Bayesian nonparemetric MRF and Entropy estimation for robust image filtering

J. I. de la Rosa Vargas, *Member, IEEE*, O. Gutíerrez, J. J. Villa, E. González,
E. de la Rosa, and Gilles Fleury

*Universidad Autónoma de Zacatecas - Unidad Académica de Ingeniería Eléctrica,
Lab. de Procesamiento Digital de Señales, Av. López Velarde No.* 821*, Col. Centro,
C.P. 98068, Zacatecas, Zacatecas.*
*e-mail:*ismaelrv@ieee.org

Abstract

We introduce an approach for image filtering in a Bayesian framework. In this case, the probability density function (pdf) of the likelihood function is approximated using the concept of non-parametric or kernel estimation. The method is complemented using Markov random fields, for instance the Semi-Huber Markov random field (SHMRF), which is used as prior information into the Bayesian rule, and the principal objective of it is to eliminate those effects caused by the excessive smoothness on the reconstruction process of signals which are rich in discontinuities. Accordingly to the hypothesis made for the present work, it is assumed a limited knowledge of the noise pdf, so the idea is to use a non-parametric estimator to estimate such a pdf and then apply the entropy to construct the cost function for the likelihood term. The previous idea leads to the construction of new Maximum a posteriori (MAP) robust estimators, and considering that real systems are always exposed to continuous perturbations of unknown nature. Some promising results have been obtained from two new MAP entropy estimators (MAPEE) for the case of robust image filtering, where such results have been compared with respect to the classical median image filter.

1. Introduction

The main objective of this investigation is to propose new robust algorithms to deal with signal filtering. The data restoration approaches or recuperation of a signal to its original condition given a degraded signal, passes by reverting the effects caused by noise and some times a distortion functional which must be estimated. One useful idea in Bayesian estimation is to construct a Maximum a posteriori (MAP) of the modes or so called estimator of true data by using Markov random fields (MRF's) [3], [16]. The evolution of this idea has caused the development of algorithms which consider new models of contextual information which is lead by the MRF's and the final aim is the restoration of signals. The idea proposed in this work is based in a robust scheme which could be adapted to reject outliers, tackling situations where noise is present in different forms during the signal acquisition process. In the case of classical MAP filters, usually the additive Gaussian noise hypothesis is considered, however in some applications this noise is non-Gaussian or unknown (with some partial knowledge, that is the case of some optical metrology experiments which give at the output signals with speckle noise) [2]. This is a source of information which imposes a key rule in the signal processing context (the contextual or spatial information in two dimensional signals), that represents the likelihood function or correlation between data, for example, in the context of image processing the intensity values of a well specified neighborhood of pixels. Also, the modelling when using MRF's takes into account such spatial or data interaction and it was introduced and formalized in [3] where it is shown the powerfulness of these statistical tools [4], [5], [6], [16], [22]. The signal modeling in the context of the present work lead us to assume a limited knowledge about the noise pdf $p(\boldsymbol{e})$, (see Eq. (3), where $p(\boldsymbol{e}) = p(y|x)$), so we propose to use the data ($\boldsymbol{e}$) itself to obtain a non-parametric Entropy Estimate (EE) of the log-likelihood pdf ($\widehat{p}_{n,h}(e)$) [8]. Then the log-likelihood will be optimized together with a log-MRF to obtain the MAP signal estimation. A variety of applications in signal processing and instrumentation are based in statistical modelling analysis. One of the most used is the linear regression model

$$\boldsymbol{y} = \boldsymbol{x}^\top \boldsymbol{\theta} + \boldsymbol{e}, \text{ with } \boldsymbol{e} \sim p(\boldsymbol{e}), \tag{1}$$

or the multi-variable model in the case of images

$$y_{i,j} = x_{i,j}^\top \theta_{i,j} + e_{i,j}, \text{ with } e \sim p(e), \tag{2}$$

ISBN: 978-607-95476-6-0

where y represents the response (e.g. observed data, or acquired data), to x explicative variables (e.g. data without distortions) for $i = 1, \ldots, N$ and $j = 1, \ldots, M$, and a system response parameterized by θ which is associated to data (y, x). In some applications θ are functional parameters which will be estimated by an identification procedure if x are known, but if θ 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 noise or residuals e variables are independent random processes identically distributed accordingly to $p(e)$.

