Proceedings of the
3rd International Conference on Inverse Problems in Engineering:
Theory and Practice, June 13-18, 1999, Port Ludlow, Washington, USA

REGULARIZATION TOOLS AND MODELS FOR IMAGE AND SIGNAL RECONSTRUCTION

Jérôme Idier
Laboratoire des Signaux et Systèmes
Supélec, 91192 Gif-sur-Yvette cedex, France
idier@lss.supelec.fr

*

ABSTRACT The present paper proposes a synthetic overview of regularization techniques for the reconstruction of piecewise regular signals and images. The stress is put on Tikhonov penalized approach and on subsequent non-quadratic and half-quadratic generalizations. On one hand, a link is made between the detection-estimation formulation and the non-convex penalization approach. On the other hand, it is highlighted that convex penalizing functions provide a good edge-preserving compromise between quadratic regularization and the numerically burdensome detection-estimation approach.

INTRODUCTION: TIKHONOV REGULARIZATION

Generalities

In general, solutions of ill-posed problems are brought by restricting the class of admissible solutions, given suitable prior knowledge. One of the most basic regularization method has been introduced by Tikhonov and Arsenin (1977). Assume that an unknown univariate signal x^* is to be estimated from the data $\boldsymbol{y} = Ax^* + noise$, where A is a linear operator. Tikhonov approach requires the choice of a stabilizing functional $\|Px\|^2$, where P is also a linear operator. Then, an estimated solution is $\hat{x}$ that minimizes the objective functional

$$\mathcal{J}(x) = \|\boldsymbol{y} - Ax\|^2 + \lambda \|Px\|^2 , \qquad (1)$$

where λ is a regularization parameter, which controls the trade-off between the regularity of a solution and its fidelity to the data. It can be shown that under mild conditions the criterion $\mathcal{J}$ is strictly convex and a unique solution exists.

The first term of $\mathcal{J}(x)$ in (1) is a quadratic norm that penalizes the discrepancy between the "output" for admissible functions x and given data. It is an optimal choice in a certain statistical sense when the noise is assumed to be a realization of a white centered Gaussian random process. Other common choices derive from different statistical assumptions. For instance, a Poisson distribution may be preferable for *corpuscular* imaging, as encountered in CCD image processing, X-ray tomography, positron emission tomography. When statistical information lacks, quadratic penalties are still useful to define solutions *in the least squares sense*. The present paper rather focuses on the second term of $\mathcal{J}(x)$, *i.e.,* on the construction of stabilizing functionals.

In Tikhonov original contribution, the stabilizing functional is given by

$$\|Px\|^2 = \sum_{r=0}^{R} \int c_r(t) \left|x^{(r)}(t)\right|^2 dt,$$

where the weights c_r are strictly positive functions and $x^{(r)}$ indicates the rth derivative of x. It is clear that such a choice corresponds to the prior knowledge that the estimated signal x is smooth to a certain extent.

Multivariate extensions have been proposed. More specifically, the two-dimensional case has been developed for the purpose of image restoration. For instance, Hunt's pioneering work (Hunt, 1977) is devoted to image deblurring based on the 2D discrete counterpart of (1) and on fast computation in the Fourier domain.

Limitation

A great deal of inverse problems involve piecewise homogeneous objects, such as blocky signals, images with well-separated regions, isolated defects within homogeneous media. In such context, methods based on Tikhonov approach often fail to detect or even to preserve the expected discontinuities.

As an illustration, consider the following simulated 1D experiment: let $\boldsymbol{y} = [y_0, \ldots, y_N]^{\mathrm{t}}$ represent a noisy data vector sampled from a *piecewise smooth* univariate function x^* on $[0, 1]$: $y_n = x^*(n/N) + b_n$. x^* and $\boldsymbol{y}$ are depicted on Figure 1. With $R = 2, c_0 = c_1 = 0, c_2 = 1$, and $\|\boldsymbol{y} - Ax\|^2 = \sum_{n=0}^{N} |y_n - x(n/N)|^2$,

