WEIGHTED MYRIAD FILTERS: A ROBUST FILTERING FRAMEWORK DERIVED FROM ALPHA-STABLE DISTRIBUTIONS *

Juan G. Gonzalez *Gonzalo R. Arce*

Department of Electrical Engineering
University of Delaware, Newark, DE 19716
E-mail correspondence: *gonzalez@ee.udel.edu*
WWW: *http://www.asel.udel.edu/ gonzalez/homepage.html*

ABSTRACT

In this paper we introduce a robust and nonlinear filtering framework: Weighted Myriad Filtering. Much like the Gaussian assumption has motivated the development of linear filtering theory, the formulation of *myriad filters* is motivated by the statistical properties of α-stable processes. *Weighted Myriad Filters* have a solid theoretical basis, are inherently more powerful than weighted median filters, and are very general subsuming traditional linear FIR filters. The foundation of the proposed filtering algorithms lies in the definition of the sample myriad as the location estimate for a class of α-stable distributions. In turn, the myriad has been discovered as the location parameter estimated by the sample myriad. This paper addresses some theoretical properties of myriad filters. The superior performance of myriad filters in impulsive environments is illustrated in the problem of robust synchronization by means of a "myriad phase lock loop".

1. INTRODUCTION

A large number of algorithms used in signal processing and communications rely on the fundamental assumption that the underlying signals and noise obey a Gaussian distribution. This assumption is often justified in practice due to the central limit theorem. The Gaussian assumption, however, often proves invalid as the noise encountered in practice frequently exhibits impulsive behavior. It is widely recognized that the performance of linear filters rapidly degrade in the presence of outliers.

Many approaches to the robust filtering problem have been formulated recently. Among these, the approach that has received considerable attention is that of weighted median filters [5]. Today, due to its sound underlying theory, weighted median filters are increasingly being used – software and hardware implementation of WMF filters are now common in image processing commercial products. The applications of weighted median filters, however, have not significantly spread beyond the field of image processing. This can be explained from the fact that median-based algorithms can be derived as optimal estimators in the presence of Laplacian noise, which is not a good model for the processes found in practice. An important shortcoming of the sample median is that its output value is constrained to be the same as that of one of the input samples. In effect, when compared to the mean, it is well known that the median loses as much as 40% efficiency as a location estimator [2].

In this paper we introduce a robust and nonlinear filtering framework derived from α-stable distributions. Alpha-stable random variables form an important class of processes, impulsive in nature, that obey a *Generalized Central Limit Theorem.*[1] Thus, like Gaussian random processes, α-stable processes can arise in practice as a result of physical principles.

Recently, considerable research efforts have been given to several problems in communications, within the framework of α-stable distributions [3, 4]. The filtering problem, however, has only been superficially addressed [1, 3].

Much like the Gaussian assumption has motivated the development of linear filtering theory, *myriad filtering theory* is motivated by the statistical properties of α-stable processes. *Weighted Myriad Filters* have a solid theoretical basis, are inherently more powerful than weighted median filters, and are very general subsuming traditional linear FIR filters. The foundation of the proposed filtering algorithms lies in the definition of the sample myriad as the location estimate of a class of α-stable distributions. In turn, the myriad has been discovered as the location parameter estimated by the sample myriad.

2. THE MYRIAD: A "NATURAL" LOCATION PARAMETER

To begin, it is crucial to look at the fundamentals of parameter estimation and at the characteristics of stochastic processes which may arise in real world applications. The processes we are mostly concerned are those that are impulsive in nature. These processes can be natural, as well as man-made, and include lighting in the atmosphere, motor ignitions, switching transients in power lines, and multiple access interference signals. While Gaussian models are clearly inappropriate, the class of α-stable distributions has been proven to accurately model impulsive-type processes [3]. On a first-order analysis, symmetric α-stable processes are characterized by their distribution having a characteristic function

$$\varphi(\omega) = exp(j\beta\omega - k|\omega|^{\alpha}), \quad (1)$$

where α is the characteristic exponent restricted in the range $0 < \alpha \leq 2$, β is the real-valued location parameter, and k is the dispersion of the distribution [3]. Two important cases of α-stable distributions emerge for $\alpha = 1$ and $\alpha = 2$, namely the Cauchy and Gaussian distributions respectively. Unfortunately, there are no closed form expressions for general α-stable symmetric distributions other than the Cauchy and the Gaussian. The parameter α determines how impulsive the process is. The smaller the parameter α is, the heavier the tails of the distribution. Alpha-stable processes are very useful in modeling signals encountered in practice since these obey two important properties:

*THIS WORK HAS BEEN PARTIALLY SUPPORTED BY A GRANT OF THE U.S. ARMY RESEARCH LABORATORY (ARL).

[1] Alpha-stable distributions are the only class of distributions that can be the limit for sums of continuous i.i.d. random variables.

0-7803-3192-3/96 $5.00©1996 IEEE

- Stability: $x_1, x_2, \cdots, x_n$ are independent α-stable random variables if and only if, for any constants $a_1, \cdots, a_n$, the linear combination $\sum_i a_i x_i$ is also an α-stable random variable.
- Generalized central limit theorem: The family of stable distributions contains all limiting distributions of sums of independent and identically distributed random variables.

