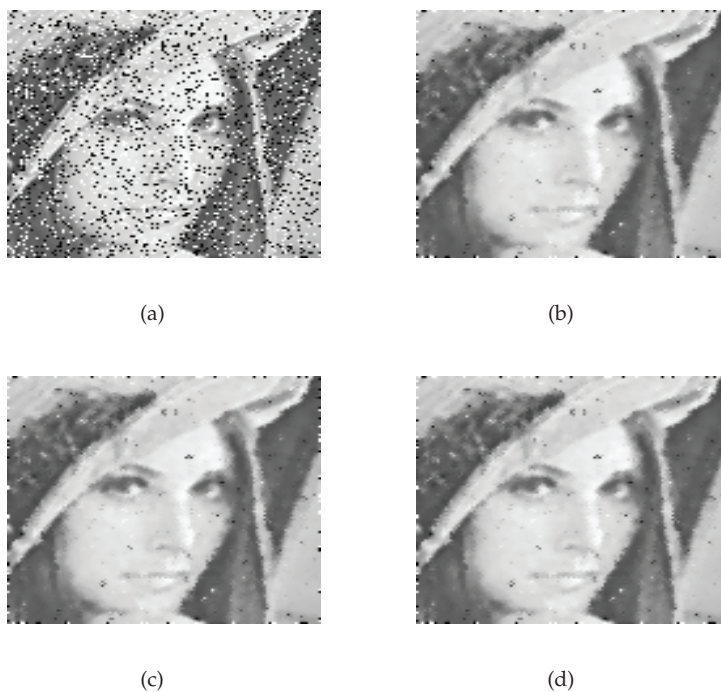


Fig. 10. Results for Lena image: (a) describes the noisy Lena image, for impulsive noise; (b) filtered image using MAPEE1 (with Normal kernel); (c) filtered image using MAPEE2 (with cosine kernel); and, (d) filtered image using MAPEE3 (with Hilbert kernel).

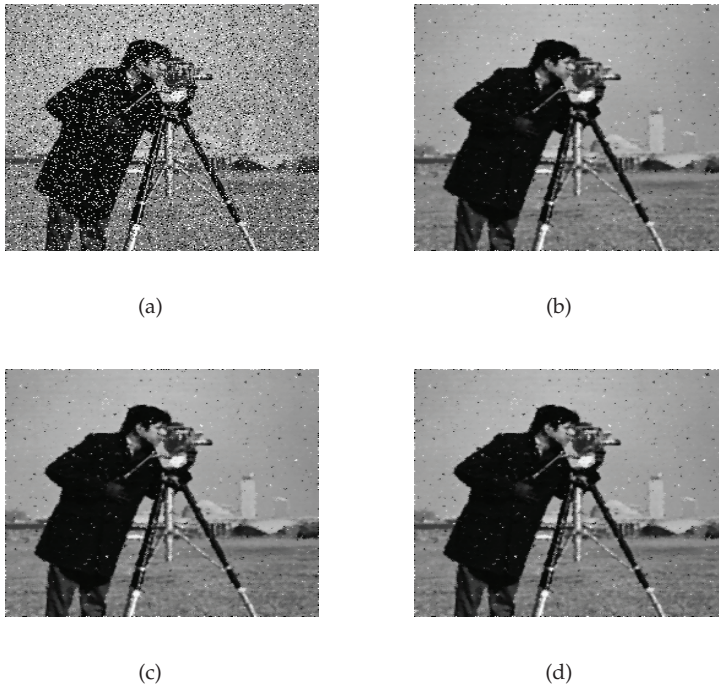


Fig. 11. Results for Cameraman image: (a) describes the noisy Cameraman image, for impulsive noise; (b) filtered image using MAPEE1 (with Normal kernel); (c) filtered image using MAPEE2 (with cosine kernel); and, (d) filtered image using MAPEE3 (with Hilbert kernel).