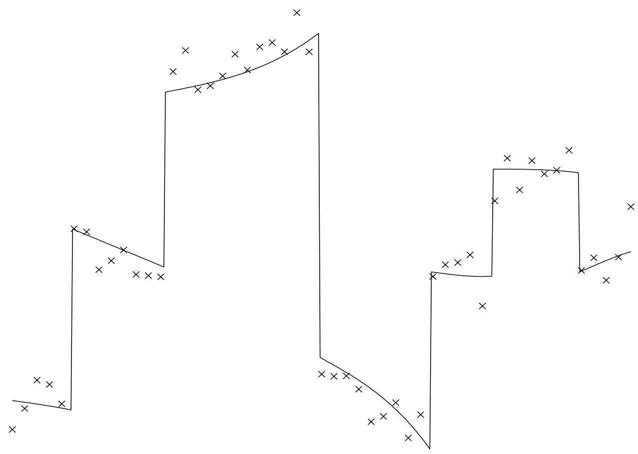

Figure 1: A piecewise smooth univariate function x^*, and a noisy data series $(y_n), n = 0, \ldots, N = 50$, sampled from x^*.

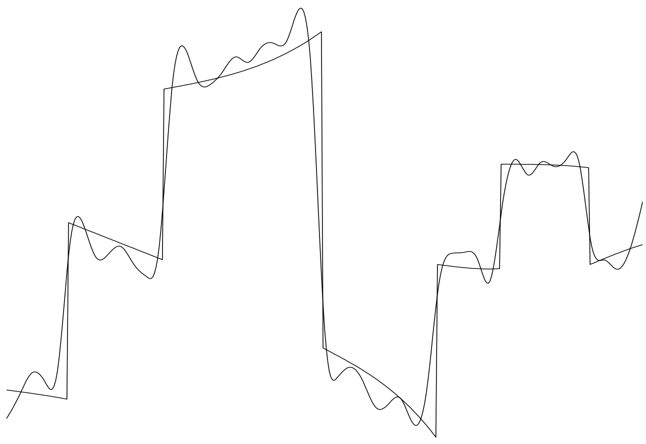

Figure 2: Smooth linear estimate $\widehat{\boldsymbol{x}}_\lambda$ obtained as the minimizer of (3), $M = 400$ for the optimal value of λ in the L_1 sense: $E(\widehat{\boldsymbol{x}}_\lambda, x^*) = 18.16\%$.

(1) reads

$$\mathcal{J}(x) = \|\boldsymbol{y} - Ax\|^2 + \lambda \int_0^1 \left|x''(t)\right|^2 \, dt. \tag{2}$$

Let us introduce a discrete approximation for (2), based on finite differences:

$$J_M(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \lambda M^3 \sum_{m=1}^{M-1} |2x_m - x_{m-1} - x_{m+1}|^2 , \tag{3}$$

where $\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 = \sum_{n=0}^N \left|y_n - x_{nM/N}\right|^2$, $\boldsymbol{x} = [x_0, \ldots, x_M]^{\mathrm{t}}$, and $M+1$ is a multiple of $N+1$. Define the estimated vector $\widehat{\boldsymbol{x}}_\lambda$ as the minimizer of J_M. It is a classical result that pointwise convergence of $\widehat{\boldsymbol{x}}_\lambda$ towards the unique minimizer of (2) holds when $M \to \infty$. Figure 2 depicts $\widehat{\boldsymbol{x}}_\lambda$ for $M = 400$ and for the "best" value of λ in the L_1 sense, *i.e.,* the one that minimizes $E(\widehat{\boldsymbol{x}}_\lambda, x^*)$, with

$$E(\boldsymbol{x}, x^*) = \sum_{m=0}^{M} |x_m - x^*(m/M)| . \tag{4}$$

This procedure is artificial, since it requires the knowledge of x^*, but it is adopted here to allow fair comparisons.

The solution $\widehat{\boldsymbol{x}}_\lambda$ is not satisfying, because it is uniformly smooth, whereas the expected solution is only piecewise smooth. In comparison, the simple estimate obtained from $\boldsymbol{y}$ by linear interpolation reaches a lower L_1-norm error value of 17.04%.

If the number P and the positions $\boldsymbol{t} = [t_1, \ldots, t_P]$ of the discontinuities were known, then a convenient approach would be to replace $\mathcal{J}$ by

$$\mathcal{J}_{\boldsymbol{t}}(x) = \|\boldsymbol{y} - Ax\|^2 + \lambda \sum_{p=0}^{P} \int_{t_p}^{t_{p+1}} \left|x(t)''\right|^2 \, dt,$$

with $t_0 = 0$ and $t_{P+1} = 1$. The discrete counterpart reads

$$J_{\boldsymbol{b}}(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \lambda M^3 \sum_{m=1}^{M-1} b_m \, |2x_m - x_{m-1} - x_{m+1}|^2 , \tag{5}$$

where $\boldsymbol{b} = [b_1, \ldots, b_{M-1}]$ is a binary vector of *edge variables*: $b_m = 0$ corresponds to the presence of a discontinuity at the mth position. In practice, this approach is very limited, because the discontinuities are usually not known. In the following sections, we introduce the main ideas brought in the signal and image processing community to tackle the problem of restoring piecewise regular functions.