The Bayesian formulation is introduced in section 2, where it is also depicted the log-likelihood approximated by Entropy estimation. The proposed nonparametric procedure for the Entropy estimation is led by classical kernel estimators, it will be introduced in section 3. The principal apport of this work is also explained in section 3, where two different MAP-Entropy estimators (MAPEE) are proposed and used for the case of image filtering. Section 4 presents a comparison of the MAPEE estimators showing the performance and the improvement of estimation results when one takes into account the change of MRF, moreover the MAPEE estimation is compared with respect to an iterative median filtering. Finally, some concluding remarks are given in section 5.

2. Bayesian filtering and log-likelihood approximated by EE

The problem of signal 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

Maximum A Posteriori (MAP) estimator is given by:

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

in this case, the estimator is regularized by using a Markov random field function (MRF) $g(x)$ which model all prior information as a whole probability distribution, where $\mathbb{X}$ is the set of data x capable to maximize $p(x|y)$ (or minimize $-p(x|y)$), and $p(y|x)$ is the likelihood function from the obtained data y given x.

2.1. Markov random fields

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

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

where Z is a normalization constant, $\mathbb{C}$ is a set of "cliques" c or local neighborhoods of pixels, and $V_c(x)$ is a weighting function given over the local group of points c. Generally, the "cliques" correspond to the sets of neighborhoods of pixels if $\forall s, r \in c$, s and r are neighbors, and one can construct a neighborhood system called ∂s; for the 8 closest neighbors $\partial s = \{r : |s - r| < 2\}$. The Markov random fields have the capacity to represent various image sources.

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 [5], [16].

2.2. Likelihood pdf Entropy estimators (EE)

A classical procedure to estimate x when θ is known (from Eq. (1) and (2)), is based on a cost function or criterion $\mathcal{J}(x)$ which varies in function $\psi(\cdot)$ of residuals or noise $\boldsymbol{e}(x)$, where:

$$\boldsymbol{e}(x) = \boldsymbol{y} - \boldsymbol{x}^\top\boldsymbol{\theta}, \tag{5}$$

and so

$$\mathcal{J}(x) = \sum_{i=1}^{N}\sum_{j=1}^{M} \psi(e_{i,j}(x)). \tag{6}$$

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

$$\widehat{x}_{\mathbf{ML}} = \arg\min_{x\in\mathbb{X}} \left[-\sum_{i=1}^{N}\sum_{j=1}^{M} \log p(e_{i,j}(x))\right]. \tag{7}$$

The ML estimator is optimal when all information about the distribution $p(e)$ is accessible. When the knowledge about $p(e)$ is imprecise or wrong, the estimator $\widehat{x}_{\mathbf{ML}}$ is possibly suboptimal. Moreover, under certain circumstances, in image processing filtering and 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 a more effective approach called Maximum A Posteriori (MAP) estimator which reduces noise and smoothness at the same time. Our proposition for a new MAP scheme is to use both, a Generalized Gaussian MRF introduced by Bouman and Sauer in [5],[22], and the Semi-Huber proposition in [17], together with one of the three kernel estimators used in [8] to obtain cost functionals or criterions based

ISBN: 978-607-95476-6-0

on the entropy of the approximated likelihood function (first term of Eq. (3)) $\widehat{p}_{n,h}(e)$. Thus, $-\log p(y|x)$ is built on the basis of the Entropy of an estimate (EE) version $\widehat{p}_{n,h}(e)$ of the distribution $p(\boldsymbol{e})$. A first proposition is due to Pronzato [20], [21], and [25].

Thus, the approximation is obtained using the classical kernel estimators which uses the empirical distribution of the random vector $e_1(x), \ldots, e_n(x)$, the next expression denote such estimators:

$$\widehat{p}_{n,h}(e) = \widehat{p}_{n,h}(e|e_1(x), \ldots, e_n(x)) = \frac{1}{n}\sum_{i=1}^{n} K_h\left(e - e_i\right). \tag{8}$$

This expression assumes the hypothesis that $p(e)$ is symmetric, two times differentiable and positive, indeed, it is assumed that $K(\cdot)$ is a kernel weighted function which satisfies some imposed conditions treated in the work of Masry [19] and subsequently taken back by Devroye [11]–[14], Berlinet [1], and Loader [18] 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: $h_n \to 0$ and $nh_n \to \infty$ when $n \to \infty$. The strong uniform consistency of $\widehat{p}_{n,h}(e)$ and its convergence toward $p(e)$, depend on a convenable procedure of bandwidth selection. For instance, a simple and faster procedure to bandwidth selection could be the technique proposed and developed by Terrell [23], [24]. In the two dimensional kernel cases the previous idea has been extended in this work according to the following equation:

$$\begin{aligned}\widehat{p}_{n,h}(e) &= \widehat{p}_{n,h}(e|e_{1,1}(x), \ldots, e_{n,n}(x)) \\ &= \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n} K_h\left(e - e_{k,l}\right).\end{aligned} \tag{9}$$

