A MAP estimation algorithm using IIR recursive filters ⋆

João M. Sanches⋆⋆ Jorge S. Marques

IST/ISR, Torre Norte, Av. Rovisco Pais, 1049-001, Lisbon, Portugal

Abstract. The MAP method is a wide spread estimation technique used in many signal processing problems, e.g., image restoration, denoising and 3D reconstruction. When there is a large number of variables to estimate, the MAP method often leads to a huge set of linear or non-linear equations which must be numerically solved using time consuming algorithms.
This paper proposes a fast method to compute the MAP estimates in large scale problems, based on the solution of a linear set of equations by low pass filtering the ML solution. A family of space varying IIR filters with data dependent coefficients is derived from the MAP criterion. This approach can be extended to other types of filters derived under different assumptions about the prior or using other design strategies. The filter approach proposed in this paper is much faster than the classic solution and provides additional insights about the structure of the problem.
Experimental results are provided to assess the performance of the proposed methods with Gaussian and non Gaussian noise models.

1 Introduction

The maximum likelihood (ML) method is widely used to estimate signals from noisy data. For instance, in image processing, the ML method is often used to solve problems of image restoration [1], denoising [2], deblurring [3], 2D and 3D reconstruction [4]. In medical imaging, statistical methods [5], e.g., the ML method, became popular techniques especially after the work of Shepp and Vardi on emission tomography [6].

However, the ML estimates tend to be noisy and the convergence of the algorithm is slow due to the ill-conditioned nature of the estimation problem [7]. In the case of 2D or 3D reconstruction, the lack of data is the main difficulty, since we do not have enough information to obtain an accurate estimate for each unknown variable [8]. To overcome these difficulties the ML criterion is often replaced by Bayesian methods (e.g., the MAP method) using prior knowledge about the unknown variables. This leads to an improvement of the numerical

⋆ This work was partially supported by FCT in the scope of project POSI-33726-CPS-2000

⋆⋆ corresponding author: João Sanches, IST/ISR, Torre Notre, Av. Rovisco Pais, 1049-001 Lisboa, Portugal, Email:jmrs@alfa.ist.utl.pt, Phone:+351 21 8418195

stability of the optimization algorithms as to better estimates of the unknown variables if the prior distribution is correctly chosen [9, 10].

Gibbs distributions are one of the most popular choices to represent the prior information. This class of priors leads to simple formulations of the estimation problem. The equivalence with Markov random fields allows to easily obtain the joint probability distribution from a set of local distributions [11].

This paper shows that the MAP estimation of signals can be obtained by filtering the ML estimate with space varying IIR filters. This result is derived assuming a Gibbs prior with quadratic potential functions. The proposed solution is based on two recursive filters with space varying coefficients: a causal IIR filter and an anti-causal IIR filter. It is shown, that the cutoff frequencies of the proposed filters are adaptively adjusted according to the number of observations and to sufficient statistics of the signal.

The MAP filters proposed in this paper are applied to reconstruction problems from noisy observations corruped by Gaussian and non Gaussian (Rayleigh) noise. Experimental results are provided to compare the proposed solution to the classic algorithms.

2 Problem Formulation

Let $X = \{x_i\}$ be a sequence of N unknown variable to be estimated and $Y = \{y_i\}$ a sequence of observations. Each element of Y, y_n is, itself, a set of n_i observations of x_i. In typical problems of image restoration $n_i = 1$, which means that there is one observation per pixel. On the contrary, in 3D reconstruction, the number of observations per voxel varies from voxel to voxel. For instance, in free-hand 3D ultrasound the number of observations associated to non inspected voxels is zero ($n_i = 0$). On the contrary if given voxel is intersected by several cross sections $n_i > 1$.

In this paper the MAP method is used to estimate X from the observations Y. This method estimates X by minimizing an energy function,

$$\hat{X} = arg\min_{U} E(Y, X) \tag{1}$$

where

$$E(X, Y) = -l(X, Y) - \log p(X) \tag{2}$$

$l(X, Y) = \log(p(Y|X)$ is the likelihood function, and $p(X)$ is the a prior distribution associated to the unknown variables.