THE DETECTION-ESTIMATION APPROACH

Principle

A fruitful idea was introduced to cope with discontinuities in the field of computer vision (Mumford and Shah, 1985; Blake and Zisserman, 1987). It consists in *jointly* considering the estimation problem of x and the detection problem of $\boldsymbol{t}$ or $\boldsymbol{b}$. For this purpose, jointly minimizing $\mathcal{J}_{\boldsymbol{t}}(x)$ in $(x, \boldsymbol{t})$ or $J_M(\boldsymbol{x}, \boldsymbol{b})$ in $(\boldsymbol{x}, \boldsymbol{b})$ is not adequate, since it is not difficult to see that this strategy would lead to a maximal number of discontinuities. This is not so if a "price to pay" $\alpha > 0$ is imposed for each discontinuity. The resulting **augmented criterion** to minimize is either $\mathcal{K}(x, \boldsymbol{t}) = \mathcal{J}_{\boldsymbol{t}}(x) + \alpha P$, or

$$K(\boldsymbol{x}, \boldsymbol{b}) = J_{\boldsymbol{b}}(\boldsymbol{x}) - \alpha \sum_{m=1}^{M-1} b_m \tag{6}$$

for the discrete counterpart. Here, only the number of discontinuities is penalized. Accounting for various kinds of information can also make penalization depend on their relative positions. For instance, neighboring discontinuities can be over-penalized if the "price to pay" is a decreasing function of the distance between them. When such "position-dependent" penalization is adopted, the edge variables are said *interacting*. For the purpose of image segmentation, specific forms of interacting edge variables have been introduced to produce closed contours (Geman and Geman, 1984; Marroquin et al., 1987; Jeng and Woods, 1991).

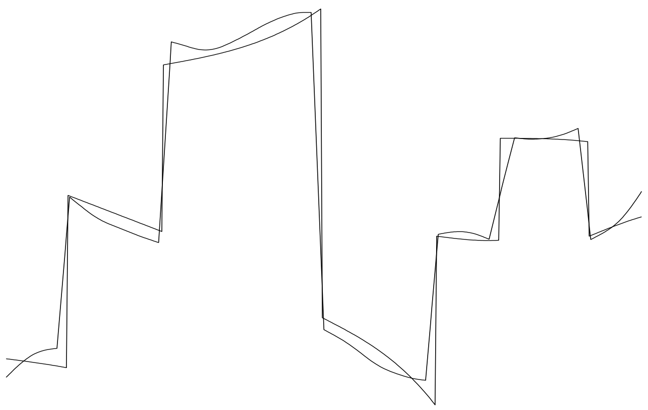

Figure 3: Piecewise smooth estimate $\widehat{\boldsymbol{x}}_{\lambda,\alpha}$ obtained as the joint minimizer of (6) for optimal values of λ, α in the L_1 sense: $E(\widehat{\boldsymbol{x}}_{\lambda,\alpha}, x^*) = 16.22\%$.

In the field of image processing and computer vision, many researchers worked within the detection-estimation framework during the eighties (Mumford and Shah, 1985; Terzopoulos, 1986; Blake and Zisserman, 1987). The common features of their contributions are

- the presence of binary edge variables in an augmented form of criterion. Such variables can be referred to as **hidden variables** w.r.t. the observation process, since they do not enter the observation equation. From both conceptual and practical viewpoints, handling hidden variables constitutes a very powerful tool to incorporate sophisticated prior knowledge.
- In most cases, the augmented criterion is **half-quadratic** (HQ): a function K will be said HQ if it depends on two sets of variables, say, $\boldsymbol{x}$ and $\boldsymbol{b}$, so that K is a quadratic function of $\boldsymbol{x}$ (but not of $(\boldsymbol{x}, \boldsymbol{b})$).

Drawbacks

The main drawback of the detection-estimation scheme lies in its computational load. Numerically heavy methods are often required to tackle the combinatorial problem induced by the binary variables. Most of them are based on relaxation principles, either in a stochastic framework (*e.g., simulated annealing* – Geman and Geman, 1984) or in a deterministic framework (*e.g.,* continuation methods such as *graduated non convexity* (GNC) – Blake and Zisserman, 1987).

