

MODELING NON-GAUSSIAN NOISE FOR ROBUST IMAGE ANALYSIS

Sio-Song Ieng

ERA 17 LCPC, Laboratoire des Ponts et Chaussées, 23 avenue de l'amiral Chauvin, B.P. 69, 49136 Les Ponts de Cé, France
Sio-Song.Ieng@equipement.gouv.fr

Jean-Philippe Tarel

ESE, Laboratoire Central des Ponts et Chaussées, 58 Bd Lefebvre, 75015 Paris, France
Tarel@lpc.fr

Pierre Charbonnier

ERA 27 LCPC, Laboratoire des Ponts et Chaussées, 11 rue Jean Mentelin, B.P. 9, 67035 Strasbourg, France
Pierre.Charbonnier@equipement.gouv.fr

Keywords: Image Analysis, Statistical Approach, Noise Modeling, Robust Fitting, Image Grouping and Segmentation, Image Enhancement.

Abstract: Accurate noise models are important to perform reliable robust image analysis. Indeed, many vision problems can be seen as parameter estimation problems. In this paper, two noise models are presented and we show that these models are convenient to approximate observation noise in different contexts related to image analysis. In spite of the numerous results on M-estimators, their robustness is not always clearly addressed in the image analysis field. Based on Mizera and Müller's recent fundamental work, we study the robustness of M-estimators for the two presented noise models, in the fixed design setting. To illustrate the interest of these noise models, we present two image vision applications that can be solved within this framework: curves fitting and edge-preserving image smoothing.

1 INTRODUCTION

In computer vision, it is common knowledge that data are corrupted by non-Gaussian noise, outliers and may contain multiple statistical populations. It is a difficult task to model observed perturbations. Several parametric models were proposed in (Huang and Mumford, 1999; Srivastava et al., 2003), and sometimes based on mixtures (Hasler et al., 2003). Non-Gaussian noise models imply using robust algorithms to reject *outliers*. The most popular techniques are Least Median Squares (LMedS), RANSAC and Iterative Reweighted Least Squares (IRLS). The first two algorithms are close in their principle and achieve the highest *breakdown point*, i.e. the admissible fraction of outliers in the data set, of approximately 50%. However, their computational burden quickly increases with the number of parameters. In this paper, we focus on the IRLS algorithm, which is an extension of least-squares allowing non-Gaussian noise models, see (Huber, 1981). IRLS means Iterative Reweighted Least-Squares, where the weight λ is a particular function of the noise model at the value of the residual. One may argue that IRLS algorithm is a deterministic algorithm that only converges towards

a local minimum close to its starting point. This difficulty can be circumvented by the so-called Graduated Non Convexity (GNC) strategy. The IRLS algorithm, even with the GNC strategy, is usually very fast compared to LMedS and RANSAC and it is also able to achieve the highest breakdown point.

Indeed, let us consider the linear problem:

$$Y = XA + B \quad (1)$$

where $Y = (y_1, \dots, y_n) \in \mathbb{R}^n$ is a vector of observations, $X = (\mathbf{x}_1, \dots, \mathbf{x}_n) \in \mathbb{R}^{n \times d}$ the design matrix, $A \in \mathbb{R}^d$ the vector of unknown parameters that will be estimated by the IRLS algorithm, and $B = (b_1, \dots, b_n) \in \mathbb{R}^n$ the noise. The noise is assumed independent and identically distributed but not necessarily Gaussian. We consider the *fixed design* setting, i.e. in (1), X is assumed non-random. In that case, as demonstrated in (Mizera and Müller, 1999), certain M-estimators attain the maximum possible breakdown point. However, if X cannot be assumed non-random, the breakdown point of M-estimators drops towards zero (Meer et al., 2000). This underlines how important the way computer vision problems are formulated is. To our opinion, many problems can be expressed in the fixed design setting, allowing to apply the fast and efficient IRLS algorithm.

The paper is organized as follows. In Sec. 2, we present the non-Gaussian noise models we found of interest in diverse contexts, which is shown in Sec. 3. Then in Sec. 4, we prove that M-estimators that achieve the maximum breakdown point of approximately 50% can be built based on these probability distribution families. Finally in Sec. 5, we illustrate the interest of these theoretical results, by presenting two applications, casted in the fixed design setting: curves fitting for lane-marking shape estimation and edge-preserving image smoothing.

2 NON-GAUSSIAN NOISE MODELS

We are interested in parametric functions families that allow a continuous transition between different kinds of probability distributions. We here focus on two simple parametric probability density functions (pdf) of the form $pdf(b) \propto e^{-\rho(b)}$ suitable for the IRLS algorithm, where \propto denotes the equality up to a factor.

