Digital Image Restoration Using Autoregressive Time Series Type Models*

HÉCTOR ALLENDE O.[1]
JORGE GALBIATI R.[2]
RONNY VALLEJOS A.[3]

[1]Universidad Técnica Federico Santa María,
Casilla 110-V, Valparaíso, Chile
[2]Universidad Católica de Valparaíso,
Casilla 4059, Valparaíso, Chile
[3]Universidad de Valparaíso,
Casilla 5030, Valparaíso, Chile

Abstract. We consider an non-symmetric half plane autoregressive image, where the image intensity of a point is a linear combination of the intensitites of the eight nearest points located on one quadrant of the coordinate plane, plus a normal white noise innovations process. Two types of contaminations are considered. Innovation outliers, where a fraction of innovations are corrupted with a heavy tailed outlier generation process, and additive outliers, where a fraction of observations are corrupted. We develop a GM-estimator for the robust estimation of parameters of a contamined autoregressive image model, based on time series GM-estimators introduced by Denby & Martin (1979) applied to the restoration of radar generated images. Ordinary least-squares estimators are asymptotically efficient with a non-contamined gaussian process, like the one considered here. M-estimators behave better when innovation outliers are present, but are very sensitive to additive outliers. A simulation study is carried out, which shows that the GM-estimator introduced here has a better performance with an additive outlier contamined image model than M-estimators and ordinary least squares estimators.

Keywords: GM-estimators, Image Restoration Robust Estimation, Two-Dimensional Autoregressive Models.

1 Introduction

Restoration of an image in the presence of noise is one of the fundamental problems in image processing. Parametric representations of two-dimensional processes suitable for this problem, have been well studied. However, in these models, the image intensity array is assumed to be a two-dimensional Gaussian Process. There are many image restoration methods based on the Gaussian assumption. For instance, Chellepa and Kashyap (1982) used spatial interaction models to represent image intensity arrays and restored images obtained with minimum mean square

* This work was supported by FONDECYT grant Nº 1960521

error criterion. However, when the image is contamined with outliers, the estimated parameters obtained from the Gaussian model do not appear to be appropiate. A more realistic assumption for the image model is a contaminated Gaussian noise.

The importance of the ε-contaminated models has been legitimated by numerous publications about applied works in the area of image processing and image analysis. See for instance Kashyap and Eom (1988).

We develop a restoration method based on a robust image model in this work. In the proposed method, the image intensity array is represented by a causal autoregressive model. A robust parameter estimation algorithm and a data cleaning procedure is applied to restore contaminated images. The restoration algorithm based on the robust modelling is tested with several simulated images.

Our contributions are threefold. We first develop an algorithm for the robust estimation of parameters of an image model in which the innovations process is a mixture of a Gaussian and an outlier process. It is a GM-estimator. We prove the convergence and confirm the convergence via simulation. Next we consider the robust estimation of the parameters of a model where the image obeying the model is not available, the corrupting innovations process being a mixture of a Gaussian process and an outlier process.

We develop an algorithm to recover the parameters of the model from a noisy image. The procedure involves alternate parameter estimation and data cleaning.

We provide intuitive reasons for the convergence of the procedure and confirm our intution by several simulations. Finally, we use the above results to restore an image corrupted by different types of outliers.

2 The additive outlier in nonsymmetric half-plane autoregressive models

Consider a nonsymmetric half-plane autoregressive two dimensional model with additive outliers. Assume that the image intensity of an image follows the nonsymmetric half-plane model. Let (i, j) be an index for the coordinate location, and $y(i, j)$ be the intensity at the coordinate (i, j).

Let us define a nonsymmetric half-plane (NSHP) model as follows:

$$\Omega_{-} := \{(i, j) : (i = 0 \text{ and } j < 0) \text{ or } (i < 0 \text{ and } j \text{ is arbitrary}\}. \tag{2.1}$$

Let u and v be indexes for two-dimensional coordinate locations. One important property of Ω_{-} is that if $u \in \Omega_{-}$ and $v \in \Omega_{-}$ then $(u+v) \in \Omega_{-}$.