For sake of simplicity let us assume that y_i is normal distributed (later we will consider other distributions), with $p(y_i) = N(x_i, \sigma^2)$ corresponding to the following observation model

$$y_i = x_i + w_i \tag{3}$$

with $p(w_i) = N(0, \sigma^2)$. If the observations are independent, the log-likelihood function is given by

$$l = C - \frac{\beta}{2} \sum_{i,k} (y_{ik} - x_i)^2 \tag{4}$$

where $\beta = 1/\sigma^2$ and y_{ik} is the kth observation of the unknown x_i.

The prior $p(X)$ used in (2) plays an important role in the estimation process when there is lack of information about the variables X (n_i small), since the ML estimates are very poor in this case [8].

In this paper we will consider that $p(x)$ is a Gibbs distribution with quadratic potential functions [12,11]. This is equivalent to assuming that the vector X is a Markov random field [13,11]. Therefore

$$p(X) = \frac{1}{Z} e^{\sum_i V_i(X)} \tag{5}$$

where Z is the partition function and $V_i(X)$ is the potential function associated to the i-th unknown. Assuming that X is a 1D signal and assuming that each variable x_i has two neighbors, x_{i-1}, x_{i+1},

$$p(X) = \frac{1}{Z} e^{-\frac{\alpha}{2} \sum_i (x_i - x_{i-1})^2} \tag{6}$$

The parameter α defines the strength of the links among neighbors and it is pre-defined. Therefore, the energy function to be minimized is

$$E(Y, X) = \frac{\beta}{2} \sum_{i,k} (y_{i_k} - x_i)^2 + \frac{\alpha}{2} \sum_i (x_i - x_{i-1})^2 \tag{7}$$

The constants C and Z were discarded because they do not contribute to the solution.

3 Optimization

To minimize (7) the following stationary conditions must be met

$$\frac{\partial E(Y, X)}{\partial x_i} = 0 \tag{8}$$

which lead to the following set of linear equations

$$x_i = (1 - k_i) x_i^{ML} + k_i \bar{x}_i \qquad i = 1, ..., N \tag{9}$$

where x_i^{ML} is the maximum likelihood estimation of x_i, k_i is a parameter that depends on the data and $\bar{x}_i$ is the average value of the neighbors of x_i

$$x_i^{ML} = \frac{1}{n_i} \sum_{k=1}^{n_i} y_{i_k}^2 \tag{10}$$

$$k_i = \frac{1}{1 + \frac{\beta n_i}{2\alpha}} \tag{11}$$

$$\bar{x}_i = \frac{x_{i-1} + x_{i+1}}{2} \tag{12}$$

To minimize the border effects it is assumed that $x_0 = x_2$ and $x_{N+1} = x_{N-1}$. Taking this into account, (9) can be written as follows

$$Ax = b \tag{13}$$

where

$$A = \begin{bmatrix} 1 & -k_1 & 0 & 0 & 0 \ldots & 0 & 0 & 0 & 0 \\ -k_2/2 & 1 & -k_2/2 & 0 & 0 \ldots & 0 & 0 & 0 & 0 \\ 0 & -k_3/2 & 1 & -k_3/2 & 0 \ldots & 0 & 0 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & 0 \ldots & -k_{N-2}/2 & 1 & -k_{N-2}/2 & 0 \\ 0 & 0 & 0 & 0 & 0 \ldots & 0 & -k_{N-1}/2 & 1 & -k_{N-1}/2 \\ 0 & 0 & 0 & 0 & 0 \ldots & 0 & 0 & -k_N & 1 \end{bmatrix}$$

and

$$b = [(1 - k_1)x_1^{ML}, (1 - k_2)x_2^{ML}, ..., (1 - k_N)x_N^{ML}]^T \tag{14}$$

The estimation of (13) amounts to the solution of a linear system of equations which can be performed by using either iterative (e.g. Gauss elimination method) or non iterative methods (e.g. Gauss-Seidel method). Since the number of unknowns is often very large (e.g. on the order of a million) iterative methods are preferred since they provide an approximate solution with acceptable computational effort. In order to obtain an efficient solution the structure of A can be considered.

In the next section we will show that the system (13) can be solved using two space varying IIR filters, obtained from (9).

