BAYESIAN ESTIMATION OF ABRUPT CHANGES CONTAMINATED BY MULTIPLICATIVE NOISE USING MCMC

Jean-Yves TOURNERET, Michel DOISY and Manuel MAZZEI

National Polytechnic Institute of Toulouse
ENSEIHT/GAPSE, 2 rue Camichel, 31071 Toulouse, France
Tel (33)(0) 5 61 58 83 14 email: tournere@len7.enseeiht.fr

ABSTRACT

The paper addresses the estimation of abrupt changes which are contaminated by multiplicative Gaussian noise. The marginal mean a posteriori or marginal maximum a posteriori estimators can be derived for estimating the position of a single abrupt change. However, these estimators have optimization or integration problems for multiple abrupt changes. The paper solves these optimization problems by using Markov Chain Monte Carlo methods.

1. INTRODUCTION

Increasing interest is being shown in many signal processing applications for abrupt change estimation and detection. These applications include segmentation, fault detection or monitoring (for an overview see [1] and references therein). Most of these studies have been carried out for signals contaminated by additive noise. However, the observed process can also be corrupted by multiplicative noise. Some examples of multiplicative noise occur in image processing (speckle) or communication (fading channels). This paper focuses on edge detection in Synthetic Aperture Radar (SAR) images contaminated by multiplicative speckle noise.

Abrupt change and noise parameters can be estimated using the Maximum Likelihood (ML) method [10]. However, the resulting Maximum Likelihood Estimator (MLE) has serious limitations, especially when multiple abrupt changes occur. First, since no prior knowledge about the parameters is used, the ML estimates may lie outside the realistic parameter range. Second, the MLE is sensitive to over-parametrization [8]. Finally, the MLE can face optimization problems. This paper uses Bayesian inference for abrupt change location estimation. The marginal posterior probability density function (pdf) of abrupt change locations is used to derive the Marginal MEan A Posteriori (MMEAP) and Marginal MAximum A Posteriori (MMAAP) estimators. Unfortunately, the implementation is difficult with multiple abrupt changes. Markov Chain Monte Carlo (MCMC) methods are then used to simulate the marginal abrupt change location posterior pdf and to compute the MMEAP and MMAAP estimators.

2. PROBLEM FORMULATION

The SPECtral ANalysis (SPECAN) algorithm (combined with parametric spectral estimation) was shown to offer spatial resolution improvement and speckle noise reduction [3][6]. The SPECAN performs a line-by-line processing of the SAR image. Following the conventional SAR processor (i.e. matched filter + SPECAN), the received signal can be modelled by:

$$x_n = b_n s_n \qquad n = 1, ..., N \tag{1}$$

b_n, s_n and x_n are the multiplicative noise, the uncorrupted and corrupted line of the SAR image respectively. The properties of b_n and s_n in SAR image processing are defined by:

- The terrain reflectivity is usually modelled by a complex random variable whose real and imaginary parts are zero-mean Gaussian (considering the very large number of image cells in the radar field of view and invoking the central limit theorem [2][3]). Consequently, the *multiplicative noise* b_n is modelled by a zero mean Gaussian variable with variance σ^2.
- The uncorrupted line of the SAR image s_n can be modelled by K steps, when K fields with different reflectivities are considered. Denote N as the number of samples. Denote T as the sampling period and m_{i-1} (with $m_0 = 0$ and $m_K = N$) as the sample point after which there is the ith sudden change (with amplitude A_i) in the signal ($i = 1, ..., K$). The actual change locations are $t_i = m_i T + \tau$, with $0 < \tau < T$. The uncorrupted line of the SAR image can then be defined by:

$$s_n = A_i, \quad n \in \left]m_{i-1}, m_i\right], \quad i = 1, ..., K \tag{2}$$