And NSHP autoregressive model can be written as

$$y(u) = \sum_{v \in N_1} \underline{a}_v y(u+v) + m + a(u) \tag{2.2}$$

where $a(u)$ are independent identically distributed random variables with a symmetric distribution G with mean zero and scale s_a. The density of G will be denoted by g. The $a's$ are called innovations. The neighborhood set N_1 is a subset of the nonsymmetric half-plane Ω_-. The NSHP autoregressive model (2.2) can be written in the linear model form

$$y(u) = \underline{a}^T z(u) + a(u) \tag{2.3}$$

where $\underline{a}^T$ is a parameter vector and $z(u)$ is a vector which consists of intensities of pixels in the neighborhood set N_1 and unity. The last element of the vector $z(u)$ is required to represent a constant gray level in the image.

If

$$N_1 = \{(0,-1),(-1,0),(-1,-1),(0,-2),(-1,-2),(-2,0),(-2,-1),(-2,-2)\}$$

the NSHP autoregressive model (2.2) can be rewritten as follows:

$$Y(i,j) = \underline{a}^T Z(i,j) + a(i,j) \tag{2.4}$$

where

$$\begin{aligned}[Z(i,j)]^T = \{&Y(i,j-1),Y(i-1,j),Y(i-1,j-1),Y(i,j-2),\\ &Y(i-1,j-2),Y(i-2,j),Y(i-2,j-1),Y(i-2,j-2),1\}\end{aligned}$$

The model given in (2.4) is called an eight neighbor causal autoregressive model, and this model is used in our simulation study.

Suppose now that the NSHP autoregressive process cannot be perfectly observed because a small fraction e (in practice we usually have $e \le 0.1$) of observations are distributed by an outlier-generating process $\{n(i,j)V(i,j)\}$, where $\{n(i,j)\}$ is one or zero, with $\mathrm{P}(n(i,j)=1)=e$, $\mathrm{P}(n(i,j)=0)=1-e$, and the variables $V(i,j)$ have arbitrary distribution function H. Thus the observational model is

$$X(i,j) = Y(i,j) + n(i,j)V(i,j) \qquad \begin{matrix} i=1,\dots,n \\ j=1,\dots,m \end{matrix} \tag{2.5}$$

Therefore with probability $(1-e)$ the NSHP autoregressive process $Y(i,j)$ itself is observed, and with probability e the observations $X(i,j)$ are corrupted by an error with distribution H.

It is well known that the LS estimates are asymptotically normal and asymptotically efficient when G is Gaussian and $V(i,j)=0$. However, when the innovations density is non-Gaussian

(Innovative Outliers), the above estimates are no longer efficient and heavy-tailed innovation distributions can result in large losses of efficiency.

The latter fact suggests that a good alternative to the LS estimate can be the M-estimate as proposed by Huber (1981) for the NSHP autoregressive case (Kashyap and Eom, 1988). However, the LS estimate and even the M-estimate are extremaly sensitive to the presence of additive outliers (AOs). This fact is reported by Bustos and Yohai (1986) for one dimensional autoregressive processes. In this work we present the results of a Monte Carlos simulation which shows that for a two dimensional eigth neighbor causal autoregressive model the LS and M-estimates are more sensitive to AO-s than in the case of causal autoregressive with innovative outliers.

3 Generalized M-estimates

Consider the parameter estimate in the NSHP autoregressive process. In the least squares estimation, we need to minimize the function

$$\sum_{i,j}[X(i,j)-\underline{a}^T Z(i,j)]^2 \tag{3.1}$$

with respect to $\underline{a}$, where $Z(u)$ is a vector which consists of the observations $X(i,j)$ in the neighbor set N_1.

The idea of a least squares estimation is to minimize the residuals. However, if one observation is an outlier, then the corresponding residual is very large, and the least squares estimator is not robust.