A first interesting family of pdfs is the exponential family (also called generalized Laplacian, or generalized Gaussian, or stretched exponential (Srivastava et al., 2003)):

$$E_{\alpha,s}(b) = \frac{\alpha}{s\Gamma(\frac{1}{2\alpha})} e^{-((\frac{b}{s})^2)^\alpha} \quad (2)$$

The two parameters of this family are the scale s and the power α , which specifies the shape of the noise model. Moreover, α allows a continuous transition between two well-known statistical laws: Gaussian ($\alpha = 1$) and Laplacian ($\alpha = \frac{1}{2}$). The associated ρ function is $\rho_{E_\alpha}(b) = ((\frac{b}{s})^2)^\alpha$.

As detailed in (Tarel et al., 2002), to guarantee the convergence of IRLS, $\frac{\rho'(b)}{b}$ has to be defined on $[0, +\infty[$, which is not the case for $\alpha \leq 0$ in the exponential family. Therefore, the so-called smooth exponential family (SEF) $S_{\alpha,s}$ was introduced in (Ieng et al., 2004):

$$S_{\alpha,s}(b) \propto \frac{1}{s} e^{-\frac{1}{2}\rho_\alpha(\frac{b}{s})} \quad (3)$$

where $\rho_\alpha(u) = \frac{1}{\alpha}((1+u^2)^\alpha - 1)$.

Similarly to the exponential family, α allows a continuous transition between well-known statistical laws such as Gauss ($\alpha = 1$), smooth Laplace ($\alpha = \frac{1}{2}$) and Geman & McClure ($\alpha = -1$). These laws are shown in Figure 1. For $\alpha \leq 0$, $S_{\alpha,s}$ can be always normalized on a bounded support, so it can still be seen as a pdf. In the smooth exponential family, when α is decreasing, the probability to have large, not to say very large errors (outliers) increases.

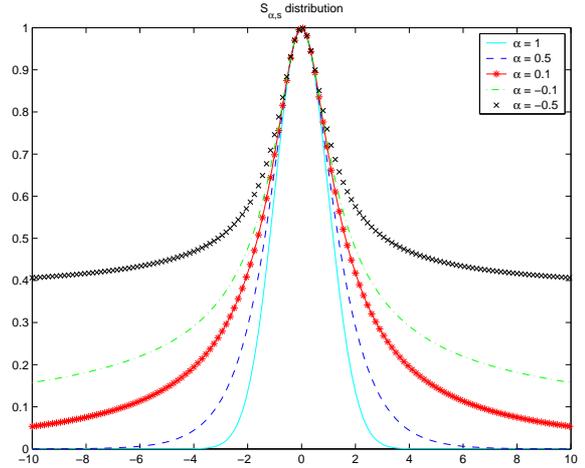


Figure 1: SEF noise models, $S_{\alpha,s}$. Notice how tails are getting heavier as α decreases.

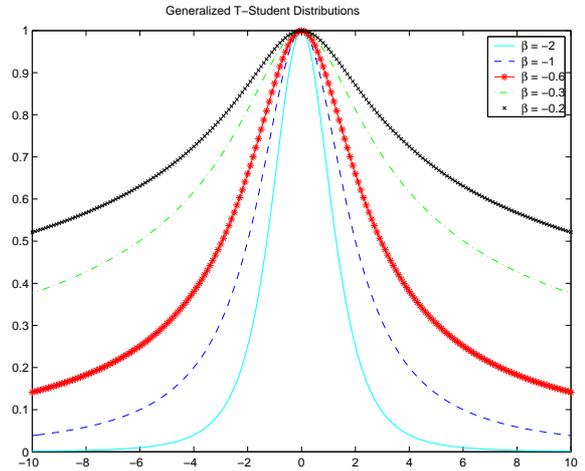


Figure 2: GTF noise models, $T_{\beta,s}$. Notice how tails are getting heavier when β is increasing towards 0.

In the IRLS algorithm, the residual b is weighted by $\lambda = \frac{\rho'(b)}{b}$. Notice that while the pdf is not defined when $\alpha = 0$, its weight does and it corresponds in fact to the T-Student law. Moreover, it is easy to show that the so-called generalized T-Student pdfs have the same weight function $\frac{1}{1+\frac{b^2}{s^2}}$ up to a factor. We define the Generalized T-Student Family (GTF) by:

$$T_{\beta,s}(b) = \frac{\Gamma(-\beta)}{\sqrt{\pi}\Gamma(-\beta-\frac{1}{2})s} (1 + \frac{b^2}{s^2})^\beta \quad (4)$$

where $\beta < 0$. This family of pdfs also satisfies the required properties for robust fitting. It is named generalized T-Student pdf (Huang and Mumford, 1999) in the sense that an additional scale parameter is intro-

duced compared to the standard T-Student pdf. Notice, that the case $\beta = -1$ corresponds to the Cauchy pdf. These laws are shown in Figure 2. The GTF can be rewritten as:

$$T_{\beta,s}(b) \propto \frac{1}{s} e^{-\frac{1}{2}\rho_{\beta}(b)} \quad (5)$$

where $\rho_{\beta}(b) = -2\beta \log(1 + \frac{b^2}{s^2})$.

