Iterative wavelet–based denoising methods and robust outlier detection

R. Ranta, V. Louis-Dorr, C. Heinrich, and D. Wolf

***Abstract*— The goal of this letter is to study convergence conditions for a previously presented iterative wavelet denoising method [1] and to shed light on its relationship with outlier rejection. This method involves a user-defined parameter, which must fulfill certain conditions in order to ensure denoising. Using generalized Gaussian modelling for the wavelet coefficients distribution, we obtain a lower bound for this parameter, adapted to the shape of the distribution. Thresholding of the wavelet coefficients can then be achieved with a parameter-free algorithm. The properties of this threshold are examined and the proposed method is compared with other classical rejection methods.**

I. Introduction

The starting point of the research presented in this letter is the iterative wavelet–denoising method initially proposed by Starck and Bijaoui [2] and Coifman and Wickerhauser [3], [4], applied by Hadjileontiadis *et al.* [5] to physiological sounds analysis. This method is particularly adapted to non-stationary transient extraction from stationary (but not necessarily Gaussian) noise [2], [5]. The denoised signal is estimated using an iterative scheme, yielding successive refinements of this signal: at each iteration, the largest wavelet coefficients of the residual noise contribute to the current estimate of the denoised signal.

From a statistical point of view, large wavelet coefficients characterizing non-stationary transients can be considered as outliers, *i.e.*, points that strongly influence second order statistics as the standard deviation σ. Iterative thresholding algorithms can then be seen as *iterative outlier deletion* schemes (see Rousseeuw and Leroy [6, p. 254]), used to separate a subset of (small) noise-representative wavelet coefficients from the (large) outliers, further on used to reconstruct the denoised signal (one must notice that the method does not separate noisy artifacts from informative transients).

In a previous work [1], we have shown that if no best basis procedure is considered (as in [4]), iterative denoising (or, equivalently, iterative outlier deletion) may be seen as a fixed-point algorithm determining independent thresholds for each scale. The goal of this letter is to analyse and determine convergence conditions for this fixed-point algorithm by introducing generalized Gaussian (GG) modelling of the wavelet coefficients, without any prior information on the noise.

Manuscript received May 15, 2004; revised February 24, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. See May-Phoong.

R. Ranta, V. Louis-Dorr, and D. Wolf are with the Centre de Recherche en Automatique de Nancy – CRAN UMR CNRS-INPL-UHP 7039, 54516 Vandoeuvre, France (e-mail:radu.ranta@ensem.inpl-nancy.fr, valerie.louis@ensem.inpl-nancy.fr didier.wolf@ensem.inpl-nancy.fr).

C. Heinrich is with the Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection – LSIIT UMR CNRS-ULP 7005, 67400 Illkirch-Graffenstaden, France (e-mail: christian.heinrich@ensps.u-strasbg.fr).

Digital Object Identifier 10.1109/LSP.2005.851267

This letter is organized as follows. In the second section, we recall the iterative denoising principle, its fixed-point interpretation and the subsequent algorithm [1]. In the third section, we analyse the convergence of this algorithm to a non null fixed-point (*i.e.* denoising threshold) in the light of the probability law of the wavelet coefficients. The fourth section proposes modelling of the coefficients by GG probability density function (pdf) and introduces a convergence condition for the fixed-point algorithm depending on the GG parameters. The properties of the method both from outlier rejection and denoising points of view, as well as some simulation results, are presented in the Section V, followed by the conclusion.

II. Iterative wavelet–based denoising methods

We consider the model $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{n}$, where $\boldsymbol{z}$ is the given discrete–time signal to be denoised, $\boldsymbol{x}$ is the noise–free unknown version of $\boldsymbol{z}$ and $\boldsymbol{n}$ the noise. Orthogonal wavelet decomposition of $\boldsymbol{z}$ is written as

$$\boldsymbol{z} = \sum_{p,j} w_z^{j,p}\ \boldsymbol{\psi}^{j,p} + \sum_{p} w_z^{M,p}\ \boldsymbol{\phi}^{M,p},$$

where j is the scale, p the position, $\boldsymbol{\psi}$ the wavelet, $\boldsymbol{\phi}$ the scaling function and M the analysis depth [7].