Similarly, the class of M-estimators proposed by Kashyap and Eom (1988) for causal autoregressive processes, defined by minimizing the function of a finite sample of observations

$$Q(\underline{a},s) = \frac{1}{mn}\sum_{i,j}\left[r\left(\frac{X(i,j)-\underline{a}^T Z(i,j)}{s}\right)+\frac{1}{2}\right]s \tag{3.2}$$

is robust for innovative outliers, when the function r has bounded influence. But the situation is totally different when the contamination model is given by (2.5), that is, when the autoregressive model is disturbed by additive outliers. This suggests introducing a more general class of robust estimators, known as GM-estimators, which are an extension of the M-estimators, obtained by assigning a weight function to the observations of the model. The residuals $X(i,j)-\underline{a}^T Z(i,j)$ in a NSHP autoregressive model contamined with additive outliers may be very large. The way the GM-estimators reduces this effect is by introducing smaller weights to larger residuals. A GM-estimator for the parameters $\underline{a}^T$ and s of model (2.5) is the solution to the problem of minimizing the non-quadratic function defined by

$$Q(\underline{a}, s) = \frac{1}{mn} \sum_{i,j} l_{ij} t_{ij} \left[r \left(\frac{X(i,j) - \underline{a}^T Z(i,j)}{l_{ij} s} \right) + b \right] s \tag{3.3}$$

where r is a differentiable function, convex, symmetric with respect to the origin, with bounded derivative, and such that $r(0) = 0$. The l_{ij} and t_{ij} are the weights corresponding to the respective $Z(i, j)$. In order to obtain consistency of the scale estimate at the normal model, we consider $b = E_f[c]$ (see Huber 1981).

The GM-estimator is obtained by solving the equations

$$\sum_{i,j} t_{ij} y \left[\frac{X(i,j) - \underline{a}^T Z(i,j)}{l_{ij} s} \right] Z^T(i,j) = \underline{0} \tag{3.4}$$

$$\sum_{i,j} l_{ij} t_{ij} \left[c \left(\frac{X(i,j) - \underline{a}^T Z(i,j)}{l_{ij} s} \right) - b \right] = 0 \tag{3.5}$$

where $y(x) = \dfrac{\partial r(x)}{\partial x}$, $c(x) = x y(x) - r(x)$ and y is a bounded and continuous function. There are several proposals for the choice of y due to the fact that the robustness of the procedure and the rate of convergence of the procedure depends on these functions: the Huber hard-limiter type, given by $y_H(x) = \operatorname{sgn}(x).\min\{|x|, c\}$ and Tuckey's redescending bisquare function given by

$$y_B(x) = \begin{cases} x[1 - x/a]^2 & ,|x| \le a \\ 0 & ,|x| > a \end{cases}$$

Typical values for the adjusting constant c in y_H range from 1.5 and 2.0 and for a in y_B range from 4.5 and 6.

The principal types of GM-estimators are:

i) Mallows type, where $l_{ij} = 1$ and $t_{ij} = \dfrac{y(b_{ij}/c_r)}{b_{ij}/c_r}$ with $b_{ij} = p^{-1} z^T_{(i,j)} \hat{C}^{-1} Z_{(i,j)}$, where $\hat{C}^{-1}$ is a robust estimate of C^{-1} and C is the a priori unknown covariance matrix for the NSHP autoregressive process, which may be expressed as $C(\underline{a})$. The construction of $\hat{C}^{-1}$ is described by Martin (1980).

ii) Schweppe type $l_{ij} = t_{ij} = \dfrac{y(b_{ij}/c_r)}{b_{ij}/c_r}$.

4 Implementation of GM-estimates

Assuming that an estimate of C^{-1} required to construct the weights t_{ij}, is available, then good approximate solutions of equations (3.4) and (3.5) can be conveniently obtained by using an iterately-weighted-least-squares (IWLS) techniques similar to that described by Martin (1980).

It may be shown that the estimating equations (3.4) and (3.5) have a unique solution when y is strictly monotone.

The GM-estimation of the NSHP autoregressive model under regularity conditions preserve the properties of consistency and asymptotic normality of the unidimensional autoregressive models. But they also have their computation difficulties, because they involve the minimization of a non quadratic function of multiple parameters. To obtain the GM-estimator of $\underline{a}$ and s we use the algorithm known as IWLS, whose convergence is established in Huber (1981).