The parameters of the GTF are s and β ($\beta < 0$). They play exactly the same role as s and α in the SEF. For $-\frac{1}{2} \leq \beta < 0$, as previously the pdf is defined only for a bounded support.

3 IMAGE NOISE MODELING

We have used with success SEF and GTF for noise modeling in different image analysis applications. We now illustrate this on two particular examples, where geometric and photometric perturbations are considered in turn, and modeled using the smooth exponential family (SEF).

3.1 Geometry

In this experiment, we have taken a set of 150 images of the same marking with different perturbations such as stones, shadows and so on, see Figure 3. The ground-truth position of the lane marking center was obtained by hand. It is shown in black in Figure 3.

Then, for each of these images, local marking centers are extracted, see (Ieng et al., 2004) for details, and the horizontal distances to each feature center with respect to the reference are collected. The MLE approach is then applied to estimate α and s at the same time by a non-linear minimization with a gradient descent. The best parameters within SEF are $\alpha = 0.05$ and $s = 1.1$. As shown in Figure 4, these parameters seems to lead to a nice noise model.

3.2 Photometry

We performed a similar experiment for photometric information. We have taken a set of 19 images of the same scene with different light conditions as shown on Fig 5. The ground truth is not easy to build contrary to the previous section. As a consequence, we compute, for each pair of images, the histogram of the differences in intensity between the two images, for each pixel. Rather than to sample the pdf of the perturbations, we sample the autocorrelation function of the pdf.

After fitting of the SEF model, we obtained a quite accurate model of the autocorrelation function as shown in Fig 6, with $\alpha = 0.02$ and $s = 9$.

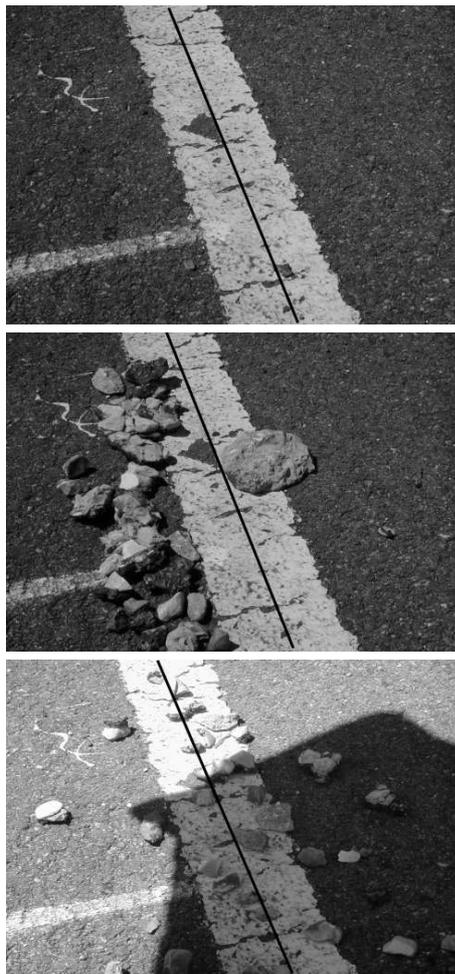


Figure 3: Three of the 150 images of the same marking with different perturbations. The black straight line shows the reference marking position.

4 ROBUSTNESS STUDY

Following (Mizera and Müller, 1999), in the fixed design case, robustness of an M-estimator is characterized by its breakdown point which is defined as the maximum percentage of outliers the estimator is able to cope with:

$$\varepsilon^*(\hat{A}, Y, X) = \frac{1}{n} \min \{m : \sup_{\tilde{Y} \in B(Y,m)} \|\hat{A}(\tilde{Y}, X)\| = \infty\} \quad (6)$$

where \tilde{Y} is a corrupted data set obtained by arbitrary changing at most m samples, B is the set of all \tilde{Y} : $B(Y,m) = \{\tilde{Y} : \text{card}\{k : \tilde{y}_k \neq y_k\} \leq m\}$ and $\hat{A}(\tilde{Y}, X)$ is an estimate of A from \tilde{Y} . It is important to notice that the previous definition is different from the one proposed in (Hampel et al., 1986) which is not suited

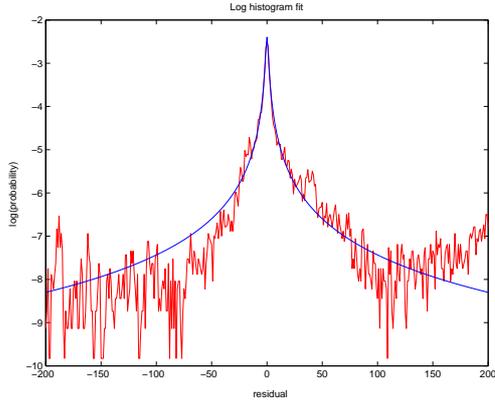


Figure 4: Log histogram of the residual errors in feature centers collected from 150 images after local marking extraction. The obtained distribution is well approximated by the best fit SEF (3) with parameters $\alpha = 0.05$ and $s = 1.1$.

to the fixed design setting.