As pointed out by several authors (Bouman and Sauer, 1993; Li and Huang, 1995), the lack of stability may constitute another weakness of the detection-estimation approach. Mathematically, it corresponds to the fact that the estimated signal $\widehat{\boldsymbol{x}}_{\lambda,\alpha}$ (obtained as the joint minimizer of (6)) is not necessarily a continuous function of the data, as illustrated by Figure 4, following an example given in Li and Huang (1995). In other words, the third Hadamard condition (Tikhonov and Arsenin, 1977) is not fulfilled by $\widehat{\boldsymbol{x}}_{\lambda,\alpha}$, so the problem is still ill-posed.

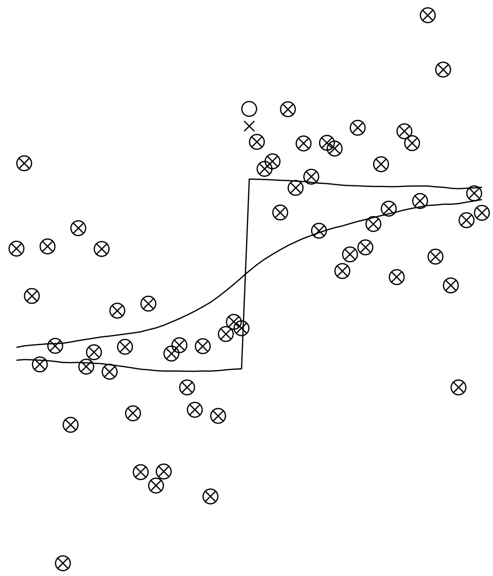

Figure 4: Instability of $\widehat{\boldsymbol{x}}_{\lambda,\alpha}$ as a function of data: in solid line, two estimated solutions $\widehat{\boldsymbol{x}}_{\lambda,\alpha}$ arising from the same data set, but for one point (data points are either depicted by circles or crosses).

Actually, $\widehat{\boldsymbol{x}}_{\lambda,\alpha}$ is "piecewise stable", and such a behavior is intrinsic to the edge detection capability of such an approach. In some cases, this is a desirable feature since it allows automatic decision-making. Otherwise, the detection-estimation approach is not recommended.

THE NON-QUADRATIC APPROACH

Nonquadratic penalization approaches have been taking much importance during the last years (Bouman and Sauer, 1993; Künsch, 1994; Li and Huang, 1995). The principle is to replace Tikhonov quadratic penalizer by an even function ϕ that could be better suited to the preservation of discontinuities. For instance,

$$\mathcal{J}_\phi(x) = \|\boldsymbol{y} - Ax\|^2 + \lambda \int_0^1 \phi(x''(t))\, dt \quad (7)$$

is a generalization of (2) for the continuous case, as well as

$$J_\phi(\boldsymbol{x}) = \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \frac{\lambda}{M} \sum_{m=1}^{M-1} \phi\left(\frac{2x_m - x_{m-1} - x_{m+1}}{1/M^2}\right), \quad (8)$$

for the discrete counterpart. In order to preserve sharp edges between homogeneous regions, it is quite clear that ϕ should be growing slower than a parabola, so that the cost associated to large steps be reduced. Two main families of such functions have been proposed in the litterature:

- "L_{21}" functions, *i.e.,* convex, continuously differentiable, asymptotically linear functions with a quadratic behavior near 0. A typical example is the hyperbola branch (Figure 5(b))
$$\phi(u) = \sqrt{s^2 + u^2},\ s > 0.$$
The corresponding minimizer of (8) is depicted by Figure 6. It provides a quite good compromise between the preservation of smooth *plateaux* and sharp edges. Other families of convex functions, such as "L_p" functions $\phi(u) = |u|^p$, $1 < p < 2$, yield the same qualitative behavior (Bouman and Sauer, 1993).

- "L_{20}" functions, *i.e.,* asymptotically constant functions with a quadratic behavior near 0. A typical example is (Geman and McClure, 1987; Figure 5(c))
$$\phi(u) = \frac{u^2}{r^2 + u^2},\ r > 0.$$
Such functions are nonconvex. As a mathematical consequence, global minimization makes sense in the finite-dimensional setting of (8), but not in the continuous setting of (7). The minimizer of (8) is not represented here. It has the same piecewise smooth character as the solution obtained with the detection-estimation approach (Figure 3). Other families of nonconvex functions, such as concave, increasing functions on $\mathbb{R}_+$, *e.g.,* $\phi(u) = |u| / (r + |u|)$, $r > 0$, have been recommended (Geman and Reynolds, 1992).