3. BAYESIAN DATA ANALYSIS

The parameter a posteriori probability density function (pdf) is used in the Bayesian formalism. This a posteriori pdf is the product of the likelihood function conditioned on the parameters and the parameter priors. The Gaussian likelihood function of $x = (x_1, ..., x_N)^t$ (where t denotes transposition), conditioned on the abrupt change and noise parameters (i.e. abrupt change locations $m = (m_1, ..., m_{K-1})^t$, abrupt change amplitudes $\mu = (\mu_1, ..., \mu_K)^t$ and noise standard deviation σ) is defined by:

$$\begin{aligned} p(x|\,\theta) &= \prod_{k=0}^{K-1} \prod_{i=m_k+1}^{m_{k+1}} \frac{1}{\sqrt{2\pi\sigma^2\mu_{i+1}^2}} \exp - \frac{x_i^2}{2\mu_{k+1}^2\sigma^2} \\ &= \frac{\left(2\pi\sigma^2\right)^{-N/2}}{\prod\limits_{k=0}^{K-1} |\mu_{k+1}|^{m_{k+1}}} \exp - \sum_{k=0}^{K-1} \frac{S_{k+1}}{2\mu_{k+1}^2\sigma^2} \end{aligned} \quad (3)$$

with $\theta = (\sigma, \mu^t, m^t)^t$ and $S_{k+1} = \sum\limits_{i=m_k+1}^{m_{k+1}} x_i^2$. The parameter priors can be chosen using the a posteriori transformation invariance principle or the a priori maximum entropy principle as described in [7]. However, these priors can lead to intractable computations. This study uses the following priors for the abrupt change location estimation problem (the abrupt change number is first assumed known): 1) *Uniform priors* are chosen for the abrupt change amplitudes. These non-informative prior densities express ignorance about the value of the parameter vector μ, 2) *The Jeffrey's prior* for the standard deviation σ is a common choice in Bayesian estimation. This prior pdf corresponds to a uniform pdf in a logarithmic scale, 3) m is uniform over all ordered subsequences of $(1, ..., N)$ of length $K - 1$ (with $K << N$). Using Bayes' theorem, the posterior pdf density of the parameters, given the data and with the previous priors, can be expressed as:

$$p(\theta|\,x) \propto \frac{1}{\sigma} p(x|\,\theta) \quad (4)$$

where "$\propto$" means "proportional to". The next step in the proposed Bayesian inference removes the so called "nuisance parameters" by integration of the posterior pdf density (4) (this operation is often referred to marginalization). Marginal inference has been used successfully in signal processing, mainly because it reduces the Bayesian inference complexity. Using the gamma integral $\Gamma(t) = \int_0^{+\infty} u^{t-1} e^{-u}\, du$, the following result can be obtained:

$$p(m|\,x) \propto \prod_{k=0}^{K-1} \Gamma\left(\frac{l_{k+1} - 1}{2}\right) (S_{k+1})^{(1-l_{k+1})/2} \quad (5)$$

where $l_{k+1} = m_{k+1} - m_k$. The marginal pdf $p(m|\,x)$ is used to derive the MMEAP and MMAAP estimators, in the case of a single abrupt change. However, when multiple abrupt changes occur, these two estimators face integration or optimization problems. For instance, the implementation of the MMAAP estimator requires to examine all possible abrupt change configurations. The number of these configuration is $C_N^{K-1} = \frac{N!}{(K-1)!(N-K+1)!}$, which can be very large (for instance $C_{500}^5 \approx 10^9$). Consequently, it is not always feasible to examine all possible abrupt change configurations. Moreover, the problem is much more complicated when the abrupt change number is unknown, since the marginal density has to maximized with respect to m and K. To cope with the previous problems, the next section proposes a simulation approach using Markov Chain Monte Carlo (MCMC) methods.

4. SIMULATION USING MCMC

This section addresses the Bayesian estimation of the number and the location of multiple abrupt changes (defined in eq. (2)) using MCMC methods. Note that the abrupt change number is unknown. Define an indicator vector $\omega = (\omega_1, ..., \omega_N)^t$ such that:

$$\begin{cases} \omega_j = 1 & \text{if there is an abrupt change at lag j} \\ \omega_j = 0 & \text{otherwise} \end{cases} \quad (6)$$