IWLS algorithm

Let $X(i,j)$ be the observations of the contamined causal autoregressive model defined in (2.5) and let $\underline{a}^{(0)}$ and $s^{(0)}$ be the initial values, e a tolerance value and weights l_{ij} , t_{ij}, $i=\overline{1,n}$, $j=\overline{1,m}$, starting values.

1. Set $k=0$.
2. At the k-th iteration of $\underline{a}^{(k)}$ obtain the residual

$$r^{(k)}(i,j)=X(i,j)-\underline{a}^{(k)}Z(i,j), i=1,\cdots,n, j=1,\cdots,m$$

3. Compute the new value of s using $\hat{s}=1.483Med\{|r^{(0)}_{(i,j)}-Med|r^{(0)}_{(i,j)}||\}$
4. Compute the weights W_{ij}, from $r(i,j)$, l_{ij}, and t_{ij} for the Mallows or the Schweppe type GM-estimators.

$$W_{ij}^{(k)}=\begin{cases} t_{ij}\,y\left(\dfrac{r^{(k)}_{(i,j)}}{l_{ij}\hat{s}}\right)\Big/\left(\dfrac{r^{(k)}_{(i,j)}}{\hat{s}}\right) & if \quad r^{(k)}_{(i,j)}\neq 0,\ \ l_{ij}\neq 0 \\ t_{ij}/l_{ij} & if \quad r^{(k)}_{(i,j)}=0,\ \ l_{ij}\neq 0 \\ 1 & if \quad r^{(k)}_{(i,j)}=l_{ij}=t_{ij}=0 \\ 1 & if \quad r^{(k)}_{(i,j)}\neq 0,\ \ l_{ij}=t_{ij}=0,\ \ y_1(t)=t \\ 0 & if \quad r^{(k)}_{(i,j)}\neq 0,\ \ l_{ij}=t_{ij}=0,\ \ y_1 \ \ is\ \ bounded \end{cases}$$

where $i=\overline{1,n}$, $j=\overline{1,m}$. Define $W^{(k)}$ as a diagonal matrix with $W_{ij}^{(k)}$ as its $[(n-1)(j+1)+i-1]$-th diagonal element.

5. Solve $\sum_{ij}[r^{(k)}(i,j) - Z^T(i,j)\underline{t}^{(k)}]^2 W_{ij}^{(k)} = \min$

for $\underline{t}^{(k)}$, the solution given by

$$\hat{\underline{t}}^{(k)} = [Z^T W^{(k)} Z]^{-1} Z^T W^{(k)} \cdot X - \underline{a}^{(k)}$$

where the rows of Z^T are the $Z(i,j)$ defined in (2.4), and X is the vector of observations.

6. Compute the new value of $\underline{a}$, $\underline{a}^{(k+1)} = \underline{a}^{(k)} + l\,\hat{\underline{t}}^{(k)}$ with $0 < l < 2$, an arbitrary relaxation constant.
7. Repeat 2 to 6 until the stopping rule: $\| \underline{a}^{(k)} - \underline{a}^{(k+1)} \| = | l\,\underline{t}^{(k)} | < e\hat{s}$ is reached.

5 Applications to Image Restoration

Restoration of an Image in the presence of noise is one of the fundamental problems in image processing. The image degradation process can be modeled by the observational model (2.5). We assume that the observation $X(i,j)$ is corrupted by a contamined process which contains a small fraction of additive outliers.

There are many image restoration methods based on the Gaussian noise assumption. Chellapa and Kashyap (1982) used a spatial interaction model to represent an image intensity array and restored images with minimum mean square criterion. Geman and Geman (1984) used the equivalence of Markov random field and Gibbs distribution and restored images by a stochastic relaxation method with maximum a posteriori criterion. Wu (1985) used a multimensional Kalman filtering approach and nonsymmetric half plane autoregressive model.

Unfortunately, most image restoration methods based on the Gaussian assumption are not effective to impulse noise.