## 8. Conclusions and comments

Some advantages on the use of GGMRF as prior information into the Bayesian estimation (MAP1) 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 assumptions. The edge preserving is also assured, over all when  $p \rightarrow 1$  and obviously it depends on the choice between the interval  $1 < p < 2$ . The use of other semi-quadratic or half-quadratic potential functions such as SH (MAP2) also presents some advantages: the convexity of these functions is relaxed with respect to the GGMRF. This fact gives as a result the decrement of the computation time, which is a good advantage over the GGMRF, since the tuning of hyper-parameters in the GGMRF is more complicated, one has more degrees of freedom. However, this problem can be solved as argued by Idier (Champagnat and Idier, 2004) and Rivera (Rivera and Marroquin, 2003) by implementing more sophisticated algorithms. The robustness in both cases is assured. Moreover, the advantages presented by those estimators could overcome the disadvantages, by experimenting largely with them, and establishing the aim to implement more sophisticated algorithms with the compromise to reduce computation time and better quality in restoration as recently exposed by Ruimin Pan (Pan and Reeves, 2006).

On the other hand, the performance and improvement of the MAPEE estimators for non-Gaussian noise filtering, could be classified in terms of simplicity and in terms of filtering quality. The obtained results for the three proposed estimators are favorable in general in the sense of robustness, and the fastest convergence is obtained for the MAPEE1 and MAPEE3 estimators (in the case of the MAPEE2 estimator  $q = 40$ , and for MAPEE3,  $k = 2$ ,  $d = 1$ , and  $V_d = 0.7071$ . The value of  $q$  was changed alternatively in the range of  $30 \leq q \leq 90$ , and some performance improvement has been noticed). A general scheme for MAPEE estimators has been introduced and it works in real frameworks, where the nonlinearity conditions of some systems could be present. For future works it exists the interest to implement procedures of Bayesian estimation into high level programming that will be characterized into algorithms to be used in DSP cards, and tasks such as image reconstruction and segmentation. Also, one can probe to use MAPEE to processing real signals issued from real instrumentation or signal processing problems. The final objective of this work has been to contribute with a series of software tools for image analysis, focused for instance, in optical instrumentation tasks such as those treated in the works (Villa et al., 2005)–(Villa et al., 2010b) and (Rivera and Marroquin, 2004; Rivera, 2005) obtaining competitive results in filtering and reconstruction.

## 9. Acknowledgements

We would like to express our deep gratitude to the Secretaría de Educación Pública (SEP) of Mexico, this work was partially supported through the PIFI Mexican Program (PIFI 2010-2011).

## 10. References

- Allain, M.; Idier, J. and Goussard, Y. (2006). On global and local convergence of half-quadratic algorithms, *IEEE Trans. Image Processing*, Vol. 15, No. 5, pp. 1130–1142.
- Andrews, H. C. and Hunt, B. R. (1977). *Digital image restoration*, Prentice-Hall, Inc., New Jersey.