The number of abrupt changes is a function of $\omega \in \Omega = \{0,1\}^N$ since $\sum\limits_{j=1}^{N} \omega_j = K - 1$. Consequently, the Gaussian likelihood function conditioned to the parameter vector $\theta = (\sigma, \mu^t, m^t, K)^t$ (defined in eq. (3)) can be expressed as a function of the parameters σ, μ and ω. Denote as $p(x|\sigma, \mu, \omega)$ the corresponding conditional pdf. The Jeffrey's and uniform priors are used for the noise standard deviation and the abrupt change amplitudes, as previously. The random variables ω_j are assumed independent with Bernoulli priors with parameter λ. The marginal posterior density $\widetilde{p}(\omega|x)$ can then be easily determined (using Bayes theorem). A Markov chain $\Theta^n = (\Theta_i^n)_{i=1...N}$ is constructed on Ω by the Metropolis-Hastings algorithm with stationary distribution $\widetilde{p}(\omega|x)$. A chain element is a vector of the form $\Theta^n(\omega)$. The Markov chain is constructed as follows: 1) the chain starts from a random starting point $\Theta^0 \in \Omega$, 2) a value z^{n+1} is drawn from $q(z^{n+1}|\Theta^n)$, where q is an arbitrary transition probability function on Ω, 3) z^{n+1} is then accepted as Θ^{n+1} with probability:

$$\alpha\left(\Theta^n, z^{n+1}\right) = \min\left\{1, \frac{\widetilde{p}(z^{n+1}|x)q\left(\Theta^n|z^{n+1}\right)}{\widetilde{p}(\Theta^n|x)q\left(z^{n+1}|\Theta^n\right)}\right\} \quad (7)$$

or equivalently, if *Rand* is the outcome of a uniform drawing on $[0,1]$:

$$\begin{cases} \Theta^{n+1} = z^{n+1} & \text{if } Rand < \frac{\widetilde{p}(z^{n+1}|x)q\left(\Theta^n|z^{n+1}\right)}{\widetilde{p}(\Theta^n|x)q\left(z^{n+1}|\Theta^n\right)} \\ \Theta^{n+1} = \Theta^n & \text{otherwise} \end{cases} \quad (8)$$

It is well-known that $\widetilde{p}(\omega|x)$ is the stationary and asymptotic distribution of Θ^n, if $q(\Theta^{'}|\Theta)$ is irreducible and aperiodic. In other words, $\Theta^n \underset{n\to\infty}{\overset{d}{\to}} \Theta^\infty$, where $\overset{d}{\to}$ denotes the convergence in distribution and $\widetilde{p}(\omega|x)$ is the distribution of Θ^∞. After a sufficiently long (so-called) burn-in, Markov Chain elements Θ^n can be used to approximate the MMAAP and MMEAP estimators. Denote as N_{MC} and N_{bi} the total number of Markov Chain runs and the number of burn-in iterations. The ergodic theorem for Markov chains (analog of the law of Large Numbers) yields:

$$\overline{\Theta}^{N_{MC}} = \frac{1}{N_{MC} - N_{bi}} \sum_{n=1}^{N_{MC}-N_{bi}} \Theta^n \underset{N_{MC}\to\infty}{\overset{d}{\to}} E[\Theta^\infty] \quad (9)$$

Eq. (9) shows that $\overline{\Theta}^{N_{MC}}$ is a good approximation of $E[\Theta^\infty] = (E[\Theta_1^\infty], ..., E[\Theta_N^\infty])^t$, for a large number of Markov Chains runs N_{MC}. Note that $E[\Theta_i^\infty]$ is the probability of an abrupt change at lag i, conditioned on the observation vector x. Consequently, the vector $\overline{\Theta}^{N_{MC}} = \left(\overline{\Theta_1}^{N_{MC}}, ..., \overline{\Theta_N}^{N_{MC}}\right)^t$ provides an estimate of the probabilities $E[\Theta_i^\infty]$. The algorithm used in the Markov Chain combines different choices of transition probabilities q in this implementation: 1) the independence sampler defined by $q(\Theta^{'}|\Theta) = q(\Theta^{'})$ is considered (the candidates are drawn independently of the current location Θ). For our experiment, q is a Bernoulli distribution (with parameter $\lambda = \frac{1}{1+e^\beta}$ and $\beta = 5$) for each abrupt change, 2) "birth" or "death" of changes (see [5]), 3) shifted abrupt change locations. To ensure convergence, the Markov chain associated with q is reversible in each of the three procedures (see [5]). Choice of different transition probabilities q increases the convergence speed since every state can be visited in few iterations.