Mizera and Müller (Mizera and Müller, 1999) also emphasize the notion of *regularly varying* functions, and described the link between this kind of regularity and robustness property. By definition, f varies regularly if there exists a r such that:

$$\lim_{t \rightarrow \infty} \frac{f(tb)}{f(t)} = b^r \quad (7)$$

When the exponent r equals zero, the function is said to vary slowly, i.e. the function is heavily tailed.

We now assume that the ρ function of the M-estimator follows the four following conditions:

1. ρ is even, non decreasing on \mathbb{R}^+ and nonnegative,
2. ρ is unbounded,
3. ρ varies regularly with an exponent $r \geq 0$,
4. ρ is *sub-additive*: $\exists L > 0, \forall t, s \geq 0, \rho(s+t) \leq \rho(s) + \rho(t) + L$.

The main result proved in (Mizera and Müller, 1999) is that the percentage ε^* is bounded by a function of r :

Theorem. Under the four previous conditions on ρ , and if $r \in [0, 1]$, then $\forall Y$ and X ,

$$\frac{M(X, r)}{n} \leq \varepsilon^*(\hat{A}, Y, X) \leq \frac{M(X, r) + 1}{n} \quad (8)$$

where, with the convention $0^0 = 0$, $M(X, r)$ is defined as:

$$M(X, r) = \min_{A \neq 0} \{ \text{card}(K) : \sum_{k \in K} |X_k^t A|^r \geq \sum_{k \notin K} |X_k^t A|^r \} \quad (9)$$

where K runs over the subsets of $\{1, 2, \dots, n\}$.



Figure 5: Three of the 19 images of the same scene with different lighting conditions.

When the exponent r is zero, the exact value of the percentage ε^* is known. The following theorem states its value.

Theorem. Under the four previous conditions on ρ and if $r = 0$, then $\forall Y$ and X ,

$$\varepsilon^*(\hat{A}, Y, X) = \frac{M(X, 0)}{n} = \frac{1}{n} \lfloor \frac{n - \mathcal{X}(X) + 1}{2} \rfloor \quad (10)$$

where $\lfloor x \rfloor$ represents the integer part of x , and $\mathcal{X}(X) = \max_{A \neq 0} \{ \text{card}\{X_k : X_k^t A = 0\}, k = 1, \dots, n \}$.

This value is also the maximum achievable value which is approximately 50%. As a consequence, M-estimators with zero r exponent are of highest breakdown point.

Finally in (Mizera and Müller, 1999), it is shown that the bounds on the percentage ε^* are related to r as a decreasing function. This is stated by the following lemma:

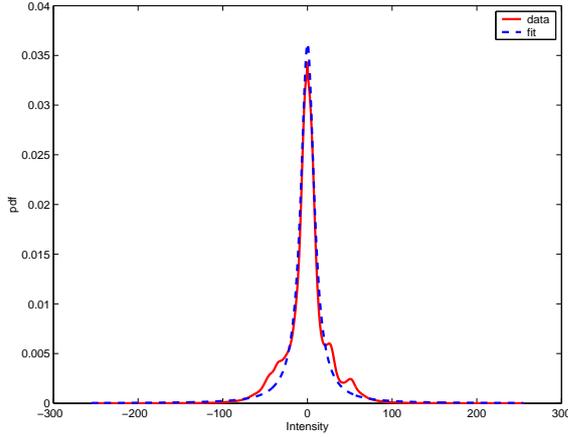


Figure 6: Histogram of the intensity differences collected from 19 images. The obtained distribution is well approximated by the best fit SEF (3) with parameters $\alpha = 0.02$ and $s = 9$.

Lemma. *If $q \geq r \geq 0$ then*

$$M(X, q) \leq M(X, r) \leq M(X, 0) = \lfloor \frac{n - \mathcal{N}(X) + 1}{2} \rfloor.$$

With these three results, the relationship between the exponent r of the function ρ and the robustness of the associated M-estimator was clearly established. As a consequence of these important theoretical results, the robustness of a large class of M-estimators can be compared just by looking at the exponent r of the cologarithm of their associated noise pdf. To illustrate this, we now apply the above results to the SEF and GTF pdfs described in Sec. 2.