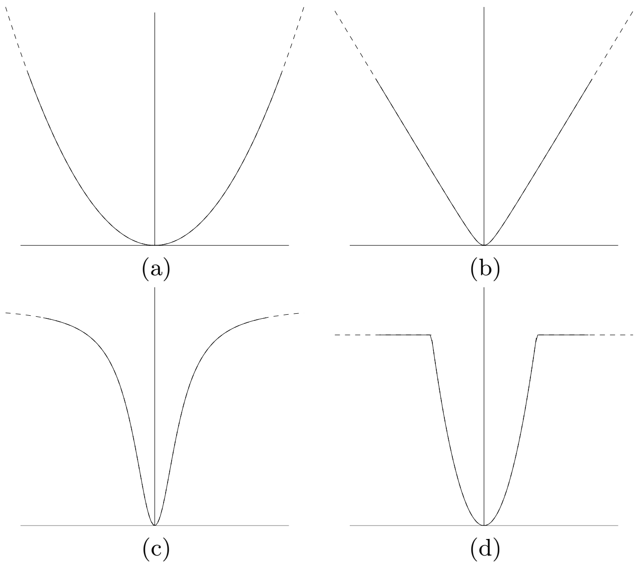

Figure 5: $\phi(u) =$ **(a)** u^2, **(b)** $\sqrt{s^2 + u^2}$, **(c)** $\frac{u^2}{r^2 + u^2}$, **(d)** $\min\{u^2, r^2\}$

Actually, the two classes of functions ϕ provide quite contrasted estimates in terms of behavior and of computational cost.

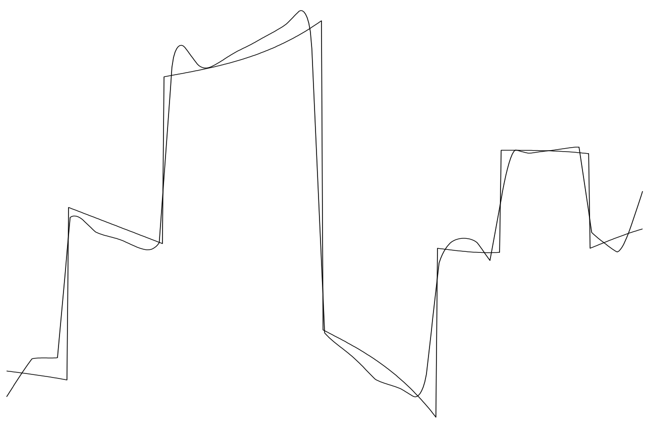

Figure 6: Smooth nonlinear estimate $\widehat{\boldsymbol{x}}_{\lambda,s}$ obtained as the minimizer of (8), with $\phi(u) = \sqrt{s^2 + u^2}$ for optimal values of λ, s in the L_1 sense: $E(\widehat{\boldsymbol{x}}_{\lambda,\alpha}, x^*) = 16.57\%$.

On one hand, the "L_{21}" approach yields convex criteria. Under simple technical assumptions, this allows to define a unique global minimizer for (8) and also for (7). Furthermore, convex criteria admit no local minima, so the convergence of standard minimization techniques is granted. Another interesting property is that the resulting solution is "robust" (Bouman and Sauer, 1993; Künsch, 1994; Li and Huang, 1995), *i.e.,* it fulfills the third Hadamard condition, contrarily to solutions obtained by detection-estimation.

On the other hand, the "L_{20}" approach shares the main characteristics of the detection-estimation approach: it is truly edge-detecting (and only piecewise stable), but it is also numerically demanding, because of the possible presence of local minima.

The similarity between the respective outputs of the detection-estimation approach and the "L_{20}" approach is not a coincidence, as explained in the next section.

HALF-QUADRATIC AUGMENTED CRITERIA

Handling HQ criteria has recently spread out as a powerful numerical device in the field of edge-preserving image restoration (Charbonnier et al., 1994; Brette and Idier, 1996; Cohen, 1996; Charbonnier et al., 1997; Vogel and Oman, 1998). As previously mentioned, HQ objective functions initially incorporated binary edge variables, either interacting or decoupled, within the detection-estimation approach. Mainly following Geman and Reynolds (1992), the present section shows that HQ criteria cover a much broader range of situations. More precisely, many non-quadratic penalization approaches can be equivalently reformulated in a HQ framework. The consequences are twofold. From the formal viewpoint, such an equivalence provides better insight into the real issues of signal and image modeling. From the practical viewpoint, new minimization tools for non-quadratic criteria derive from the HQ formulation.