- Berlinet, A. and Devroye, L. (1994). A comparison of kernel density estimates, *Publications de l'Institut de Statistique de l'Université de Paris*, Vol. 38, No. 3, pp. 3–59.
- Bertaux, N.; Frauel, Y.; Réfrégier, P. and Javidi, B. (2004). Speckle removal using a maximum-likelihood technique with isoline gray-level regularization, *J. Opt. Soc. Am. A*, Vol. 21, No. 12, pp. 2283–2291.
- Besag, J. E. (1974). Spatial interaction and the statistical analysis of lattice systems, *J. Royal Stat. Soc. Ser. B*, Vol. B-36, pp. 192–236.
- Besag, J. E. (1986). On the statistical analysis of dirty pictures, *J. Royal Stat. Soc. Ser. B*, Vol. B-48, pp. 259–302.
- Bouman, C. and Sauer, K. (1993). A generalized Gaussian image model for edge-preserving MAP estimation, *IEEE Trans. on Image Processing*, Vol. 2, No. 3, pp. 296–310.
- Cai, J.-F.; Chan R. H., and Nikolova, M. (2008). Two-phase approach for deblurring images corrupted by impulse plus Gaussian noise, *Inverse Problems and Imaging*, Vol. 2, No. 2, pp. 187–204.
- Champagnat, K. and Idier, J. (2004). A connection between half-quadratic criteria and EM algorithms, *IEEE Signal Processing Letters*, Vol. 11, No. 9, pp. 709–712.
- Chan, T. F.; Esedoglu, S. and Nikolova, M. (2006). Algorithms for finding global minimizers of image segmentation and denoising models, *SIAM J. on Applied Mathematics*, Vol. 6, No. 5, pp. 1632–1648.
- Ciuciu, P. and Idier, J. (2002). A half-quadratic block-coordinate descent method for spectral estimation, *Journal of Signal Processing*, Vol. 82, pp. 941–959.
- Ciuciu, P.; Idier, J. and Giovannelli, J.-F. (2001). Regularized estimation of mixed spectra using circular Gibbs-Markov model, *IEEE Trans. on Signal Processing*, Vol. 49, No. 10, pp. 2202–2213.
- De Campos Velho, H. F.; Ramos, F. M.; Chalhoub, E. S.; Stephany, S.; Carvalho, J. C. and De Sousa, F. L. (2007). Inverse problems in space science and technology, *Inverse Problems in Science and Engineering*, Vol. 15, No. 4, pp. 359–372.
- De la Rosa, J. I. and Fleury, G. (2002). On the Kernel selection for Minimum-Entropy estimation, *Proc. of the IEEE Instrumentation and Measurement Technology Conference*, Vol. 2, pp. 1205–1210, Anchorage, AK (USA).
- De la Rosa, J. I.; Fleury, G., and Davoust, M.-E. (2003). Minimum-Entropy, pdf approximation and Kernel selection for measurement estimation, *IEEE Trans. on Instrum. Meas.*, Vol. 52, No. 4, pp. 1009–1020.
- De la Rosa, J. I. (2004). Convergence of Minimum-Entropy robust estimators: Applications in DSP and Instrumentation, *Proc. of the XIV International Conference on Electronics, Communications, and Computers - CONIELECOMP'04*, pp. 98–103, Veracruz, Ver. (Mexico).
- De la Rosa, J. I., and Fleury, G. (2006). Bootstrap methods for a measurement estimation problem, *IEEE Trans. on Instrum. Meas.*, Vol. 55, No. 3, pp. 820–827.
- De la Rosa, J. I.; Villa, J. J. and Araiza, Ma. A. (2007). Markovian random fields and comparison between different convex criteria optimization in image restoration, *Proc. XVII International Conference on Electronics, Communications, and Computers - CONIELECOMP'07*, Cholula, Puebla, (Mexico).
- Devroye, L. (1992). A note on the usefulness of superkernels in density estimation, *The Annals of Statistics*, Vol. 20, pp. 2037–2056. URL : // jeff.cs.mcgill.ca/~luc/

- Devroye, L. (1989). The double kernel method in density estimation, *Annales de l'Institut Henri Poincaré*, Vol. 25, pp. 533–580.
- Devroye, L., and Krzyżak, A. (1999). “On the Hilbert kernel density estimate,” *Statistics and Probability Letters*, Vol. 44, pp. 299–308.
- Devroye, L. (1997). Universal smoothing factor selection in density estimation: theory and practice, *Test*, Vol. 6, pp. 223–320.
- Durand, S. and Nikolova, M. (2006). Stability of the minimizers of least squares with a non-convex regularization. Part I: Local behavior, *Journal of Applied Mathematics and Optimization*, Vol. 53, No. 2, pp. 185–208.
- Durand, S. and Nikolova, M. (2006). Stability of the minimizers of least squares with a non-convex regularization. Part II: Global behavior, *Journal of Applied Mathematics and Optimization*, Vol. 53, No. 3, pp. 259–277.
- Eğecioglu, Ö., and Srinivasan, A. (2000). Efficient non-parametric density estimation on the sphere with applications in fluid mechanics, *SIAM Journal on Scientific Computing*, Vol. 22, No. 1, pp. 152–176.
- Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-6, pp. 721–741.
- Geman, D. and Reynolds, G. (1992). Constrained restoration and the recovery of discontinuities, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 14, No. 3, pp. 367–383.
- Geman, D. and Yang, C. (1995). Nonlinear image recovery with half-quadratic regularization, *IEEE Trans. Image Processing*, Vol. 4, No. 7, pp. 932–946.
- Gibbs, A. L. (2000). *Convergence of Markov Chain Monte Carlo algorithms with applications to image restoration*, Ph.D. Thesis, Department of Statistics, University of Toronto, URL : [www.utstat.toronto.edu](http://www.utstat.toronto.edu).
- Giovannelli, J.-F.; Idier, J.; Boubertakh, R. and Herment, A. (2002). Unsupervised frequency tracking beyond the Nyquist frequency using Markov chains, *IEEE Trans. on Signal Processing*, Vol. 50, No. 12, pp. 2905–2914.
- Idier, J. (2001). Convex half-quadratic criteria and interacting auxiliary variables for image restoration, *IEEE Trans. on Image Processing*, Vol. 10, No. 7, pp. 1001–1009.
- Labat, C. and Idier, J. (2006). Convergence of truncated half-quadratic algorithms using conjugate gradient, *technical report*, IRCCyN, August (France).
- Labat, C. (2006). *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, Nantes, December (France).
- Levin, A.; Fergus, R.; Durand, F. and Freeman, W. T. (2007). Image and depth from a conventional camera with a coded aperture, Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory, SIGGRAPH, *ACM Transactions on Graphics*, August.
- Levin, A.; Fergus, R.; Durand, F. and Freeman, W. T. (2007). Deconvolution using natural image priors, Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory, SIGGRAPH.
- Loader, C. M. (1999). Bandwidth selection: classical or plug-in?, *The Annals of Statistics*, Vol. 27, No. 3, pp. 415–438.
- Masry, E. (1983). Probability density estimation from sampled data, *IEEE Trans. on Information Theory*, Vol. IT-29, No.5, pp. 697–709.