4.1 The SEF case

Let us prove that the robustness of SEF based M-estimators is decreasing with $\alpha \in]0, 0.5]$. To this end, we first check the four above conditions on the ρ_α function. Function ρ_α is clearly even, non decreasing on \mathbb{R}^+ and nonnegative. The first condition is thus satisfied. The second one is fulfilled only when $\alpha > 0$, due to the fact that ρ_α is bounded for $\alpha \leq 0$. Looking at the ratio:

$$\frac{\rho_\alpha(tb)}{\rho_\alpha(t)} = \frac{(1 + \frac{t^2 b^2}{s^2})^\alpha - 1}{(1 + \frac{t^2}{s^2})^\alpha - 1} = \frac{(\frac{1}{t^2} + \frac{b^2}{s^2})^\alpha - \frac{1}{t^{2\alpha}}}{(\frac{1}{t^2} + \frac{1}{s^2})^\alpha - \frac{1}{t^{2\alpha}}}$$

we see that when $\alpha > 0$, $\lim_{t \rightarrow \infty} \frac{\rho_\alpha(tb)}{\rho_\alpha(t)} = b^{2\alpha}$. As a consequence, the ρ_α function varies regularly and the third condition is also satisfied. For the fourth condition, we can use Huber's Lemma 4.2 (Huber, 1984), to prove the sub-additivity when $\alpha \in]0, 0.5]$. For $\alpha = 0.5$, it can also be proved that ρ_α is sub-additive. All conditions on ρ being fulfilled, the first

Theorem applies, and using the Lemma, we prove that the robustness of SEF M-estimators is increasing towards the maximum of approximately 50%, with respect to a decreasing α parameter within $]0, 0.5]$.

4.2 The GTF case

Let us prove that the robustness of GTF M-estimators is maximum, whatever the value of β . We shall first check the 4 conditions on the associated ρ function. The first two assumptions are easy to check for ρ_β . Looking at the ratio:

$$\frac{\rho_\beta(tb)}{\rho_\beta(t)} = \frac{\ln(t^2) + \ln(\frac{1}{t^2} + \frac{b^2}{s^2})}{\ln(t^2) + \ln(\frac{1}{t^2} + \frac{1}{s^2})}$$

we deduce: $\lim_{t \rightarrow \infty} \frac{\rho_\beta(tb)}{\rho_\beta(t)} = 1 = b^0$. As a consequence, the ρ_β function varies slowly, and the third condition is fulfilled. The fourth condition is proved by using Huber's Lemma 4.2 (Huber, 1984). All conditions on ρ being satisfied, the second Theorem applies, showing that GTF M-estimators achieve the highest breakdown point of approximately 50%.

5 APPLICATIONS

We now describe two applications showing how interesting the use of SEF or GTF is in applying robust algorithms to problems related to geometry and photometry.

5.1 Curve Fitting for lane-marking tracking

This application is detailed in (Tarel et al., 2002; Ieng et al., 2004). The problem of tracking a lane marking can be handled using Kalman filtering, if the lane marking is robustly fitted in each image, which can be cast in the linear framework (1). In that case, $A = (a_1, \dots, a_d)$ are the parameters of a curve within a linearly parameterized family: $y = \sum_{j=1}^d a_j f_j(x)$. The road shape features (x_i, y_i) are given by the lane-marking centers extracted using a local feature extractor (Ieng et al., 2004). In (1), the noisy observations are the y_i . The vector $\mathbf{x}_i = (f_j(x_i))_{j=1, \dots, d}$ is assumed non random. The problem is thus set in the fixed design and maximum robustness can be achieved by modeling the geometric noise using the GTF family. Within the GTF family, the IRLS algorithm is used several times to refine the curve fitting result with decreasing scale s , until the scale of the noise is reached. The initial solution is obtained with