Classical denoising methods as *VisuShrink* and *SureShrink*, introduced by Donoho and Johnstone [8], perform a one-step thresholding to separate supposedly Gaussian noise coefficients from large signal coefficients (see also [9] for different thresholding techniques). The iterative denoising scheme proposed in [2], [3], [5] writes $\boldsymbol{z} = \hat{\boldsymbol{x}}_k + \hat{\boldsymbol{n}}_k$, where k is the iteration step. The current noise estimation $\hat{\boldsymbol{n}}_k$, initialized for $k = 0$ as $\hat{\boldsymbol{n}}_0 = \boldsymbol{z}$, is decomposed to obtain the noise coefficients vector $\boldsymbol{\Omega}_{n,k}$. By thresholding $\boldsymbol{\Omega}_{n,k}$, one obtains the current "peeled off layer" $\boldsymbol{\Omega}_{\Delta x,k+1}$. The new noise coefficient vector $\boldsymbol{\Omega}_{n,k+1}$ is obtained from $\boldsymbol{\Omega}_{\Delta x,k+1} + \boldsymbol{\Omega}_{n,k+1} = \boldsymbol{\Omega}_{n,k}$.

In the case of iterative methods, threshold computing can be done either by considering a user-chosen proportion of large coefficients (which can be seen as a nonparametric outlier detection technique) or by classical ℓ_2 criteria, based on standard deviation of the coefficients σ (again a classical outlier rejection method based on Euclidean or Mahalanobis distance). Thresholding can be done independently for each scale of the decomposition or globally for the complete vector. We consider here the complete noise coefficients vector $\boldsymbol{\Omega}_{n,k}$ at iteration k and σ_k its standard deviation (scale by scale generalization is immediate). The nonstationary transients detection algorithm [2], [5], which uses an ℓ_2 iterative outlier deletion method and a unique orthogonal wavelet basis, writes:

1) compute σ_k as $(\sigma_k)^2 = \frac{1}{N}\ \|\boldsymbol{\Omega}_{n,k}\|^2$;

2) compute the threshold $T_{k+1} = F_a\sigma_k$, where F_a is a user–defined constant. A classical choice is $F_a = 3$, as in [2] and [5]. One can also recognize the threshold proposed by the RLS (*Reweighted Least Squares*) algorithm described in [6], which uses an empiric $F_a = 2.5$;
3) compute $\boldsymbol{\Omega}_{n,k+1}$, $\boldsymbol{\Omega}_{\Delta x,k+1}$ by T_{k+1}-thresholding $\boldsymbol{\Omega}_{n,k}$;
4) compute $\hat{\boldsymbol{n}}_{k+1}$ and $\Delta\boldsymbol{x}_{k+1}$ as wavelet reconstructions of $\boldsymbol{\Omega}_{n,k+1}$ and $\boldsymbol{\Omega}_{\Delta x,k+1}$ respectively, and set $\hat{\boldsymbol{x}}_{k+1} = \hat{\boldsymbol{x}}_k + \Delta\boldsymbol{x}_{k+1}$;
5) loop to the top until a certain stop criterion is reached.

The classical stop criterion is:

$$STC_{k+1} = \|\hat{\boldsymbol{n}}_k\|^2 - \|\hat{\boldsymbol{n}}_{k+1}\|^2 \leq \varepsilon, \tag{1}$$

with ε a (small) user chosen parameter. In a denoising framework, a small STC means that little or no energy is lost by the noise estimate between two steps. In an outlier rejection framework, choosing $\varepsilon = 0$ leads to the classical iterative outlier deletion algorithm described in [6, p. 254].

Adapting the scale by scale approach presented in [1] for $\varepsilon = 0$ to the complete vector of the wavelet coefficients, this algorithm boils down to a fixed point descent:

1) compute σ_k as $(\sigma_k)^2 = \dfrac{1}{N} \displaystyle\sum_{p,j} \left(w_z^{j,p}\ g_{T_k}(w_z^{j,p})\right)^2$, where $g_{T_k}(w_z^{j,p})$ are coefficients weights:

$$g_{T_k}(x) = \begin{cases} 1, & \text{if } |x| < T_k, \\ 0, & \text{if } |x| \geq T_k; \end{cases} \tag{2}$$

2) compute $T_{k+1} = F_a\sigma_k$ and loop till convergence.