4 IIR filter

Equation (9)

$$x_i = (1 - k_i)x_i^{ML} + \frac{k_i}{2}(x_{n-1} + x_{n+1}) \tag{15}$$

defines a non causal recursive filter [14] with x_i^{ML} as input. Assuming that k_i is constant the filter impulsive response, g_i, can be computed. The general form is

$$g_i = Ca^{|i|} \tag{16}$$

where C and a are computed replacing (16) in (15) and making $x_i^{ML} = \delta(i)$

$$\begin{cases} Ca^0 = (1-k) + \frac{k}{2}(Ca^{|-1|} + Ca^{|1|}) & i = 0 \\ Ca^i = \frac{k}{2}(Ca^{i-1} + Ca^{i+1}) & i \neq 0 \end{cases}$$

$$\begin{cases} C = \frac{1-k}{1-ak} \\ a^2 - \frac{2}{k}a + 1 = 0 \end{cases}$$

leading to

$$\begin{cases} a = \frac{1 \pm \sqrt{1-k^2}}{k} \\ C = \frac{1-k}{-(\pm)\sqrt{1-k^2}} \end{cases}$$

Since $0 \leq k \leq 1$ and $C > 0$, only one solution is feasible, i.e.,

$$\begin{cases} a = \frac{1-\sqrt{1-k^2}}{k} \\ C = \frac{1-k}{\sqrt{1-k^2}} \end{cases}$$

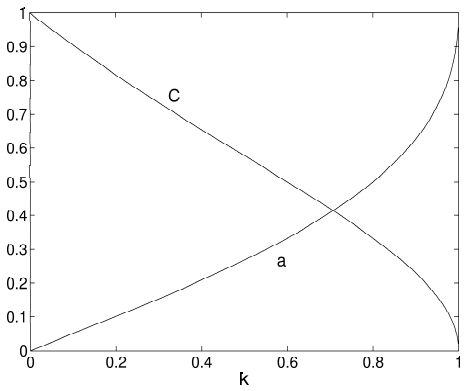

Fig. 1. Impulse response parameters.

Fig.1 shows the dependence of a and C on $k \in [0,1]$. As it can be observed, a is monotonic increasing with k and limited to the interval $[0,1]$.

Fig.2 shows the impulsive response (16) and the solution of (13) for several values of k. As it can be seen, both solutions are identical, except for k close to 1 ($k > 0.99$). These differences are due to border effects. In fact, equation (15) is not valid at $n = 1$ and $n = N$.

The impulsive response defined by (16) can not be used in a recursive way because it is not wedge supported [14]. Therefore, we will decompose it as a sum of two wedge supported impulsive responses, one causal and other anti-causal, i.e., one depending only on past inputs and outputs and other depending only on future inputs and outputs [15]. Therefore,