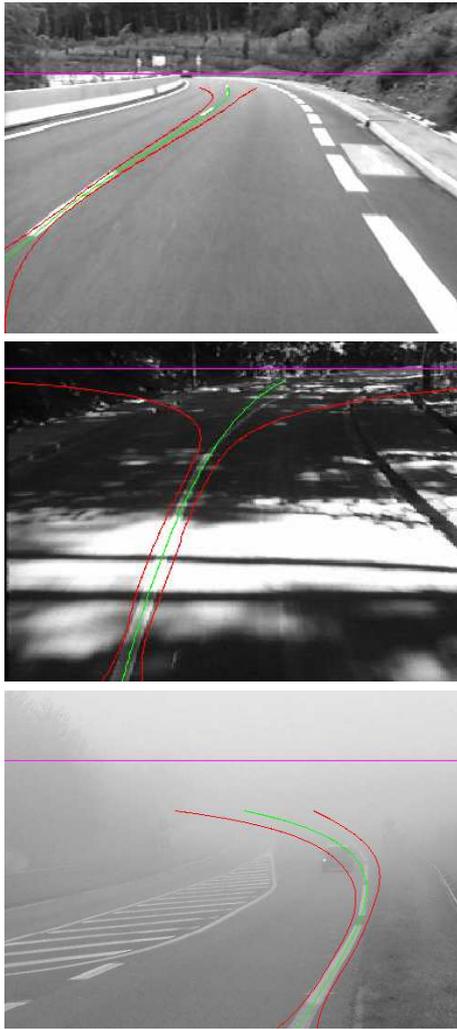
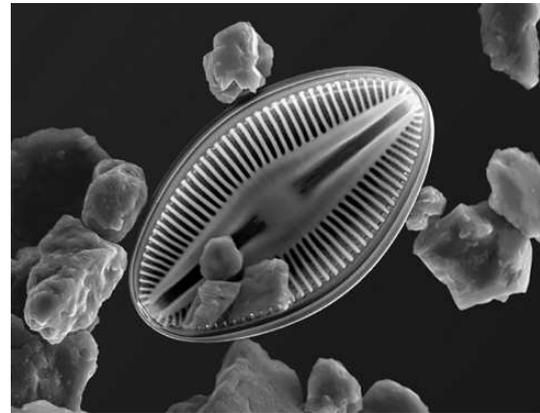


Figure 7: Three examples of detected lane-marking of degree 3 (in green). The uncertainty about the lane-marking location are also shown (in red).

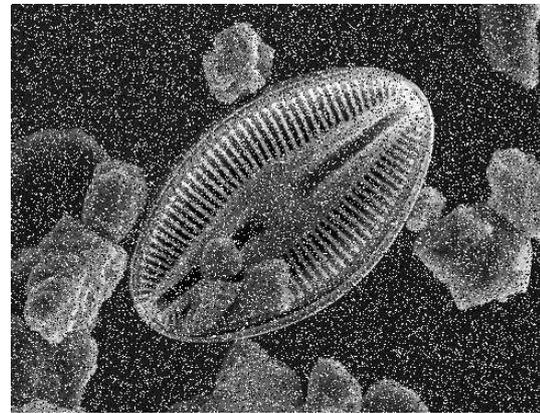
a scale large enough to imply a convex minimization problem. The SEF family can also be used if it better corresponds to the observed noise. In such a case rather than to decrease the scale, we refine the curve fitting by decreasing α step-by-step until the α of the observed noise is achieved. The initial solution is obtained with $\alpha = 0.5$ where the minimization problem is convex. These two strategies are examples of the GNC strategy.

Finally, the IRLS, with GNC strategy, allows curve tracking in real time, contrary to other robust methods such as LMedS and RANSAC. Three examples of main curve detection and tracking are shown in Fig. 7.

5.2 Edge-Preserving Image Smoothing



(a)



(b)

Figure 8: The original image (a), perturbed with 20% salt and pepper noise (b) (PSNR=11.5dB).

Image smoothing is an important topic in image analysis. Figure 8 shows an original image and the same image perturbed with photometric salt and pepper noise. As is well-known, using classical Gaussian filtering degrades edges, as shown in Figure 9(a). This motivated many works on nonlinear filtering (Astola and Kuosmanen, 1997) and edge-preserving image smoothing, see e.g. (Kervrann and Boulanger, 2006) for more references. In particular, bilateral filtering (Tomasi and Manduchi, 1998) is very intuitive because it is only a generalization of Gaussian smoothing that takes into account both spatial and intensity variations in the vicinity of each pixel. In (Elad, 2002; Kervrann, 2004), the theoretical background of bilateral filtering is explained which opens the possibility of using high breakdown point M-estimators derived from SEF and GTF noise models. Indeed, bilateral filtering can be seen as a linear estimation problem

$Y = A + B$ where Y is the observed image, A is the source image and B is the noise. Thus, bilateral filtering can be set as in (1) within the fixed design.

Let us consider a particular pixel (i, j) . Its observed intensity is $y_{i,j}$, and $a_{i,j}$ is the noiseless intensity that is to be estimated over a square neighborhood $[i - m, i + m] \times [j - m, j + m]$, assuming a non-Gaussian noise. To this end, the following error criterion is minimized:

$$e(a_{i,j}) = \sum_{k=-m}^m \sum_{l=-m}^m \rho(a_{i,j} - y_{i+k,j+l})k(k,l) \quad (11)$$