If the convergence and consistence of $\widehat{p}_{n,h}(e)$ is assumed, such that $\widehat{p}_{n,h}(e) \rightarrow p(e)$, then the entropy criterion over $\widehat{p}_{n,h}(e)$ 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 $\widehat{p}_{n,h}(e_E)$ is based on the residuals or noise extended vector which is given by: $\boldsymbol{e}_E = \{(e_{1,1}(x), \ldots, e_{n,n}(x)), -(e_{1,1}(x), \ldots, e_{n,n}(x))\}$ and on a suitable choice of h:

$$\mathcal{J}_e(x) = \mathrm{H}_A\left(\widehat{p}_{n,h}(\boldsymbol{e}_E)\right) \approx -\log p(y|x), \tag{10}$$

where $\mathrm{H}_A(f) = -\int_{-A_n}^{A_n} f(x)\log f(x)dx$. Finally, if we assume that the EE is a version of the log-likelihood function into the MAP estimator, then a fist 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 $\widehat{p}_{n,h}(e)$ of the distribution $p(e)$, that is $\mathrm{H}_A\left(\widehat{p}_{n,h}(e_E)\right)$, thus:

$$\widehat{x}_{\mathbf{MAPEE}} = \arg\min_{x\in\mathbb{X}}\left\{\mathrm{H}_A\left(\widehat{p}_{n,h}(e_E)\right) - \log g(x)\right\}. \tag{11}$$

The selection among different kernel options, permits the performance improvement of the MAPEE estimators which could be classified in terms of simplicity and in terms of filtering quality, here we choose the Hilbert kernel which is shown in next section.

3. Kernel structure and MRFs

The Hilbert kernel estimate is used [15]. The function $K_h(e) = 1/(h^d)K(e/h)$ is considered equivalent to $K(u) = 1/\|u\|^d$, where the smoothing factor h is canceled obtaining:

$$\widehat{p}_n(e) = \frac{1}{n^2}\sum_{k=1}^{n}\sum_{l=1}^{n}\frac{1}{\|e - e_{k,l}\|^d}. \tag{12}$$

The Hilbert estimates are viewed as a universally consistent density estimate whose expected performance (L_1, L_∞, pointwise) is monotone in n (at least in theory) for all densities. The consistency of this class of estimators is proved in [13](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 the one dimensional case and using the value of $k = 2$ the kernel estimate is given by:

$$\widehat{p}_n(e) = \sqrt{\frac{4}{V_d^2\pi n(n-1)\log n}\sum_{1\leq i<j\leq n}\frac{1}{\mathrm{Den}_{i,j}}}, \tag{13}$$

where $\mathrm{Den}_{i,j} = \|e - e_i\|^{2d} + \|e - e_j\|^{2d}$ and V_d is the volume of the unit ball in $\mathbb{R}^b$. 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 $\widehat{p}_n(e) \to p(e)$ at least in probability for almost all e. For a suitable choice of A_n and alternatively of h_n, 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 [20], [21], [25]. More over, in recent works the powerfulness of these nonparametric tools have been largely used for different signal processing problems [25]. The MAPEE approach proposed here takes into account the proved robustness in presence of outliers of the minimum entropy estimators proposed in [8], [9], [10]. Obtaining now the complete cost functional structure for the $\widehat{x}_{\mathbf{MAPEE}}$ estimator from the point of view of the MRF, the $\log g(x)$ used is based on: i) a generalized Gaussian MRF introduced in [5],[22], and ii) a Semi-Huber MRF function used in [17].