Considering function f defined as

$$f(x) = F_a \sqrt{\frac{1}{N}\sum_{p,j}\left(w_z^{j,p}\ g_x(w_z^{j,p})\right)^2}, \tag{3}$$

the final threshold is computed by the fixed-point descent algorithm $T_{k+1} = f(T_k)$. The values of the function f depend on the user constant F_a, which must be lower bounded in order to ensure the existence of a non null fixed point (*i.e.* denoising threshold), otherwise the estimated noise vanishes ($\hat{\boldsymbol{n}} = 0$) and the denoised estimate $\hat{\boldsymbol{x}}$ equals $\boldsymbol{z}$ [1].

III. A PROBABILISTIC APPROACH

The goal of this section is to study the existence of the aforementioned lower bound for F_a in a probabilistic framework. In fact, wavelet coefficients w can be considered as a sample issued from a zero mean symmetric continuous probability density law $p(w)$ (the zero-mean condition can easily be adapted to non-zero previously estimated means). Under this assumption, the function (3) can be rewritten as:

$$f(T) = F_a\sigma_{|w|<T} = F_a\sqrt{2\int_0^T p(w)w^2 dw}. \tag{4}$$

Proposition 1 *Consider that the wavelet coefficients follow a probability density function $p(w)$ with a zero mean, finite variance and a mode in 0. A sufficiency condition for the existence of a final threshold $T_f \in [a,b]$ with $a, b > 0$ (i.e., of a non-null fixed-point for the function (3)) is $F_a \geq \sqrt{3/2ap(a)}$.*

Proof: A function $f(x)$ continuous on an interval $[a,b]$ has a fixed point in $[a,b]$ if $\forall\ x \in [a,b], f(x) \in [a,b]$. As the function (4) is monotone increasing, we must prove that there exists an interval $[a,b]$ with $f(b) \leq b$ and $f(a) \geq a$.

a) $f(b) \leq b$: Since $p(w)$ has a finite variance and F_a is finite, $M = F_a\sigma_w$ exists and

$$\forall\ b \geq M, f(b) = F_a\sigma_{|w|<b} \leq M \leq b. \tag{5}$$

b) $f(a) \geq a$: Let $a > 0$. Considering part integration:

$$[f(a)]^2 = 2F_a^2\left[p(w)\frac{w^3}{3}\Big|_0^a - \int_0^a p'(w)\frac{w^3}{3}dw\right].$$

Under the initial hypothesis, $p(w)$ has a mode in 0. One can find a such as $p(w)$ is decreasing on $]0,a]$. Then, the derivative $p'(a) < 0$, so $\int_0^a p'(w)\frac{w^3}{3}dw < 0$ and

$$[f(a)]^2 > 2F_a^2 p(a)\frac{a^3}{3}. \tag{6}$$

Hence the implication: $2F_a^2 p(a)\frac{a^3}{3} \geq a^2 \Rightarrow [f(a)]^2 \geq a^2$, so

$$F_a \geq \sqrt{\frac{3}{2ap(a)}} \Rightarrow f(a) \geq a. \tag{7}$$

Consequently, $f(x)$ has at least one fixed point $T_f \in [a,b]$. ■

IV. GENERALIZED GAUSSIAN MODELLING

In the previous section, we have shown that under certain realistic conditions on the pdf of the wavelet coefficients, a multiplicative constant F_a greater than a defined value (7) ensures the convergence of the algorithm to a non null point. This section aims to find a precise way to compute this lower bound[1], noted F_{am}, using a generalized Gaussian (GG) model for the pdf of the wavelet coefficients $p(w)$:

$$p_{\sigma,u}(w) = \alpha e^{-|\beta w|^u} \quad \text{with} \tag{8}$$

$$\beta = \frac{1}{\sigma}\sqrt{\frac{\Gamma(3/u)}{\Gamma(1/u)}},\ \alpha = \frac{\beta u}{2\Gamma(1/u)},\ \Gamma(u) = \int_0^\infty e^{-x}x^{u-1}dx,$$

where σ is the standard deviation and $u > 0$ is the shape parameter of the probability law (u=2 for a Gaussian and u=1 for the Laplace pdf). Generalized Gaussians are widely used to model wavelet coefficients distribution (especially in natural images processing, see for example [10]) and respect the conditions of proposition 1 ($p'(w) < 0, \forall w > 0$, symmetric about the mean and finite even moments).

In (8), parameter β is positive. Inequality (7) can then be written (for any $a > 0$):

$$F_a \geq \sqrt{\frac{3}{2a\alpha}e^{(\beta a)^u}}, \text{ that is } \ a \geq \frac{3}{2\alpha F_a^2}e^{(\beta a)^u}. \tag{9}$$

The goal is to determine the lowest value of F_a (denoted F_{am}) such that there exists one $a > 0$ verifying (9). Since the mapping $q : a \mapsto q(a) = \frac{3}{2\alpha F_a^2}e^{(\beta a)^u}$ is strictly convex, it can have 0, 1 or 2 intersection points with the identity line $y(a) = a$. Thus, F_{am} is obtained when $q(a)$ is tangent (at a

[1]Condition (7) being sufficient but not necessary, one might find lower values of F_a ensuring the convergence of the algorithm.

point of abscissa a_0) to the line $y = a$. The bound F_{am} and the abscissa a_0 must then verify $q(a_0) = a_0$ (for $F_a = F_{am}$) and $q'(a_0) = 1$ (also for $F_a = F_{am}$), which writes

$$\frac{3}{2\alpha F_{am}^2} u\beta^u a_0^{u-1} e^{(\beta a_0)^u} = 1. \tag{10}$$

A straightforward computation yields the bound F_{am}:

$$F_{am} = \sqrt{\frac{3\beta}{2\alpha}(ue)^{\frac{1}{u}}} = \sqrt{\frac{3\Gamma(\frac{1}{u})}{u}(ue)^{\frac{1}{u}}}. \tag{11}$$

The lower bound F_{am} is independent of σ and depends only on the shape parameter u (see fig. 1(a)). For denoising, a different $F_{am}(j)$ can be computed for each decomposition scale j.

V. Discussion

First of all, it is important to note that the initial standard deviation estimate σ_0 that we use in our algorithm (as in [2], [5]) is a classical second order one (the square root of the variance), different from the one proposed by Rousseeuw and Leroy [6] (outlier rejection) and Donoho and Johnstone [8] (wavelet denoising), which is based on the median and equals $1.4826\,\underset{i}{\mathrm{med}}\, r_i$, where $r_i = w_i - \hat{w}_i$ are the regression residuals for the coefficients w_i (in the univariate case $\hat{w}_i$ is the estimate of the location and it equals 0 for wavelet denoising).

1) Gaussian case: For Gaussian distributions (*u*=2), the computed value of F_{am} used to estimate the rejection threshold is ≈ 2.49, which is particularly close to the empirical value of 2.5 proposed by Rousseeuw and Leroy [6, p. 17]. Using this threshold, the final "robust scale estimate" σ^* of the LMS algorithm (*Least Median Squares* [6, p. 202]), is obtained as:

$$\sigma^* = \sqrt{\frac{\sum_{i=1}^N g_T(r_i) r_i^2}{\sum_{i=1}^N g_T(r_i) - d}}, \tag{12}$$

where N is the sample size, d is the dimension of the data space, $g_T(r_i)$ are the thresholding weights (2), computed for the empiric threshold $T = 2.5\sigma_0$. Thus, after the first iteration, our σ_1 is almost similar to LMS estimate σ^*. Still, as seen in the denominator of (12), a difference between the denoising framework and the outlier rejection one is that, in the latter, the estimate σ^* is computed using only the non-zero coefficients instead of considering the total number of points N, as in [1], [2] and [5]. The reason is that, from a denoising point of view, the estimate of σ is related to the energy of the signal. Moreover, our iterative algorithm computes an initial threshold $T_1 = F_{am}\sigma_0$, which is used further on to compute new estimates of σ_k. For $\sigma_0 = 1$, the final threshold (after convergence) is $T_K \approx 2.29$. It is interesting to notice that another outlier rejection threshold proposed in [6, p. 260] is $\sqrt{\chi^2_{d,0.975}}$, which equals 2.24 in univariate framework $d = 1$.