Image restoration is an estimation of original intensity $Y(i,j)$ from the observation $X(i,j)$. For a small size image, the original image intensity can be modeled by a causal autoregressive model. If the original image intensity follows a causal autoregressive model, then the original image intensity can be easily restored by data cleaning with robust parameter estimation. The data cleaning procedures removes outliers at each iteration without degrading the original signal.

The restoration method based on the robust image model has an advantage over conventional methods such as median filter or $\underline{a}$-trimmed mean filter. The robust image model based method does not produce blurred images after restoration. Conventional methods, such as median filter or $\underline{a}$-trimmed mean filter, replace every pixel by its location estimates. Because these methods are based on the constant intensity assumption, the details of the original image are significantly blurred.

This procedure is described in the following algorithm.

Restoration Algorithm Based on a Robust Model

1. Initially, set $Y^{(0)}(i,j) = X(i,j)$. Compute the initial estimate $\underline{a}^{(0)}$, $s^{(0)}$ from the contaminated observation $X(i,j)$ by the least squares algorithm.
2. Consider the k-th iteration, where $Y^{(k)}$ and $\underline{a}^{(k)}$ are available. Obtain the updated estimate $Y^{(k+1)}(\cdot)$ from $Y^{(k)}(\cdot)$ by the following recursive equation

$$r^{(k)}(i,j) = Y^{(k)}(i,j) - \underline{a}^{(k)^T} Z^{(k)}(i,j)$$

$$\hat{r}^{(k)}(i,j) = y\left[\frac{r^{(k)}(i,j)}{s^{(k)}}\right]\hat{s}$$

where y is one of the bounded and continuous functions as discussed in the GM-estimation.

3. Restore the image $Y^{(k+1)}(i,j)$ using

$$Y^{(k+1)}(i,j) = \underline{a}^{(k)} Z^{(k)}(i,j) + \hat{r}^{(k)}(i,j)$$

4. Obtain estimators $\underline{a}^{(k+1)}$ from the cleaned data $Y^{(k+1)}$ by minimizing the function

$$Q(\underline{a}, s) = \frac{1}{mn}\sum_{i,j} l_{ij} t_{ij}\left[r \left[\frac{Y^{(k+1)}(i,j) - \underline{a}^T Z^{(k+1)}(i,j)}{l_{ij}\, s}\right] + b\right] s$$

This can be computed by the IWLS Algorithm.

5. Repeat Steps 2-4 until the difference of estimates between iteration becomes small.

Simulation Study

A simulation study was conducted to observe the behaviour of the GM-estimator and compare it with the LS and M estimators.

In each case one hundred 50×50 images were generated using (2.4), with additive contamination generated by (2.5), with H a large variance s_H^2 normal distribution. The parameter values were

$$\underline{a}^T = (-0.12; 0.37; -0.16; 0.25; 0.13; 0.24; 0.40; -0.16)$$

and $s_a = 0.01$.

The following cases were simulated:

No contamination

5 % contamination , $s_H^2 = 0.1$ and 0.5

10% contamination, $s_H^2 = 0.1$ and 0.5

15%contamination, $s_H^2 = 0.1$ and 0.5

Each was run three times, estimated using LS, M and GM, respectively. The mean square error of the estimated a parameters is used as a measure of performance of the estimators. The results are shown in **Table 1**.

Table 1 - Comparison of GM-estimador, M-estimator and least squares estimator for NSHP autoregressive model with additive contaminarion. Number of runs in fach case is 100. Image size is 100×100.

Percentage of outliers	Outlier Standard Deviation	Estimator		
		LS	M	GM
		Mean Square Error		
0	0	0.0177	0.0184	0.0182
5	0.1	0.1019	0.0474	0.0473
10	0.1	0.1389	0.0760	0.0758
15	0.1	0.1572	0.0985	0.0967
5	0.5	0.2155	0.1051	0.1084
10	0.5	0.2274	0.1357	0.1375
15	0.5	0.2336	0.1555	0.1546

After observing the results of the simulation study, we conclude that robust estimators have a better performance than the LS-estimator. The difference between the M and GM estimators is less significant than expected, although the GM is slightly better in some cases.