Duality between non-quadratic and HQ augmented criteria

In the case of a HQ criterion with decoupled binary edge variables, such as $K(\boldsymbol{x},\boldsymbol{b})$ defined by (6), Blake and Zisserman (1987) pointed out that it could be considered as an *augmented equivalent* of a non-quadratic criterion $J(\boldsymbol{x})$, in the sense that

$$\min_{\boldsymbol{b}\in\{0,1\}^{M-1}} K(.,\boldsymbol{b}) = J(.).$$

More precisely, given (5) and (6), we have

$$\begin{aligned}
&\min_{\boldsymbol{b}\in\{0,1\}^{M-1}} K(\boldsymbol{x},\boldsymbol{b}) = \|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|^2 \\
&\quad+\sum_{m=1}^{M-1}\min_{b_m\in\{0,1\}}\left(\lambda M^3 b_m\,|2x_m-x_{m-1}-x_{m+1}|^2-\alpha b_m\right)\\
&=\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|^2+\sum_{m=1}^{M-1}\min\{0,\lambda M^3\,|2x_m-x_{m-1}-x_{m+1}|^2-\alpha\}\\
&=J_{\phi_r}(\boldsymbol{x})+const,
\end{aligned}$$

where J_{ϕ_r} is defined by (8) for the truncated quadratic function $\phi_r(u)=\min\{u^2,r^2\}$ (Figure 5(d)), and $r=\sqrt{\alpha/\lambda M}$.

As a consequence, J and K share the same minima, that can be seeked using any suited numerical device working on either J or K. In practice, Blake and Zisserman defined the HQ function K as the objective function, deduced J from K, and proposed a GNC approach to minimize the non-quadratic function J.

Later, D. Geman and co-workers' contributions (Geman and Reynolds, 1992; Geman and Yang, 1995) generalized Blake and Zisserman's construction to a larger class of decoupled auxiliary processes. In fact, they also reversed the construction process: they showed that there exist HQ augmented counterparts K for a wide range of non-quadratic edge-preserving criteria J, so that

$$\inf_{\boldsymbol{b}\in B} K(.,\boldsymbol{b}) = J(.) \tag{9}$$

for an appropriate set B, generally different from $\{0,1\}$.

In Geman and Reynolds (1992), the construction applies to edge-preserving functions ϕ that satisfy the following hypotheses:

$$\phi \text{ is even,} \tag{10}$$

$$\phi(\sqrt{.}) \text{ is concave on } \mathbb{R}_+, \tag{11}$$

$$\phi \text{ is continuous near zero and } C^1 \text{ on } \mathbb{R}^*. \tag{12}$$

Then the following duality relation can be obtained from convex analysis (Rockafellar, 1970):

$$\phi(u)=\inf_{b\in\mathbb{R}_+}\left(bu^2+\psi(b)\right), \tag{13}$$

where

$$\psi(b)=\sup_{u\in\mathbb{R}}\left(\phi(u)-bu^2\right). \tag{14}$$

Given (8) and (13), it is easy to deduce (9) for the HQ augmented criterion

$$K(\boldsymbol{x},\boldsymbol{b})=J_{\boldsymbol{b}}(\boldsymbol{x})+\frac{\lambda}{M}\sum_{m=1}^{M-1}\psi(b_m).$$

Such an equivalence has also been obtained for the continuous framework, in the more restrictive case of a convex function ϕ (Aubert and Vese, 1997).

Either convex or not, most regularizing functions brought in the litterature satisfy (10)-(12), as, for example, the two functions of Figure 5(b) and 5(c).

- For the hyperbola branch $\phi(u)=\sqrt{s^2+u^2}, s>0$, (14) yields

$$\psi(b)=\begin{cases} s^2b+1/4b & \text{if } b\in(0,b_\infty=1/2s]\\ 1 & \text{if } b\geq b_\infty\end{cases}$$

- For $\phi(u)=\dfrac{u^2}{r^2+u^2}, r>0$, (14) yields

$$\psi(b)=\begin{cases} (1-r\sqrt{b})^2 & \text{if } b\in(0,b_\infty=1/r^2]\\ 1 & \text{if } b\geq b_\infty\end{cases}$$

In both cases, the edge variables b_m are continuous, instead of being binary. However, in practice, their values are either close to 0 or to b_∞, especially in nonconvex cases.