2) The generalized Gaussian case: Interesting points are the evolution (for different shapes u) of the final threshold T_K (solution of (4) for $F_a = F_{am}$, fig. 1(b)), as well as the probability of "large" coefficients w (*i.e.* $|w| > T_K$). Considering zero–mean GG pdf's, this probability writes:

$$p(|w| > T_K) = 1 - 2\int_0^{T_K} \alpha e^{-(\beta w)^u} dw. \tag{13}$$

Using the substitution $x = (\beta w)^u$ and the integral formula 3.381(1) in [11]: $\int_0^t x^{\nu-1} e^{-\mu x} dx = \mu^{-\nu}\gamma(\nu, \mu t)$, the above mentioned probability becomes (see fig. 1(c)):

$$p(|w| > T_K) = 1 - \frac{\gamma\left(\frac{1}{u}, (\beta T_K)^u\right)}{\Gamma\left(\frac{1}{u}\right)}. \tag{14}$$

In the above equations, $\gamma(a, z) = \int_0^z e^{-t} t^{a-1} dt$, with $Re(a) > 0$, is the lower incomplete Gamma function. The part of the variance corresponding to the large coefficients is (fig. 1(d)):

$$var(|w| > T_K) = 1 - \frac{\gamma\left(\frac{3}{u}, (\beta T_K)^u\right)}{\Gamma\left(\frac{3}{u}\right)}. \tag{15}$$

Results similar to (14) and (15) were obtained by Pizurica *et al.* [10] in a Bayesian denoising framework. As seen in fig. 1(c), for heavy-tailed distributions (small u) the probability of having a coefficient $|w| > T_K$ increases: if the number of outliers increases, the algorithm detects more and an energetic gain appears (fig. 1(d)). Conversely, for "compact" distributions (large u), outlier detection probability tends to zero[2]. Figure 1(b)–(d) also presents outlier rejection results for another two methods: classical 3σ rejection and Rousseeuw's LMS rejection. For the latter method the final threshold is obtained after a one-step iteration ($T_F = 2.5\sigma_0$), with the initial scale estimate computed as $\sigma_0 = 1.4826\,\mathrm{med}\,|w_i|$ (the median $\mathrm{med}\,|w_i|$ for GG pdf's is computed using the same integral formula 3.381 in [11]). Our method rejects more outliers than the LMS for a shape parameter $u \gtrapprox 1.75$ ($u \gtrapprox 0.8$ for iterative 3σ rejection). For distributions having $u \gtrapprox 6.6$, neither of both classic methods rejects any point. On the contrary, for peaked distributions, our method is more conservative than the LMS. Concerning the iterative 3σ rejection, its convergence to a non null point it is not guaranteed for $u \lessapprox 0.8$, as the computed F_{am} becomes greater than 3 (fig. 1(a)).

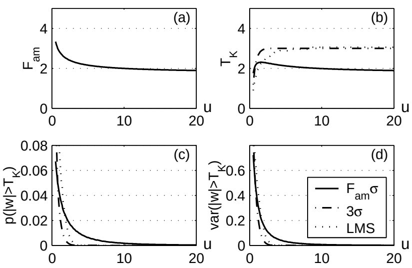

Fig. 1. For different values of the shape parameter u and a constant $\sigma = 1$: (a) evolution of F_{am}; (b) evolution of the final threshold T_K; (c) probability $p(|w| > T_K)$; (d) variance of coefficients w, $|w| > T_K$. Plots in (b)–(d) also present threshold, probability and variance evolution for two other outlier detection methods: iterative 3σ denoising [2], [5] and LMS [6].