ISBN: 978-607-95476-6-0

3.1. Generalized Gaussian MRF (GGMRF)

If one considers to generalize the Gaussian MRF (when $p = q = 2$ one has a Gaussian MRF), as proposed by Bouman [5], where the generalized potential functions can be limited such as

$$\rho(\Delta) = |\Delta|^p \,, \quad \text{for } 1 \leq p \leq 2 \tag{14}$$

obtaining the GGMRF

$$\log g(x) = -\lambda^p \left(\sum_{s \in \mathbb{S}} a_s x_s^p + \sum_{\{s,r\} \in \mathbb{C}} b_{sr} |x_s - x_r|^p \right) + \text{ct}, \tag{15}$$

where 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 it is recommended to take $a_s = 1$, since the likelihood term is not given in terms of quadratic $q = 2$ functional. In order to relax the convexity problem, the following equation has been used

$$\log g(x) = -\lambda^p \left(\sum_{s \in \mathbb{S}} a_s x_s^2 + \sum_{\{s,r\} \in \mathbb{C}} b_{sr} |x_s - x_r|^p \right) + \text{ct}, \tag{16}$$

and from Eq. (3), $\log p(y|x)$ is strictly convex and so $\widehat{x}_{\mathbf{MAPEE}}$ 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.

3.2. Semi-Huber MRF 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 [7] for the case of nonlinear regression. This proposed function is adjusted in this work in two dimensions according to the following equation:

$$\log g(x) = -\lambda \left(\sum_{\{s,r\} \in \mathbb{C}} b_{sr} \rho_1(x) \right) + \text{ct}, \tag{17}$$

where

$$\rho_1(x) = \frac{\Delta_0^2}{2} \left(\sqrt{1 + \frac{4\varphi_1(x)}{\Delta_0^2}} - 1 \right), \tag{18}$$

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

$$\begin{array}{rl} \rho_1(x) \geq 0, & \forall x \quad \text{with } \rho_1(0) = 0, \\ \psi(x) \equiv \partial \rho_1(x)/\partial x, & \text{exists}, \\ \rho_1(x) = \rho_1(-x), & \text{is symmetric}, \\ w(x) \equiv \frac{\psi(x)}{2x}, & \text{exists}, \\ \lim_{x \to +\infty} w(x) = \mu, & 0 \leq \mu < +\infty, \\ \lim_{x \to +0} w(x) = M, & 0 < M < +\infty. \end{array} \tag{19}$$

Notice that there is not necessary a scale parameter and that the potential function meet all requirements imposed by conditions (19).

Now, substituting these particular Hilbert kernel 2D estimate $\widehat{p}_n(e_E)$ and $\log g(x)$ into the equation (11) one could obtain two MAPEE estimators given by

$$\widehat{x}_{\mathbf{MAPEE}_m} = \arg\min_{x \in \mathbb{X}} \left\{ \mathrm{H}_A \left(\widehat{p}_n(e_E) \right) - \log g(x)_m \right\}, \tag{20}$$

for $m = 1, 2$, according to the two previous MRFs. The obtained results for the two proposed estimators are favorable in general in the sense of robustness.

4. Results for noise filtering

Treating the problem of filtering image noise, some estimation results were obtained for several images which were contaminated by Gamma, Beta, Uniform and impulsive noise, and there was no other type of distortions (all $\theta_{i,j} = 1$ from eq. (2)). The observation equation for this case can be written

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

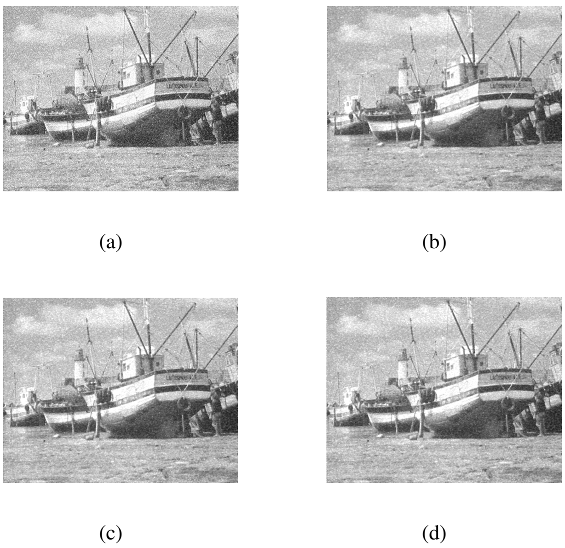
Accordingly to this last equation, one can construct the particular MAPEE estimators as proposed by equation (20) where the iterated method used to minimize the obtained criterions was the Levenberg–Marquardt method of MATLAB 2008 running in a computer with CORE i7 processor, and 4 Gbytes of RAM memory. The first experiment was made considering Gamma noise where $\alpha = 0.5, 1.5, 2.5$ 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 α and β are given such that the obtained degradation is perceptible and difficult to eliminate, Table 1 shows some filtering results in terms of the peak signal to noise ratio (PSNR) according to four probed images and considering that these were contaminated by Gamma noise. Figure 1 also shows some visual results, comparing the MAPEE estimators with respect to the classical median filter. One can see from Table 1 and figure 1, that the MAPEE estimators performance is better than median filtering for the case of Gamma noise. On the other hand, for a fourth experiment, figure 2 shows some results to remove impulsive noise (salt and pepper)