Minimization of HQ criteria

Geman and Reynolds (1992) supported the idea that minimizing K rather than J has some structural advantages thanks to half-quadraticity (see also Geman and Yang (1995)). When ϕ is not convex, they developed simulated annealing techniques based on alternate pseudo-random sampling.

Other authors developped simpler deterministic counterparts (Charbonnier et al., 1994; Brette and Idier, 1996; Cohen, 1996; Charbonnier et al., 1997; Vogel and Oman, 1998; Delaney and Bresler, 1998), that benefit from half-quadraticity by alternating updates of $\boldsymbol{x}$ given $\boldsymbol{b}$, and of $\boldsymbol{b}$ given $\boldsymbol{x}$, *i.e.,* by intertwining the two following steps from an arbitrary initial couple $(\boldsymbol{x},\boldsymbol{b})$:

- Minimize $K(\boldsymbol{x},\boldsymbol{b})$ as a quadratic function of $\boldsymbol{x}$, while $\boldsymbol{b}$ is held constant. This is a simple linear programming problem that corresponds to an adaptive version of Tikhonov regularization, since a variable factor b_m is introduced for each penalizing term.

- Since K is a separable function of the variables b_m when $\boldsymbol{x}$ is held constant, it can be minimized as a function of $\boldsymbol{b}$ in a parallel form. Moreover, the updating equation for each b_m is explicit: it can be shown that the infimum of (13) is reached at

$$\widehat{b}(u)=\begin{cases} b_\infty & \text{if } u=0,\\ \phi'(u)/2u & \text{otherwise,}\end{cases} \tag{15}$$

and the expression of ψ is not even required to compute $\widehat{b}(u)$.

When ϕ is convex, such a coordinate descent procedure has been shown to converge to the unique minimum under broad conditions (Charbonnier et al., 1997; Idier, 1999). On the other hand, a technically similar procedure is long since known in a different context, as a *reweighted least squares* method (Yarlagadda et al., 1985). Although many other optimization techniques (*e.g.,* gradient-based) can reach the minimizer of J under the same regularity conditions, the reweighted least squares approach has one remarkable practical advantage: it allows to skip from Tikhonov quadratic penalization to almost any non-quadratic extension in a straightforward way. It only requires that a variable factor b_m be introduced for each penalizing term, and that each b_m be iteratively updated according to (15).

When ϕ is not convex, the reweighted least squares technique converges to the local minima of J (Delaney and Bresler, 1998), akin to other deterministic descent algorithms. Numerically heavier techniques, such as simulated annealing, are required to avoid local minima (Geman and Reynolds, 1992; Geman and Yang, 1995).

CONCLUSION

Within the penalization framework for the reconstruction of piecewise regular objects, three alternatives have been presented. Ordered by growing complexity:

- Tikhonov original approach corresponds to *quadratic* penalizers. If the observation equation is linear, computing the corresponding solution simply amounts to linear inversion.
- *Non-quadratic convex* penalizers provide robust estimates that better perserve edges. Such solutions can be iteratively computed using deterministic descent methods such as gradient descent. In the present paper, a reweighted least squares approach has been put forward. It is also a deterministic descent method, that performs iterative linear inversion. Formally, it identifies with a coordinate descent strategy for an equivalent half-quadratic augmented criterion.
- The *detection-estimation* strategy truly allow the recovery of distinct zones with abrupt changes. However, this is achieved at the expense of increased numerical cost. On the other hand, such a formulation is the half-quadratic counterpart of *non-convex* penalization approach (more specifically, using truncated quadratic functions).

Realistic applications of the presented methodology (particularly, of the convex penalization approach) can be found in several domains, such as image deblurring (Brette and Idier, 1996), spectral analysis (Ciuciu et al., 1999), computed tomography, under the assumption of a linear observation equation (Bouman and Sauer, 1993; Charbonnier et al., 1997). Nonlinear observation equations have been also considered within the same framework, *e.g.,* for inverse scattering (Carfantan and Mohammad-Djafari, 1996) and electrical impedance tomography (Martin and Idier, 1996). However, such nonlinear cases are more complex, because convex penalizers do not necessarily induce convex objective functions anymore, since the fidelity-to-data term becomes a non-quadratic (and presumably nonconvex) function of the unknown object.