$$g_n = g_n^+ + g_n^- \tag{17}$$

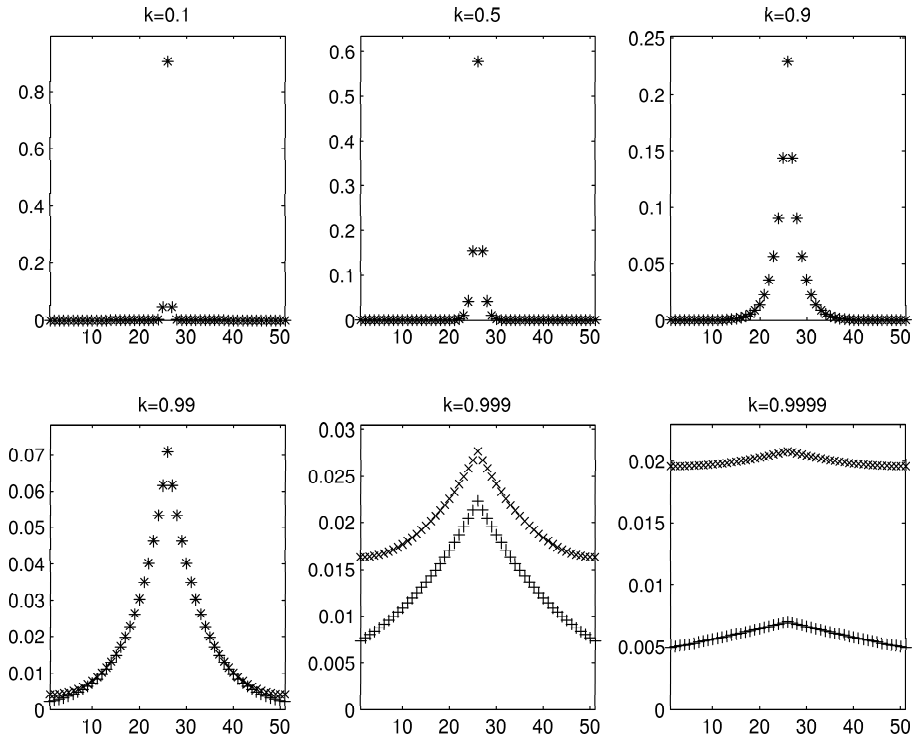

Fig. 2. Impulsive response computed from (13) (x) and from (16) (+).

where

$$g_n^+ = \begin{cases} Ca^n & n > 0 \\ \frac{C}{2} & n = 0 \\ 0 & n < 0 \end{cases}$$

$$g_n^- = \begin{cases} 0 & n > 0 \\ \frac{C}{2} & n = 0 \\ Ca^{-n} & n < 0 \end{cases}$$

where it was assumed $g_0^+ = g_0^-$ to impose symmetry. Applying the Z transform to the previous equation we obtain

$$G(Z) = G(Z)^+ + G(Z)^- \tag{18}$$

$$G(Z)^+ = \frac{C}{2}\frac{1 + aZ^{-1}}{1 - aZ^{-1}} \tag{19}$$

$$G(Z)^- = \frac{C}{2}\frac{1 + aZ}{1 - aZ} \tag{20}$$

The solution of (13) is the sum of two terms

$$x_i = x_i^+ + x_i^- \tag{21}$$

where

$$x_i^+ = g_i^+ * x_i^{ML} = \frac{C_i}{2}(x_i^{ML} + a_i x_{i-1}^{ML}) + a_i x_{i-1} \tag{22}$$

is a causal space varying recursive filter and

$$x_i^- = g_i^- * x_i^{ML} = \frac{C_i}{2}(x_i^{ML} + a_i x_{i+1}^{ML}) + a_i x_{i+1} \tag{23}$$

is an anti-causal space varying recursive filter where

$$\begin{cases} a_i = \frac{1-\sqrt{1-k_i^2}}{k_i} \\ C = \frac{1-k_i}{\sqrt{1-k_i^2}} \end{cases}$$

The MAP estimates defined in (1) can be obtained as follows. First the maximum likelihood estimates X^{ML} is computed. Then the ML estimate are filtered with a causal filter $G(Z)^+$ and with an anti-causal filter $G(Z)^-$. The solution is obtain by adding both results.

5 Frequency analysis

It is now clear that the regularization imposed by the prior is equivalent to filtering the ML estimates with a first order low-pass filter that smoothes the transitions, reducing the noise present in the maximum likelihood estimation.

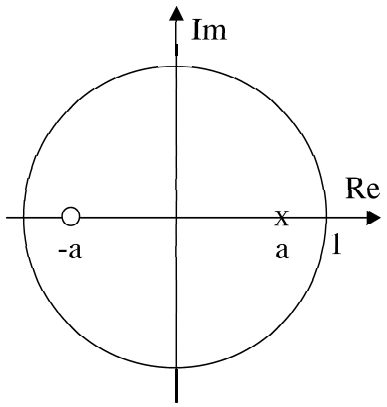

Fig. 3. Pole and zero position of $G(z)$.

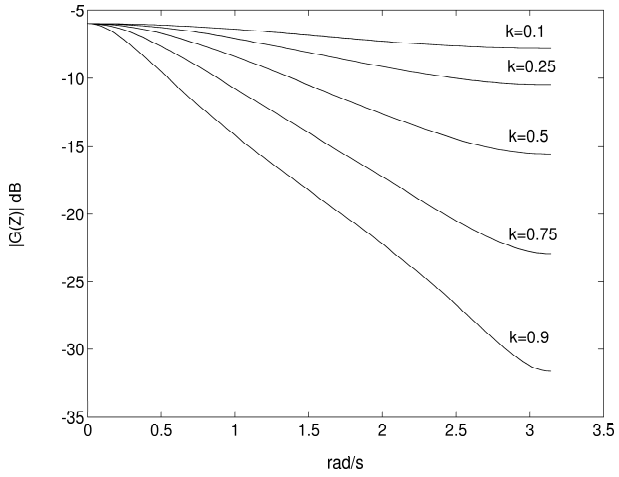

Fig. 4. Bode diagrams of $G(Z)$ for several different values of k.