- Neal, R. M. (1993). Probabilistic inference using Markov chain Monte Carlo methods, Tech. Rep., CRG-TR-93-1, Department of Computer Science, University of Toronto (CA).  
URL : [www.cs.toronto.edu/~radford](http://www.cs.toronto.edu/~radford).
- Nikolova, M. (2006). *Functionals for signal and image reconstruction: properties of their minimizers and applications*, Research report to obtain the Habilitation à diriger des recherches, CMLA (CNRS UMR 8536), ENS Cachan, Paris (France).
- Nikolova, M. (2005). Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares, *SIAM Journal on Multiscale Modeling and Simulation*, Vol. 4, No. 3, pp. 960–991.
- Nikolova, M. and Ng, M. (2005). Analysis of half-quadratic minimization methods for signal and image recovery, *SIAM Journal on Scientific computing*, Vol. 27, No. 3, pp. 937–966.
- Nikolova, M. and Chan, R. (2007). The equivalence of half-quadratic minimization and the gradient linearization iteration, *IEEE Trans. on Image Processing*, Vol. 16, No. 6, pp. 1623–1627.
- Nikolova, M.; Ng, M. K. and Tam, C.-P. (2010). Fast nonconvex nonsmooth minimization methods for image restoration and reconstruction, *IEEE Trans. on Image Processing*, Vol. 19, No. 12, pp. 3073–3088.
- Pan, R. and Reeves, S. J. (2006). Efficient Huber-Markov edge-preserving image restoration, *IEEE Trans. on Image Processing*, Vol. 15, No. 12, pp. 3728–3735.
- Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 12, No. 7, pp. 629–639.
- Pronzato, L., and Thierry, E. (2000). A minimum-entropy estimator for regression problems with unknown distribution of observation errors, *MaxEnt 2000*, Edited by A. Mohammad-Djafari, American Institute of Physics, pp. 169–180.
- Pronzato, L., and Thierry, E. (2000). A minimum-entropy estimator for regression problems with unknown distribution of observation errors, Tech. Rep. 00-08, Laboratoire I3S, CNRS-Université de Nice-Shopia Antipolis (France).
- Pronzato, L., and Thierry, E. (2001). Entropy minimization of parameter estimator with unknown distribution of observation errors, *Proc. of the IEEE International Conference in Acoustics, Speech and Signal Processing*, Vol. 6, pp. 3993–3996.
- Pronzato, L.; Thierry, E., and Wolsztynski, E. (2004). Minimum entropy estimation in semi-parametric models: a candidate for adaptive estimation?, *Proceedings of the 7th Int. Workshop on Advances in Model-Oriented Design and Analysis*, Kapellerput, Netherlands, 14-18 june, pp. 125–132. A. Di Bucchianico, H. Läuter and H.P. Wynn (eds.), Physica Verlag, Heidelberg.
- Rivera, M. and Marroquin, J. L. (2003). Efficient half-quadratic regularization with granularity control, *Image and Vision Computing*, Vol. 21, No. 4, pp. 345–357.
- Rivera, M. and Marroquin, J. L. (2004). Half-quadratic cost functions for phase unwrapping, *Optics Letters*, Vol. 29, No. 5, pp. 504–506.
- Rivera, M. (2005). Robust phase demodulation of interferograms with open or closed fringes, *J. Opt. Soc. Am. A*, Vol. 22, No. 6, pp. 1170–1175.
- Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods*, Springer Verlag, 2nd Edition.
- Sauer, K. and Bouman, C. (1992). Bayesian estimation of transmission tomograms using segmentation based optimization, *IEEE Trans. on Nuclear Science*, Vol. 39, No. 4, pp. 1144–1152.