*

REFERENCES

Aubert, G. and Vese, L. 1997. A variational method in image recovery. *SIAM J. Num. Anal.*, 34(5):1948–1979.

Blake, A. and Zisserman, A. 1987. *Visual reconstruction.* The MIT Press, Cambridge.

Bouman, C. A. and Sauer, K. D. 1993. A generalized Gaussian image model for edge-preserving MAP estimation. *IEEE Trans. Image Processing*, IP-2(3):296–310.

Brette, S. and Idier, J. 1996. Optimized Single Site Update Algorithms for Image Deblurring. In *Proc. IEEE ICIP*, pp. 65–68, Lausanne, Switzerland.

Carfantan, H. and Mohammad-Djafari, A. 1996. Beyond the Born approximation in inverse scattering with a Bayesian approach. In *2nd International Conference on Inverse Problems in Engineering*, pp. 439–446, Le Croisic, France.

Charbonnier, P., Blanc-Féraud, L., Aubert, G., and Barlaud, M. 1994. Two Deterministic Half-Quadratic Regularization Algorithms for Computed Imaging. In *Proc. IEEE ICIP*, volume 2, pp. 168–172.

Charbonnier, P., Blanc-Féraud, L., Aubert, G., and Barlaud, M. 1997. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Processing*, IP-6(2):298–311.

Ciuciu, P., Idier, J., and Giovannelli, J.-F. 1999. Markovian High Resolution Spectral Analysis. In *Proc. IEEE ICASSP*, pp. 1601–1604, Phoenix, U.S.A.

Cohen, L. D. 1996. Auxiliary variables and two-step iterative algorithms in computer vision problems. *Journal of Mathematical Imaging and Vision*, 6:59–83.

Delaney, A. H. and Bresler, Y. 1998. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Trans. Image Processing*, IP-7:204–221.

Geman, D. and Reynolds, G. 1992. Constrained Restoration and Recovery of Discontinuities. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-14(3):367–383.

Geman, D. and Yang, C. 1995. Nonlinear Image Recovery with Half-Quadratic Regularization. *IEEE Trans. Image Processing*, IP-4(7):932–946.

Geman, S. and Geman, D. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-6(6):721–741.

Geman, S. and McClure, D. 1987. Statistical Methods for Tomographic Image Reconstruction. In *Proceedings of the 46th Session of the ISI, Bulletin of the ISI*, volume 52, pp. 5–21.

Hunt, B. R. 1977. Bayesian Methods in Nonlinear Digital Image Restoration. *IEEE Trans. Communications*, C-26:219–229.

Idier, J. 1999. Convex half-quadratic criteria and interacting auxiliary variables for image restoration. Technical Report submitted to IEEE Trans. Image Processing, GPI–LSS.

Jeng, F. C. and Woods, J. W. 1991. Compound Gauss-Markov random fields for image estimation. *IEEE Trans. Signal Processing*, SP-39(3):683–697.

Künsch, H. R. 1994. Robust priors for smoothing and image restoration. *Annals of Institute of Statistical Mathematics*, 46(1):1–19.

Li, S. Z. and Huang, J. S. F. 1995. Convex MRF potential functions. In *Proc. IEEE ICIP*, volume 2, pp. 296–299.

Marroquin, J., Mitter, S., and Poggio, T. 1987. Probabilistic solution of ill-posed problems in computational vision. *J. Amer. Stat. Assoc.*, 82:76–89.

Martin, T. and Idier, J. 1996. A Bayesian non-linear inverse approach for electrical impedance tomography. In *2nd Intern. Conf. Inverse Problems in Engng.*, pp. 473–480, Le Croisic, France.

Mumford, D. and Shah, J. 1985. Boundary Detection by Minimizing Functionals. In *Proc. IEEE ICASSP*, pp. 22–26.

Rockafellar, R. T. 1970. *Convex Analysis*. Princeton University Press.

Terzopoulos, D. 1986. Regularization of Inverse Visual Problems Involving Discontinuities. *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-8(4):413–424.

Tikhonov, A. and Arsenin, V. 1977. *Solutions of Ill-Posed Problems*. Winston, Washington DC.

Vogel, R. V. and Oman, M. E. 1998. Fast, robust total variation-based reconstruction of noisy, blurred images. *IEEE Trans. Image Processing*, IP-7(6):813–823.

Yarlagadda, R., Bednar, J. B., and Watt, T. L. 1985. Fast algorithms for l_p deconvolution. *IEEE Trans. Acoust. Speech, Signal Processing*, ASSP-33(1):174–182.