The low-pass filters (19) and (20) present a 0.5 gain at d.c., a pole located at a and a zero at $-a$ (see Fig.3). The cutoff frequency, depending on the pole position, depends on the data and on the parameter α of the prior as it can be observed on (4) and (11)[1](see Figs.3,4). Therefore, the following conclusions can be stated:

1) since $0 \leq a \leq 1$, the pole of the first order filter is always inside the unit circle and the filter is always stable.
2) the parameter k (parameter a) (see eq.(11)) decreases with the number of data points n_i, i.e., the bandwidth of the filter increases with the amount of available data and decreases as the number of data points goes to zero. The algorithm compensates the lack of confidence in the data by decreasing the filter bandwidth (see Fig.4).
3) the bandwidth of the filter decreases when the regulatization parameter α increases and when the variance of the data, $\sigma^2(x) = 1/\beta$, increases.
4) the algorithm is implemented in such way that when there is no data, $n_i = 0$, $k = 1$ $(a = 1)$. In this case $x_i = \bar{x}_i$, i.e., the estimate depends only on the average value of the neighbors.

Until this point we have been working with an additive Gaussian noise. However, the method can be used with other noise models. For instance, the Rayleigh distribution is often used to model the multiplicative noise present in signals obtained using coherent radiation, e.g., LASER [16], SAAR [17] or ultrasound [18]. In a previous work, the authors have derived expressions similar to (15) in the context of 3D ultrasound by using a Rayleigh model for the multiplicative speckle noise present in the ultrasound images [19]. In the case of the Rayleigh

[1] Note that a(k) is monotonic with k.

observations (15) is still valid with

$$\begin{cases} x_i^{ML} = \frac{1}{2n_i} \sum_k (y_i^k)^2 \\ k_i = \frac{1}{1+\frac{n_i}{2\alpha (x_i^{ML})^2}} \end{cases}$$

In this case k_i depends on the number of observations, n_i, as before, but it also depends on the ML estimate. Therefore an additional property is valid.

5) the bandwidth decreases with the increase of x_i^{ML}, i.e., as stronger regularization is applied for large values of x_i^{ML} than for smaller values (the noise amplitude is larger in high intensity regions). This behaviour is a consequence of the multiplicative type of noise associated to the Rayleigh distribution.

As shown, the MAP estimation problem can be interpreted as a space varying filtering process. Adopting a Gibbs prior with a quadratic potential the MAP estimation process can be implemented by using two first order IIR filters. This approach can also be used to derive other type of filters associated, for instance, to higher order Gibbs priors, which allow the improvement of the estimation performance at transitions (see [20]).

In the next section two examples of application are presented using synthetic and real data.

6 Experimental Results

Experimental tests were performed to evaluate the algorithm in 1D and 2D signal restoration, with synthetic and real data.

Each problem is solved using the standard MAP method and the fast algorithm based on space-varying IIR filters proposed in this paper. Examples with Gaussian and non-Gaussian (Rayleigh) noise are considered.

6.1 Synthetic data

Let us consider a synthetic signal defined as follows,

$$x_i = \begin{cases} 150 & 50 \leq i \leq 100 \\ 50 & otherwise \end{cases}$$

The observation vector Y is obtained by adding Gaussian white noise $\eta_i = N(0, 20^2)$ to each sample (see Fig.5).

The MAP solution was computed by both methods, i.e, by solving (13) and by low pass filtering using (21).

Both solutions are displayed in Fig.5. Since both curves coincide they can not be distinguish. To minimize the border effects, the unknowns x_0 and x_{N+1},

used in equations (22) and (23) respectively were defined as follows

$$x_0 = \frac{1}{2w} \sum_{i=1}^{w} x_i^{ML} \tag{24}$$

$$x_{N+1} = \frac{1}{2w} \sum_{i=N-w}^{N} x_i^{ML} \tag{25}$$

i.e., x_0 and x_{N+1} are initialized with half of the average value of the ML estimates inside window with length $w = 5$.