3) Simulations: We conclude with two simulations: the first one presents the outlier rejection results of the algorithm, while the second one gives some denoising results.

Outlier rejection: We have randomly generated 9 samples of 10 000 points according to 3 pdf's (Laplacian, uniform and Gaussian), each one with 3 values of σ: 0.5, 1 and 2. A thousand outliers were generated according to two Gaussian laws (500 for each, means +/-5, $\sigma = 1$, see fig. 2). Three outlier

[2]Numerically, $F_{am} \to \sqrt{3}$ when $u \to \infty$, while the generalized Gaussian pdf tends towards an uniform law $[-\sqrt{3}, \sqrt{3}]$.

rejection algorithms (F_{am}, 3σ and LMS thresholding) were used to estimate the standard deviation σ on the remaining points. The mean absolute value of the estimation error ($|\hat{\sigma} - \sigma|$) obtained with our algorithm is about 6.5%, compared to 10.3% for 3σ thresholding and to 11.5% for LMS (see example fig. 2). Still, the LMS estimator performs better for the Gaussian sample with $\sigma = 2$, when outliers and data are highly superposed: $\hat{\sigma}_{LMS}$=2.11, $\hat{\sigma}_{F_{am}}$=2.15, $\hat{\sigma}_3$=2.21.

Denoising: We have used the four noise-free classic signals ($\boldsymbol{x}$ =*Blocks*, *Bumps*, *HeaviSine* and *Doppler*) proposed by Donoho and Johnstone [8]. They were normalized to a common arbitrary $\sigma_x = 3.5$. A first test was conducted on 2048 points signals [8] and a second one on the same signals with 1024 zeros added on both sides, thus enforcing their transient nature. Nine types of random noise $\boldsymbol{n}$ (Laplacian, uniform and Gaussian, each one with three values of σ_n: 0.5, 1 and 2) were added to the noise-free signals. As in [8], we performed a *sym8* wavelet decomposition of depth 5 and we computed denoised estimates $\hat{\boldsymbol{x}}$ using 4 hard-thresholding algorithms: iterative F_{am} and 3σ, *VisuShrink* and *SureShrink*. Results (mean square errors averaged for the 9 types of noise) are summarized in Table I: for the first test, F_{am} thresholding performs worse than *VisuShrink* but better than *SureShrink*, while it outperforms both for the second one. Further increase of the signals length by zero-padding leads to MSE of intermediate values between *VisuShrink* and *SureShrink*: more outliers are detected (mainly small scale coefficients, better represented in the sample due to the nature of the wavelet decomposition), and the signal has a noisier appearance. The same observation holds when applying iterative F_{am} thresholding independently for each scale, which again is quite natural.

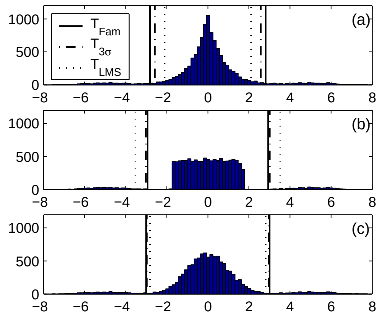

Fig. 2. Histograms and outlier rejection thresholds for simulated data ($\sigma = 1$): (a) Laplacian data ($T_{Fam} = 2.78, T_{3\sigma} = 2.58, T_{LMS} = 2.07$); (b) uniform data ($T_{Fam} = 2.93, T_{3\sigma} = 3.02, T_{LMS} = 3.54$); (c) Gaussian data ($T_{Fam} = 3.01, T_{3\sigma} = 2.98, T_{LMS} = 2.76$).

TABLE I

DENOISING RESULTS FOR THE 4 TESTED ALGORITHMS (MSE BETWEEN $\boldsymbol{x}$ AND $\hat{\boldsymbol{x}}$, AVERAGED OVER 10 SIMULATIONS FOR THE 9 TYPES OF NOISE).