5. SIMULATION RESULTS

Many simulations have been conducted to validate the previous theoretical results. This paper first considers an ideal step multiplied by a white Gaussian noise. This signal yields simultaneous mean and variance jump. The signal and noise parameters are $N = 1024$, $\mu = (1, 1+A)^t$, $m = N/2$ and $\sigma = 1$. The MCMC approach is not useful in this particular case. Indeed, the MMEAP and MAAP estimators can be derived using the marginal posterior pdf $p(m|x)$ defined in eq. (5). Fig. 1. shows a plot of $\ln p(m|x)$ for the previous signal and noise parameters. The abrupt change position MMAAP estimate is $\widehat{m} = 512$, for this particular realization. Fig. 2 shows the abrupt change position MMAAP estimate histogram, computed from 500 Monte Carlo runs. Both figures show the good performance of the Bayesian estimator.

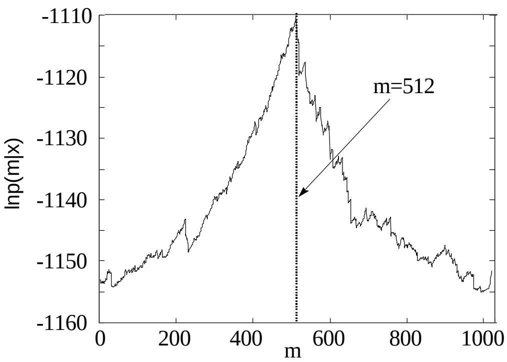

Fig. 1. Posterior pdf Logarithm

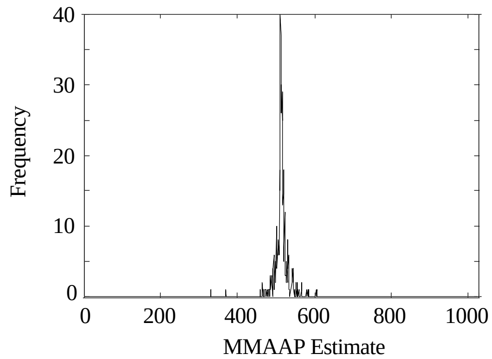

Fig. 2. Histogram of MMAAP estimates

The MMAAP estimator performance (mean square error (MSE)) is depicted in fig. 3 as a function of the number of samples for different step amplitudes. Obviously, the estimator performance increases with the step amplitude and the number of samples (except for $A = 0.1$ where the MSE is too large to be significant). This first analysis shows that the position of a step with amplitude $A \geq 0.5$ corrupted by a multiplicative white noise (with variance $\sigma^2 = 1$) can be estimated with good performance.

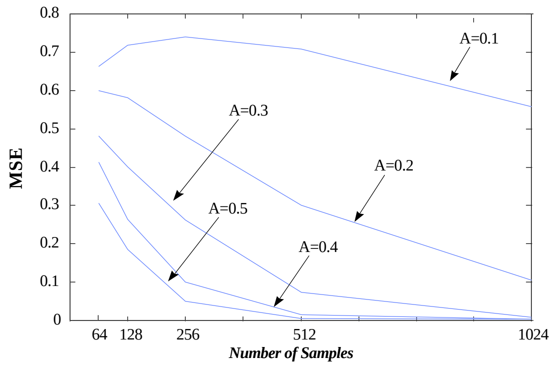

Fig. 3. Bayesian Estimate MSE

Simulations are then presented for a signal subjected to multiple abrupt changes and contaminated by multiplicative white Gaussian noise. The signal and noise parameters are $N = 500$, $m = (100, 150, 300, 450)^t$, $\mu = (1.0, 3.5, 0.5, 3.6, 0.6)^t$ and $\sigma^2 = 1$. A particular realization of this signal is plotted in fig. 4 (above). A Markov chain with invariant distribution $\widetilde{p}(\omega|x)$ is simulated. The vector $\overline{\Theta}^{N_{MC}}$, which estimates the probability of an abrupt change at each lag, is plotted in fig. 4 (below). This result was obtained for $N_{MC} = 390$ (30 burn-in iterations and 100 computation iterations for each transition probability), which only takes about 5 minutes on a PC Pentium 133Mhz in Matlab.