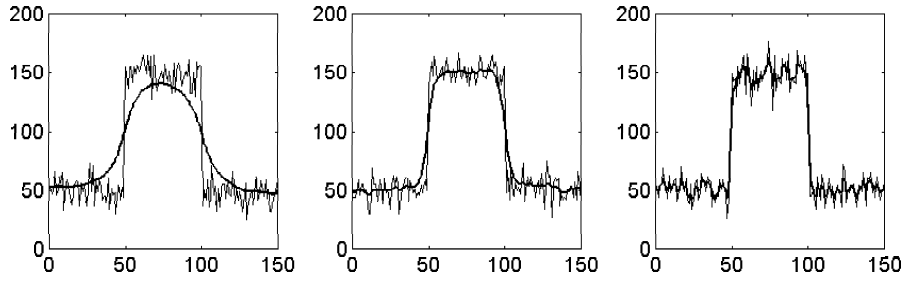

Fig. 5. Synthetic data: Noisy and filtered data for $\alpha = 1, 0.1, 0.001$.

The two solutions are not identical. However, their difference is so small that can not be observed in Fig.5. The two solutions are almost identical within the interval $[0, 150]$, except in the vicinity of the boundaries.

6.2 Ultrasound image

This example considers the problem of noise reduction in ultrasound images, using a multiplicative model for the noise (Rayleigh model). The MAP estimates of the original image was computed by both methods, i.e., by solving the linear set of equations (13), obtained by linearization of the non linear cost function (2), and by using the IIR filters defined in (21). X^{ML} and k are computed using (5).

We have used the following separable filter to process the ultrasound image

$$G(Z_1, Z_2)_{2D} = G(Z_1)G(Z_2) \tag{26}$$

where $G(Z)$ is given by (18). Separable filters allow fast filtering procedures based on two steps: in the first step the filter $G(Z)$ is applied to each column of the ultrasound image and, in a second step $G(Z)$ is applied to each row of the image obtained in the previous step. Fig.6 shows an ultrasound image (left), and the MAP estimates obtained by the IIR filter (right). The results achieved by solving (13) are nor shown since they are similar.

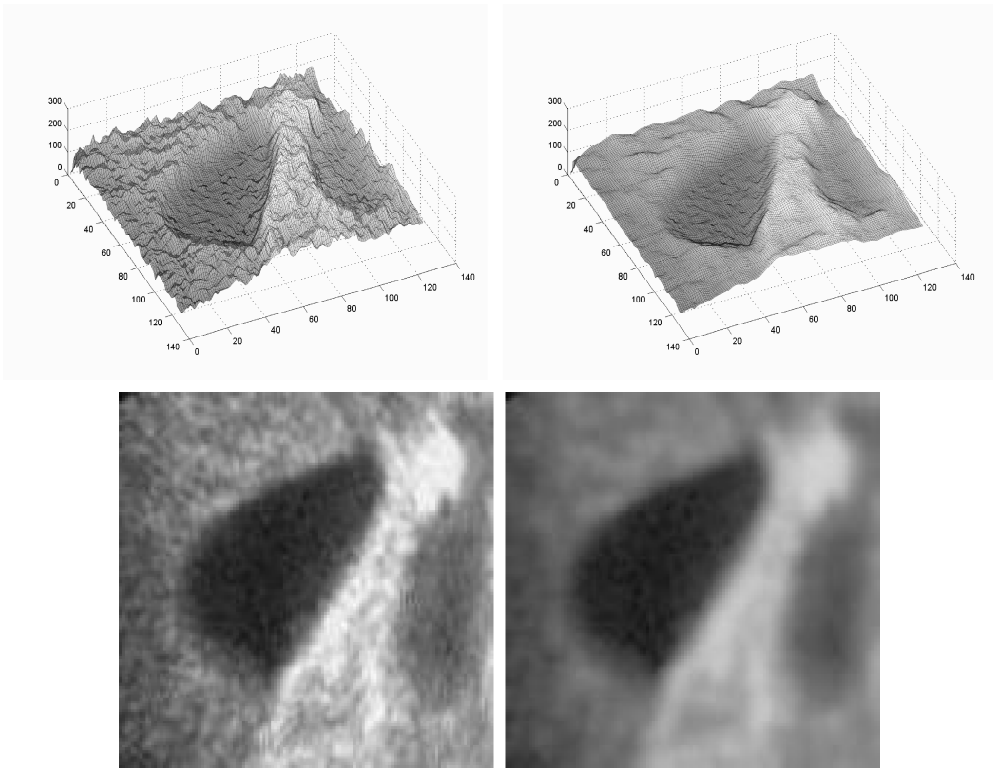

Fig. 6. First column:ultrasound image, Second column: MAP estimates .