ISBN: 978-607-95476-6-0

$\alpha = 1.5, \beta = 2$	PSNR	MAPEE$_1$	MAPEE$_2$	Median
Im. synthetic	noise	22.2	22.2	22.2
35×35	filt.	24.4	24.2	24.1
Im. Lena	noise	16.3	16.3	16.3
120×120	filt.	17.8	17.7	17.3
Im. Cameraman	noise	16.4	16.4	16.4
256×256	filt.	18.5	18.5	18.0
Im. Boat	noise	16.3	16.3	16.3
512×512	filt.	18.8	18.7	18.5

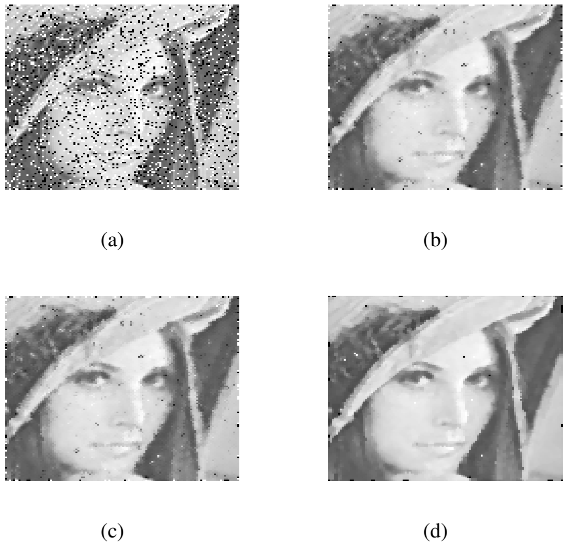
Table 1. PSNR (in dB) obtained on evaluating the filtering capacity of the different MAPEE estimators for Gamma $\mathcal{G}(\alpha, \beta)$ noise, with $\sigma_a = 10$.

for the case of the Lena image, one can see that the performance is similar to the classical median filter (for this type of noise the median filtering is optimal). Also, figure 3 shows similar filtering results for another probe image such as the Cameraman (in both cases, for the kernel estimator it was chosen $k = 2$, $d = 1$, and $V_d = 0.7071$). On the other hand, the choice of the different MRF parameters such as p and Δ_0 can help to improve the filtering results.



(a) (b) (c) (d)

Figure 1. Results for Boat image: (a) describes the noisy image, for Gamma noise; (b) filtered image using MAPEE$_1$ (GGMRF) for $p = 1.5$; (c) filtered image using MAPEE$_2$ (SHMRF) for $\Delta_0 = 10$; and, (d) filtered image using Median filter.



(a) (b) (c) (d)

Figure 2. Results for Lena image: (a) describes the noisy image, for impulsive noise; (b) filtered image using MAPEE$_1$ (GGMRF) for $p = 1.5$; (c) filtered image using MAPEE$_2$ (SHMRF) for $\Delta_0 = 10$; and, (d) filtered image using Median filter.

5. Conclusions

The selection among the different parameter options of the kernel and MRFs, permits the performance improvement of the MAPEE estimators which could be classified in terms filtering quality. The obtained results for the two proposed estimators are favorable in general in the sense of robustness, and it is compared with respect to classical median filtering. A general scheme for MAPEE estimators has been introduced and it was particularized for the case of robust filtering and it was also used for image segmentation. For future works it exists the interest to implement procedures of MAPEE estimation into high level programming that will be characterized into algorithms to be used in DSP cards, and tasks such as image reconstruction (e.g. deconvolution) and segmentation, also one can change the MRF and the optimization procedures to decrease the times of computation.

References

[1] A. Berlinet, and L. Devroye, "A comparison of kernel density estimates," *Publications de l'Intitut de Statistique de l'Université de Paris*, Vol. 38, No. 3, pp. 3–59, 1994.

ISBN: 978-607-95476-6-0

(a)

(b)

(c)

(d)

Figure 3. Results for Cameraman image: (a) describes the noisy image, for impulsive noise; (b) filtered image using MAPEE$_1$ (GGMRF) for $p = 1.5$; (c) filtered image using MAPEE$_2$ (SHMRF) for $\Delta_0 = 10$; and, (d) filtered image using median filter.