- Terrell, G. P. (1990). The maximal smoothing principle in density estimation, *Journal of the American Statistical Association*, Vol. 85, pp. 470–477.
- Terrell, G. P., and Scott, D. W. (1985). Oversmoothed nonparametric density estimation, *Journal of the American Statistical Association*, Vol. 80, pp. 209–214.
- Villa, J. J.; De la Rosa, J. I.; Miramontes, G. and Quiroga, J. A. (2005). Phase recovery from a single fringe pattern using an orientational vector field regularized estimator, *J. Opt. Soc. Am. A*, Vol. 22, No. 12, pp. 2766–2773.
- Villa, J. J.; Quiroga, J. A. and De la Rosa, J. I. (2009). Regularized quadratic cost-function for directional fringe pattern filtering, *Optics Letters*, Vol. 37, No. 11, pp. 1741–1743.
- Villa, J. J.; Rodríguez-Vera, R.; Quiroga, J. A.; De la Rosa, J. I. and González, E. (2010). Anisotropic phase-map denoising using a regularized cost-function with complex-valued Markov-random-fields, *Optics and Lasers in Engineering*, Vol. 48, No. 6, pp. 650–656.
- Villa, J. J.; Quiroga, J. A.; Servín, M.; Estrada, J. C. and De la Rosa J. I. (2010). N-dimensional regularized fringe direction-estimator, *Optics Express*, Vol. 18, No. 16, pp. 16567–16572.
- Wolsztynski, E.; Thierry, E., and Pronzato, L. (2004). Minimum entropy estimation in semi parametric models, *Proc. of the IEEE International Conference in Acoustics, Speech and Signal Processing - ICASSP'2004*, 17-21 may, Montreal, Canada, Vol. II, pp. 1045–1048.
- Wolsztynski, E.; Thierry, E., and Pronzato, L. (2004). Consistency of a Minimum-Entropy Estimator of Location, Research Report I3S/RR-2004-38-FR Projet TOPMODEL, Laboratoire I3S, CNRS-Université de Nice-Shopia Antipolis (France), 30 pages.
- Wolsztynski, E.; Thierry, E., and Pronzato, L. (2005). Estimation semi-paramétrique robuste par minimisation de l'entropie des résidus, application en traitement d'images, *Proc. GRETSI'05*, pp. 149–152, Louvain-la-Neuve (Belgium).
- Wolsztynski, E.; Thierry, E., and Pronzato, L. (2005). Minimum Entropy Estimators in Semiparametric Regression Problems, *Proc. 11th Int. Symp. on Applied Stochastic Models and Data Analysis*, pp. 882–889, 17-20 mai, Brest (France), J. Jansen and P. Lenca (eds.).
- Wolsztynski, E.; Thierry, E., and Pronzato, L. (2005). Minimum entropy estimation in semi parametric models, *Signal Processing*, Vol. 85, pp. 937–949.



## **Advanced Image Acquisition, Processing Techniques and Applications I**

Edited by Dr. Dimitrios Ventzas

ISBN 978-953-51-0342-4

Hard cover, 170 pages

**Publisher** InTech

**Published online** 14, March, 2012

**Published in print edition** March, 2012

"Advanced Image Acquisition, Processing Techniques and Applications" is the first book of a series that provides image processing principles and practical software implementation on a broad range of applications. The book integrates material from leading researchers on Applied Digital Image Acquisition and Processing. An important feature of the book is its emphasis on software tools and scientific computing in order to enhance results and arrive at problem solution.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

José I. De la Rosa, Jesús Villa, Ma. Auxiliadora Araiza, Efrén González and Enrique De la Rosa (2012). A Comparative Study of Some Markov Random Fields and Different Criteria Optimization in Image Restoration, Advanced Image Acquisition, Processing Techniques and Applications I, Dr. Dimitrios Ventzas (Ed.), ISBN: 978-953-51-0342-4, InTech, Available from: <http://www.intechopen.com/books/advanced-image-acquisition-processing-techniques-and-applications-i/a-comparative-study-of-some-markov-random-fields-and-different-criteria-optimization-for-image-resto>

# **INTECH**

open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821