Fig.7 shows the matrix of coefficients k_i. As noted before, the Rayleigh distribution, corresponding to a multiplicative type of noise, make the coefficients depend not only on the amount of data, but also on the data itself (see (5)). Therefore, the lighter zones of Fig.7 correspond to regions where the cutoff frequency of the IIR filter is smaller and consequently the regularization effect imposed by the prior is higher. On the contrary, in the darker regions, corresponding to the darker regions on the original ultrasound image, the regularization effect imposed by the prior is smaller.

Fig.8 shows the image obtained by computing the absolute difference of the images computed by the matrix inversion method and by the filtering method. The difference is small, except at the transitions and at the borders. The signal do noise ratio is $SNR = S_A - S_\Delta = 43.64dB$. S_A is the energy of the image X_A estimated using equation (13) and computed as $S_A = 10\log_{10}(X_A.X'_A)$. $S_\Delta = 10\log_{10}[(X_A - X_F)(X_A - X_F)']$ is the energy of the error image displayed in Fig.8. X_F is the image estimated using equation (21). The largest difference between X_A and X_F is observed at the origin due borders effect (as expected).

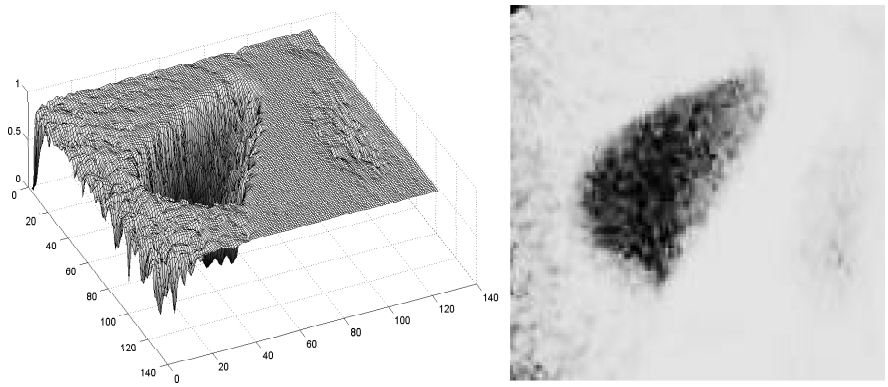

Fig. 7. Matrix of coefficients k_i.

7 Conclusions

This paper formulates the MAP estimation of signals as a filtering problem and shows that the MAP estimates can be obtained by a two step approach. The first step computes the ML estimates of the signal using standard estimation methods. In the second step the ML estimates are low pass filtered by an IIR filter. This can be performed by using two recursive filters: a forward causal filter and a backward anti-causal filter.

The filter coefficients depend on the number of observations available for each unknown variable. In the case of non Gaussian noise (Rayleigh distributed), the coefficients also depend on the sufficient statistics of the signal.

The numerical methods presented in the paper showed that the proposed algorithm based on recursive filtering is faster and produces results identical to the ones obtained by solving the MAP equations using stand numerical methods. Furthermore, this approach can be extended to higher order filters corresponding to other Gibbs priors with better performances at transitions [20]. This can also be seen as a useful tool for the design of prior distributions based on filtering theory.

References

1. V.E.Johnson, W.H. Wong, X. Hu and C. Chen, Image Restoration Using Gibbs Priors: Bondary Modeling, Treatment of Blurring, and selection of Hyperparameters, IEEE PAMI, vol.13, no.5, May 1991.
2. D.L. Snyder, M.I Miller, L.J. Thomas and D.G. Politte, Noise and Edge Artifacts in Maiximum-Likelihood Reconstructions for Emission Tomography, IEEE Trans. on Medical Imaging, vol. MI-6, no.3, September 1987.
3. A.K. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1989.