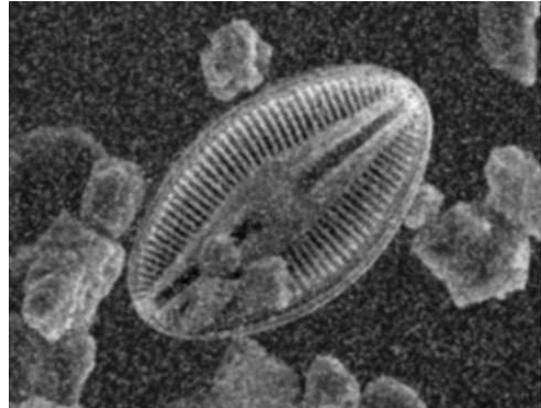
where ρ characterizes the noise model along intensities and k is a decreasing function w.r.t. the distance to the origin. This k allows to take into account the spatial distribution of the pixels and most of the time, a Gaussian kernel is used (Tomasi and Manduchi, 1998). One iteration of bilateral filtering directly consists of applying the IRLS algorithm derived from (11), for each image pixel.

In (Tomasi and Manduchi, 1998; Elad, 2002; Kervrann, 2004), it is suggested to use functions from the M-estimator theory, for ρ . ρ_α is an interesting candidate because it allows continuous transition between different kinds of pdfs, with increasing robustness when α decreases. Figure 9 displays the results obtained on the noisy image shown in Figure 8 for bilateral filtering with SEF and different values of α . When $\alpha = 1$, the obtained filter is equivalent to a weighted mean with Gaussian weights. When $\alpha = 0.5$, $\rho(x)$ approximates $|x|$ and hence, the filter behaves as a spatially-weighted L_1 norm filter. As can be seen on Figure 9, the lower the value of α , the better the restoration quality. However, when $\alpha < 0.5$, ρ_α becomes non-convex and the estimator can get stuck in a local minimum, resulting in poor results as show on Figure10(a).

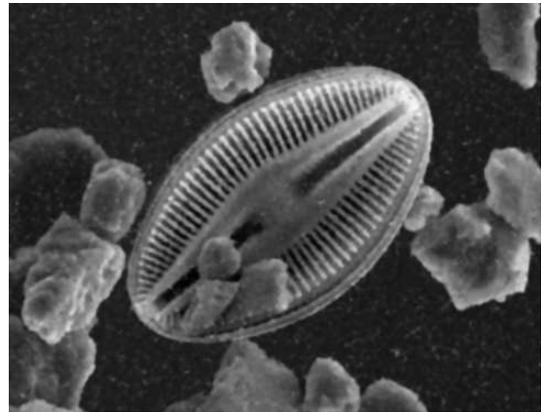
Similarly to (Kervrann, 2004) where the scale parameter is decreased regularly, we propose to decrease α step-by-step in a continuation strategy. This is another illustration of the effectiveness of the GNC strategy, see Figure 10(b).

6 CONCLUSION

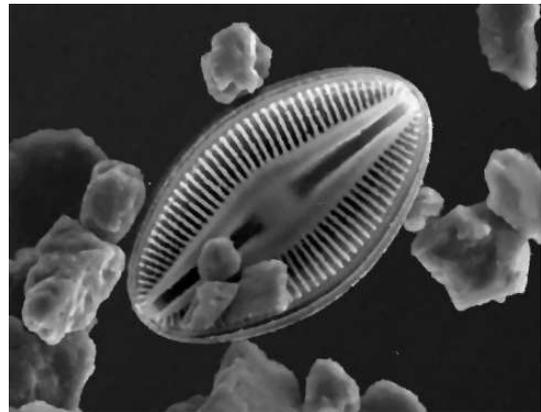
In this paper, we applied Mizera and Müller's fundamental work on M-estimators breakdown point calculation in the field of image analysis. In the *fixed design* setting, they shown that certain M-estimators can achieve maximum robustness. Using their results, we discussed the robustness of M-estimators based on two non-Gaussian pdfs families, that we introduced under the names of SEF and GTF. In particular, we shown that the GTF noise model leads to



(a)

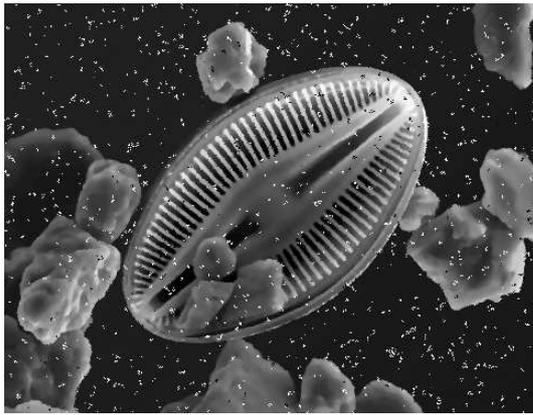


(b)