	Blocks	*Bumps*	*HeaviSine*	*Doppler*
Test 1 – Original Signals [8]				
F_{am}	4.0117	4.5774	1.4094	2.0085
3σ	3.3536	3.6833	1.8107	2.2572
VisuShrink	3.4991	4.1804	1.0049	1.8386
SureShrink	4.2392	4.8318	2.2043	2.9815
Test 2 - Zero-padded signals				
F_{am}	2.0052	2.0447	0.9465	1.2317
3σ	2.1572	2.2584	1.4166	1.6380
VisuShrink	2.1919	2.5256	0.8654	1.3134
SureShrink	2.9444	3.3099	1.7489	2.2924

To conclude, it is important to note that iterative denoising/outlier rejection techniques, should be used for non-stationary transients extraction from stationary noise (as in [5]). The convergence condition we introduced, though it is not a necessary condition because of the minoration (6), favours a quasi-maximal information extraction. On the contrary, it does not guarantee a completely noise-free signal, as many "false alarms" may exceed the minimal threshold. Further treatments, taking into account *a priori* knowledge on the noise, may be used to separate informative outliers from *e.g.* impulsive noise.

VI. CONCLUSION

In this communication, the iterative fixed-point wavelet denoising method studied in [1] is related to more general outlier rejection techniques. Under certain conditions (reliable generalized Gaussian modelling of the wavelet coefficients), we propose a parameter free method for threshold computation. This method adapts to the shape of the distribution law and it ensures convergence to a non null point: it yields a non void robust estimation subset of representative points or, in a wavelet denoising framework, a non zero noise estimate. Under Gaussian assumptions, our method relates to classical outlier rejection methods as LMS- or χ^2-based thresholds [6].

In a wavelet denoising framework, the present method outperforms classical thresholding methods as *VisuShrink* ($T = \sigma_0\sqrt{2 \ln N}$) or *SureShrink* [8] under certain conditions (transient non-stationary signals), and it can be envisaged as an alternative for *SureShrink* in general situations. Moreover, it can be seen as a part of a more complete processing method, dedicated to a quasi-maximum information extraction.

In a more general framework, our method ensures a quasi optimal identification for "common", close to the (previously estimated) mean, points: this property is useful, for example, in robust parameter estimation in clustering algorithms.

REFERENCES

[1] R. Ranta, C. Heinrich, V. Louis-Dorr, and D. Wolf, "Interpretation and improvement of an iterative wavelet–based denoising method," *IEEE Signal Processing Letters*, vol. 10, no. 8, pp. 239–241, 2003.

[2] J.-L. Starck and A. Bijaoui, "Filtering and deconvolution by the wavelet transform," *Signal Processing*, vol. 35, pp. 195–211, 1994.

[3] R. Coifman and M. Wickerhauser, "Adapted waveform de-noising for medical signals and images," *IEEE Engineering in Medicine and Biology Magazine*, vol. 14, no. 5, pp. 578–586, 1995.

[4] ——, "Experiments with adapted wavelet de-noising for medical signals and images," in *Time-Frequency and Wavelets in Biomedical Engineering*, M. Akay, Ed. IEEE Press, 1998, pp. 323–346.

[5] L. Hadjileontiadis, L. Liatsos, C. Mavrogiannis, T. Rokkas, and S. Panas, "Enhancement of bowel sounds by wavelet-based filtering," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 7, pp. 876–886, 2000.

[6] P. Rousseeuw and A. Leroy, *Robust regression & outlier detection*. John Wiley & Sons, 1987.

[7] S. Mallat, *A wavelet tour of signal processing*. Academic Press, 1999.

[8] D. Donoho and I. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[9] A. Antoniadis, J. Bigot, and T. Sapatinas, "Wavelet estimators in nonparametric regression: a comparative simulation study," *Journal of Statistical Software*, vol. 6, no. 6, pp. 1–83, 2001.

[10] A. Pizurica, V. Zlokolica, and W. Philips, "Noise reduction in video sequences using wavelet-domain and temporal filtering," in *SPIE Conference "Wavelet Applications in Industrial Processing"*, Providence, Rhode Island, USA, October 2003.

[11] I. Gradshteyn and I. Ryzhik, *Table of integrals, series and products*. Academic Press, 1980.