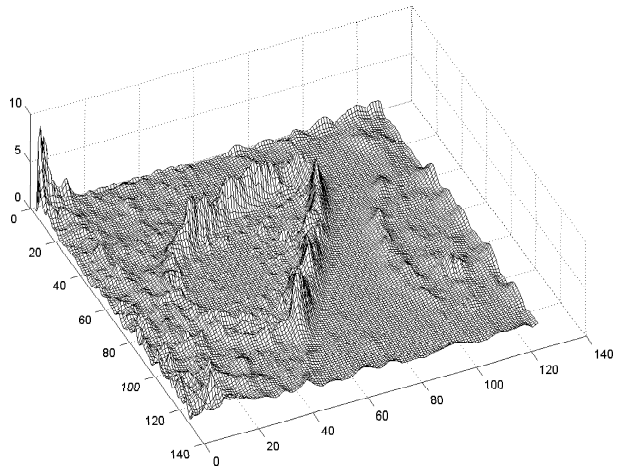

Fig. 8. Error image($SNR = 43.64dB$).

4. E.S. Chornoboy, C.J. Chen, M.I. Miller, T.R. Miller and D.L. Snyder, An Evaluation of Maxim Likelihood Reconstruction for SPECT, IEEE Trans. on Medical Imaging, vol.9, no.1, March 1990.
5. T. A. Gooley and H. H. Barret, Evaluation of Statistical Methods of Image Reconstruction Through ROC Analysis, IEEE Trans. on Medical Imaging, vol.11, no.2, June 1992.
6. L.A.Shepp and Y.Vardi, Maximum Likelihood Reconstruction for Emission Tomography, IEEE Trans. on Medical Imaging, vol.MI-1, no.2, October 1982.
7. T.Herbert and R. Leahy, A Generalized EM Algorithm for 3-D Bayesian Reconstruction from Poisson Data Using Gibbs Priors, IEEE Trans. on Medical Imaging, vol.8, no.2, June 1989.
8. A. K. Katsaggelos, Digital Image Restoration, Springer Series in Information Sciences, Springer-Verlag, 1991.
9. D.S. Lalush and B.M.W. Tsui, Simulation Evaluation of Gibbs Prior Distributiions for Use in Maximum A Posteriori SPECT Reconstructions, IEEE Trans. on Medical Imaging, vol.11, no.2, June 1992.
10. P.J. Green, Bayesian Reconstructions From Emission Tomography Data Using a Modified EM Algorithm, IEEE Trans. on Medical Imaging, vol.9, no.1, March 1990.
11. S. Geman and D. Geman, Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images, IEEE Trans on Pattern Analysis and Machine Intelligence, vol.PAMI-6, no.6, pp. 721-741, November 1984.
12. S.Z.Li, Close-Form Solution and Parameter Selection for Convex Minimization-Based Edge-Preserving Smoothing, IEEE Trans. on PAMI, vol. PAMI-20, no.9, pp.916-932, September 1998.
13. Jorge S. Marques, Pattern Recognition. Statistical and Neuronal Approches, IST Press, 1999.
14. Jae. S. Lim, Two-Dimensional Signal and Image Processing, PTR Prentice Hall, Englewood Cliffs, New Jersey.
15. R. Deriche. Using Canny's criteria to derive a recursively implemented optimal edge detector. The International Journal of Computer Vision, 1(2):167-187, May 1987.

16. J. Abbot and F. Thurstone, Acoustic Speckle: Theory and Experimental Analysis, Ultrasound Imaging vol.1, pp.303-324, 1979.
17. J. Dias, T. Silva, J. Leitão,Adaptive Restoration of Speckled SAR Images Using a Compound Random Markov Field, Procedings ICIP98 Chicago,vol.II,pp.79-83,October 1998.
18. C. Burckhardt, Speckle in Ultrasound B-Mode Scans, IEEE Trans. on Sonics and Ultrsonics, vol. SU-25, no.1, pp.1-6, January 1978.
19. J. Sanches and J. S. Marques, A Fast MAP Algorithm for 3D Ultrasound, Proceedings Third International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition, EMMCVPR2001, pp.63-74, Sophia Antipolis, France, September 2001.
20. J. Sanches and J. S. Marques, A MAP IIR Filter for 3D Ultrasound, ICIP2002, Rochester, New York, pp.II-957-960,September 2002.