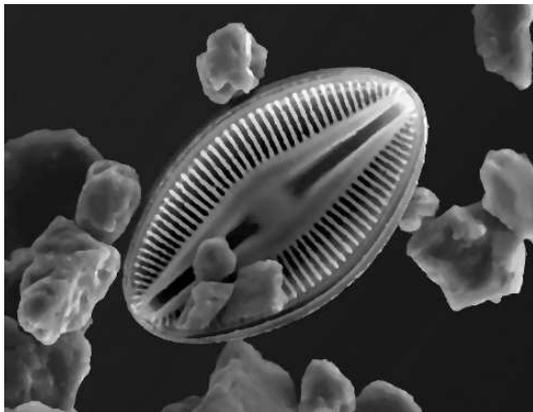


(c)

Figure 9: On the noisy image in Figure 8(b), a bilateral filtering is applied with different values of α within SEF: Gaussian-weighted mean filtering in (a) with $\alpha = 1$ (PSNR=20.3dB), $\alpha = 0.75$ in (b) (PSNR=25.2dB), and equivalent to weighted L_1 norm filtering in (c) with $\alpha = 0.5$ (PSNR=28.1dB). Notice how results are improved when α decreases.



(a)



(b)

Figure 10: The result after bilateral filtering with SEF and $\alpha = 0.25$ is in (a) (PSNR=19.6dB). Outliers still remains due to the fact that the corresponding p function is non convex. However, a better result is obtained in (b) using the GNC strategy (PSNR=28.1dB).

M-estimators that achieve the maximum breakdown point of approximately 50%, and that the robustness associated with SEF increases towards the maximum as α decreases towards 0.

In the second part of this paper, we illustrated how useful these results are in the context of image analysis: SEF and GTF approximate models seems to correctly fit observed noise pdfs in diverse applications and contexts. Moreover, many image analysis problems can be seen as parametric estimation or clustering. In the applications we shown (curve fitting and edge-preserving image smoothing), we observed the advantage of varying the noise model, progressively introducing robustness, with the so-called GNC strategy. We therefore believe that the SEF and GTF families can also be used with advantages in many other image analysis algorithms.

REFERENCES

- Astola, J. T. and Kuosmanen, P. (1997). *Fundamentals of nonlinear digital filtering*. CRC Press, Boca Raton, New York, USA, ISBN 0-8493-2570-6.
- Elad, M. (2002). The origin of the bilateral filter and ways to improve it. *IEEE Transactions on Image Processing*, 11(10):1141–1151.
- Hampel, F., Rousseeuw, P., Ronchetti, E., and Stahel, W. (1986). *Robust Statistics*. John Wiley and Sons, New York, New York.
- Hasler, D., Sbaiz, L., Ssstrunk, S., and Vetterli, M. (2003). Outlier modeling in image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(3):301–315.
- Huang, J. and Mumford, D. (1999). Statistics of natural images and models. In *Proceedings of IEEE conference on Computer Vision and Pattern Recognition (CVPR'99)*, pages 1541–1547, Ft. Collins, CO, USA.
- Huber, P. J. (1981). *Robust Statistics*. John Wiley and Sons, New York, New York.
- Huber, P. J. (1984). Finite sample breakdown of m- and e-estimators. *Annals of Statistics*, 12:119–126.
- Ieng, S.-S., Tarel, J.-P., and Charbonnier, P. (2004). Evaluation of robust fitting based detection. In *Proceedings of European Conference on Computer Vision (ECCV'04)*, pages 341–352, Prague, Czech Republic.
- Kervrann, C. (2004). An adaptive window approach for image smoothing and structures preserving. In *Proceedings of European Conference on Computer Vision*, pages 132–144, Prague, Czech Republic.
- Kervrann, C. and Boulanger, J. (2006). Unsupervised patch-based image regularization and representation. In *Proc. European Conf. Comp. Vision (ECCV'06)*, Graz, Austria.
- Meer, P., Stewart, C. V., and Tyler, D. (2000). Robust computer vision: An interdisciplinary challenge. *Computer Vision and Image Understanding: CVIU*, 78(1):1–7.
- Mizera, I. and Müller, C. (1999). Breakdown points and variation exponents of robust m-estimators in linear models. *The Annals of Statistics*, 27(4):1164–1177.
- Srivastava, A., Lee, A., Simoncelli, E., and Zhu, S. (2003). On advances in statistical modeling of natural images. *Journal of Mathematical Imaging and Vision*, 18(1):17–33.
- Tarel, J.-P., Ieng, S.-S., and Charbonnier, P. (2002). Using robust estimation algorithms for tracking explicit curves. In *European Conference on Computer Vision (ECCV'02)*, volume 1, pages 492–507, Copenhagen, Denmark.
- Tomasi, C. and Manduchi, R. (1998). Bilateral filtering for gray and color images. In *Proceedings of 6th International Conference on Computer Vision*, pages 839–846, New Delhi, India.