References

[1] Allende, H. and Heiler, S. “Recursive Generalized M-estimates for Autorregresive Moving Average Models”. Journal of Time Series Analysis, Vol. 13, N° 1. 1992.

[2] Boente, G.; Fraiman, R. and Yohai, V. J. “Qualitative Robustness for General Stochastics Process”, Technical Report N° 26. Department of Statistics. University of Washington. Seattle. WA. (1982).

[3] Bustos, H.O. “General R-Estimates for contaminated ph-order autorregresive processes: consistency and asymptotic normality”, Z. Wahrscheinlich. Verw. Geb. 49, 491-504. 1982.

[4] Bustos, H. O. and Yohai, V. J. “Robust estimates for ARMA models”. J. Anv. Statist. Assoc. 81(393), 155-68. 1986.

[5] Denby, J. and Martin, R.D. “Robust estimation of the firstorder autorregressive Parameter”. J. Am. Statist. Assoc. 74(365), 1979.

[6] Dutter, R. “Robuste Regression”. Bericht N^{o}135, Math. Statist. Eidgenoessische Techn. Hochschule. Zürich 1980.

[7] Dutter, R. and Huber, P.J. “Numerical Methods for the Nonlinear Robust Regression Problem”, J. Statist. Comput. Simul. Vol. 13,2, 79-114. 1981.

[8] Hampel, F.R. “A General Qualitative Definition of Robustness”, Ann. Math. Statist. 43, 1971.

[9] Hampel, F.R. “Robust Estimation: a Condensed Partial Survey”. Z. Wahrscheinlich. Verw Geb. 27. 1973.

[10] Hampel, F.R. “*Robust Statistics*”, New York: Wiley. 1986

[11] Huber, P. J. “*Robust Statistics*”. New York: Wiley, 1981.

[12] Kashyap, R. L. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol PAMI-4. 1982.

[13] Kashyap, R. L. and Eom, K. B. “Robust Image Techniques with an Image Restoration Application” IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. 36, N^o 8, pp. 1313-1325, Aug. 1988.

[14] Kleiner, B.; Martin, R.D. and Thomson, D.J. “Robust Estimation of Power Spectra” J. Roy. Soc., Series B, Vol. 41, N^o 3, pp. 313-351, 1979.

[15] Technical Report C.L. Mallows “On some topics in robustness”. Technical Memorandum, Bell Laboratories, Murray Hill. NJ. 1975.

[16] Martin, R.D. and Zeh, J.E. “Generalized M-Estimates for Autorregresions, Incluiding Small-Sample Efficency Robustness”. Techn. Rep. N^o 214, Department of Electrical Engineering, University of Washington, Seattle. 1978.

[17] Martin, R.D. “*Robust estimation for Autorregressive Models*”, In: Brillinger D.R. and G.,C. Tiao (Eds.): Directions in Time Series. Inst. Math. Statist. Publications, Haywood, C.A., 1980.

[18] Martin, R.D. and Yohai, V. J. “*Robustness in time series and estimatin ARMA models*”. Handbook of Statistics Vol. 5 (eds. E. J. Hanann, P.R. Krishnaiah and M.M. Rao), 1985.

[19] Rousseuw, P. J. “Least median of Squares regression”. J.Am. Assoc. 79, 1984.

[20] Zeh, J.E. “*Efficiency Robustness of Generalized M-Estimates for Autorregression and their use in determining outliers Type*”. Ph.D. Thesis. University of Washington. Seattle. 1979.

[21] Chellepa, R.and Kashyap, R.L. “Digital Image Restoration Using Spatial Interaction Models”. IEEE Trans. on Acoustics, Speech and Signal Processing, Vol 30 N^o 3, pp 461-472, June 1982.

[22] Geman, S. and Geman, D. “Stochastic relaxation, Gibbs distributions, and Bayesian restoration of images”. IEEE Trans. Pattern Anal. Machine Intell. Vol 6 pp 721-741 Nov. 1984.

[23] Wu, Z. “Multidimensional state-space model Kalman filtering with application to image restoration”. IEEE Trans. Acoust. Speech. Signal Processing Vol 33 pp 1253-1263 Oct. 1985.