[2] 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*, Vol. 21, No. 12, pp. 2283–2291, 2004.

[3] J. E. Besag, "Spatial interaction and the statistical analysis of lattice systems", *J. Royal Stat. Soc. Ser. B*, Vol. B-36, pp. 192–236, 1974.

[4] J. E. Besag, "On the statistical analysis of dirty pictures", *J. Royal Stat. Soc. Ser. B*, Vol. B-48, pp. 259–302, 1986.

[5] C. Bouman, and K. Sauer, "A Generalizaed Gaussian Image Model for Edge-Preserving MAP Estimation", *IEEE Trans. on Nuclear Science*, Vol. 99, No. 4, pp. 1144–1152, 1992.

[6] K. Champagnat, and J. Idier, "A Conection Between Half-Quadratic Criteria and EM Algorithms", *IEEE Signal Processing Letters*, Vol. 11, No. 9, pp. 709–712, Sept. 2004.

[7] J. I. De la Rosa, and G. Fleury, "Bootstrap methods for a measurement estimation problem," *IEEE Trans. Instrum. Meas.*, Vol. 55, no. 3, pp. 820–827, June 2006.

[8] J. I. De la Rosa, G. Fleury, and M.-E. Davoust, "Minimum-entropy, pdf approximation and kernel selection for measurement estimation," *IEEE Trans. on Instrumentation and Measurement*, Vol. 52, No.4, pp. 1009–1020, 2003.

[9] J. I. De la Rosa, "Convergence of minimum-entropy robust estimators: Applications in DSP and instrumentation," *Proc. of the IEEE XIV International Conference on Electronics, Communications and Computers (CONIELE- COMP'04)*, pp. 98–103, 2004.

[10] J. I. De la Rosa, "Convergencia de estimadores a mínimo de entropía robustos: aplicaciones en instrumentación y al PDS," *Computación y Sistemas*, Vol. 10, No. 2, pp. 159–171, 2006.

[11] L. Devroye, "A note on the usefulness of superkernels in density estimation," *The Annals of Statistics*, Vol. 20, pp. 2037–2056, 1992.

[12] L. Devroye, "The double kernel method in density estimation," *Annales de l'Institut Henri Poincaré*, Vol. 25, pp. 533–580, 1989.

[13] L. Devroye, and A. Krzyzȧk, "On the Hilbert kernel density estimate," *Statistics and Probability Letters*, Vol. 44, pp. 299–308, 1999.

[14] L. Devroye, "Universal smoothing factor selection in density estimation: Theory and practice," *Test*, Vol. 6, pp. 223–320, 1997.

[15] Ö. Eğecioğlu, and A. Srinivasan, "Efficent non-parametric density estimation on the sphere with applications in fluid mechanics," *SIAM Journal on Scientific Computing*, Vol. 22, No. 1, pp. 152–176, 2000.

[16] S. Geman, and C. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-6, pp. 721–741, Nov. 1984.

[17] O. Gutiérrez, I. de la Rosa, J. Villa, E. González, and N. Escalante, "Semi-Huber potential function for image segmentation," *Optics Express*, Vol. 20, No. 6, pp. 6542–6554, 2012.

[18] C. M. Loader, "Bandwidth selection: classical or plug-in?," *The Annals of Statistics*, Vol. 27, No. 3, pp. 415–438, 1999.

[19] E. Masry, "Probability density estimation from sampled data," *IEEE Trans. on Information Theory*, Vol. IT-29, No. 5, pp. 697–709, 1983.

[20] L. Pronzato, and E. Thierry, "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, July 8-13, 2000.

[21] L. Pronzato, and E. Thierry, "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, March 2001.

[22] K. Sauer, and C. Bouman, "Bayesian Estimation of Transmission Tomograms Using Segmentation Based Optimization", *IEEE Trans. on Image Processing*, Vol. 2, No. 3, pp. 296–310, July 1993.

[23] G. P. Terrell, "The maximal smoothing principle in density estimation," *Journal of the American Statistical Association*, Vol. 85, pp. 470–477, 1990.

[24] G. P. Terrell, and D. W. Scott, "Oversmoothed nonparametric density estimation," *Journal of the American Statistical Association*, Vol. 80, pp. 209–214, 1985.

[25] E. Wolsztynski, E. Thierry, and L. Pronzato, "Minimum entropy estimation in semi parametric models," *Signal Processing*, Vol. 85, pp. 937–949, 2005.

ISBN: 978-607-95476-6-0