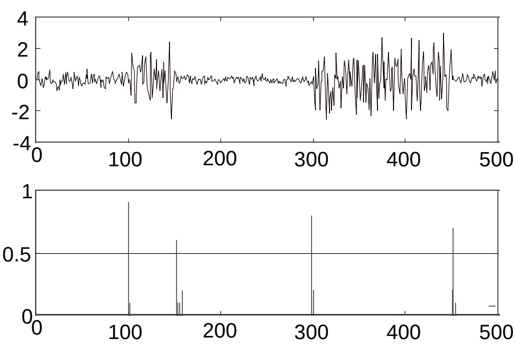

Fig. 4. Abrupt Change Position Estimate

Clearly, a suitable threshold allows the estimate of the abrupt change positions. It is interesting to note that the total simulation time (burn in + computation of $\overline{\Theta}^{N_{MC}}$) is approximately 5 minutes on a pentiumThe central limit theorem for ergodic averages can be used to obtain quantitative results from MCMC output. Indeed, under a geometrically ergodic convergence (see chapter 4 in [4]), the Markov chain $\Theta^n = (\Theta_i^n)_{i=1,\ldots,N}$ satisfies the following property:

$$\sqrt{N_{MC}}\left(\overline{\Theta}_i^{N_{MC}} - E\left[\Theta_i^{\infty}\right]\right) \underset{N_{MC}\rightarrow\infty}{\overset{d}{\rightarrow}} \mathcal{N}\left(0, \sigma_i^2\right) \quad (10)$$

The difficult problem of estimating σ_i^2 is reviewed in [4] and is currently under investigation.

6. CONCLUSION

Bayesian inference was used successfully to estimate the position of abrupt changes embedded in multiplicative white Gaussian noise. The marginal maximum a posteriori estimator showed good performance for single abrupt change estimation. Unfortunately, the implementation is difficult for multiple abrupt changes. Instead, the Metropolis Hastings algorithm was used to simulate Markov chains whose invariant distribution was the marginal a posteriori abrupt change pdf (conditionally to the observation vector). The algorithm provided an estimate of the probability of an abrupt change at each lag. Comparison of these probabilities to a convenient threshold allowed the abrupt change positions to be estimated.

Acknowledgments

The authors wish to thank M. Lavielle for his helpful discussion.

7. REFERENCES

[1] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall, 1993.

[2] A. C. Bovik, "On detecting Edges in Speckle Imagery," *IEEE Trans. ASSP*, vol. 36, pp. 1618-1627, 0ct. 1988.

[3] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar*, Wiley Interscience, 1991.

[4] C. Geyer, "Practical Markov Chain Monte Carlo," Statist. Sci., 7, pp. 473-482, 1992.

[5] P. J. Green, "Reversible Jump MCMC Computation and Bayesian Model Determination," *Biometrika*, Vol. 82, pp. 711-732, 1995.

[6] G. Mesnager, C. Lambert-Nebout and F. Castanié, "Spatial Resolution Improvement using a SPECAN Approach in SAR Data Processing," *Proc. of IGARSS'95*, July 1995.

[7] A. Quinn, "Novel Parameter Priors for Bayesian Signal Identification," *Proc. of ICASSP'97*, Munich, April 1997.

[8] J. O Ruanaidh and W. J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*, Statistics and Computing, Springer Verlag, 1996.

[9] M. Sack and al, "Application of Efficient Linear FM Matching Filtering Algorithms to SAR Processing," *Proc. IEE*, vol. 132, Feb 1985.

[10] J. Y. Tourneret and M. Chabert, "Off-line Detection and Estimation of Abrupt Changes corrupted by Multiplicative Colored Gaussian Noise," *Proc. of ICASSP'97*, Munich, April 1997.