The class of α-stable processes, obeying the above properties, can thus arise in practice as a result of physical principles. Laplacian impulsive models exploited by WMF, on the other hand, cannot be equally justified.

In order to develop robust and computationally tractable algorithms, it is logical to consider the maximum likelihood (ML) location estimation of a heavy-tailed α-stable distribution for which we have a closed form expression, namely the Cauchy distribution. Given a set of i.i.d. samples $x_1, x_2, \cdots, x_N$ obeying the Cauchy distribution with scaling factor k

$$f(x;\beta) = \frac{k}{\pi} \frac{1}{k^2 + (x-\beta)^2}, \qquad (2)$$

the ML location estimator is the value $\hat{\beta}_k$ which maximizes the likelihood function

$$L(x_1, \cdots, x_N; \theta) = \prod_{i=1}^{N} f(x_i;\beta) = \left(\frac{k}{\pi}\right)^N \prod_{i=1}^{N} \frac{1}{k^2 + (x_i - \beta)^2}.$$

This is equivalent to minimizing $\prod_{i=1}^{N} [k^2 + (x_i - \beta)^2]$. Thus, given $k > 0$, the ML location estimate, hereafter referred to as the sample *myriad*, is given by

$$\hat{\beta}_k = myriad\{k; x_1, \cdots, x_N\} = \arg\min_{\beta} \prod_{i=1}^{N} \left(k^2 + (x_i - \beta)^2\right). \qquad (3)$$

The parameter k determines the scale variability of the data, and for the Cauchy distribution it is equal to the interquartile range. If k is set to be very large, this implies that all the data are within the interquartile range of the Cauchy distribution model. As a result, the estimator considers the data to be "well behaved" (no outliers), in which case a desirable estimator of location would be the sample mean. Notably, the myriad satisfies:

Property 1 (Linear Property):

$$\lim_{k\to\infty} \hat{\beta}_k = \lim_{k\to\infty} myriad\{k; x_1, \cdots, x_N\} = mean\{x_1, \cdots, x_N\}. \qquad (4)$$

Proof: See Appendix A.

The structure of the sample myriad when k tends to zero, is also very interesting. In this case, the estimator considers practically all the data as unreliable, where most of the observations are modeled as lying outside the interquartile range. In this sense the most repeated value, if unique, will be considered as the most reliable indicator of location, and $\hat{\beta}_0$ becomes a mode-like estimator, which is expected to be highly resistive to the presence of outliers. When the most repeated value is not unique, it can be shown that the sample myriad $\hat{\beta}_0$, hereafter referred to as the *mode-myriad*, reduces to

$$\hat{\beta}_0 = \arg\min_{x_j \in M} \prod_{i=1, x_i \neq x_j}^{N} |x_i - x_j|, \qquad (5)$$

where M is the set of most repeated values.

Although the mode-myriad always lies on the most repeated values or most crowded neighborhood, it can be shown that it does not necessarily converge to the mode when the underlying distribution is not symmetric. The mode-myriad has several properties of interest. For instance, it is always equal to one of the input samples. This makes it very appropriate for image processing applications since it has the nice edge and detail preserving properties of selection-type filters [6]. The mode-myriad is also a shift and scale invariant estimate.

It is important to note that, for a general myriad, the availability of k as a tunable parameter allows the estimator to acquire some "intelligence", in the sense that the degree of linear (large k) or robust behavior (small k) can be inferred from the data by estimating an adequate value for k. To illustrate this, it is instructive to look at the behavior of the sample myriad shown in Fig. 1. The solid line shows the values of the myriad as a function of k for the data set $= \{0, 1, 3, 6, 7, 8, 9\}$. It can be appreciated that as k increases, the myriad tends asymptotically to the sample average. On the other hand, as k is decreased, the myriad favors the value 7 which indicates the location of the cluster formed by the samples $6, 7, 8, 9$. This is a typical behavior of the myriad for small k: It tends to favor values where samples are more likely to occur or cluster. The term "myriad" was coined as a result of this characteristic of the parameter.

The dotted line shows how the sample myriad is affected by an additional observation of value 100. For large values of k, the estimator is very sensitive to this new observation. On the contrary, for small k, the data variability is assumed to be very small, and the new observation is considered as an outlier, not influencing significantly the value of the myriad.

More interestingly, if the additional observations are the very large data $800, -500, 700$, (dashed curve), the myriad is practically unchanged for moderate values of k. This behavior exhibits a very desirable outlier rejection property, not found in median-type estimators.

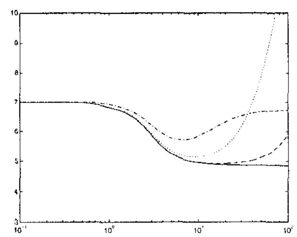

Figure 1. Values of the myriad as a function of k for the following data sets: (solid:)Original data set $= 0, 1, 3, 6, 7, 8, 9$; (dash-dot:)Original set plus an additional observation at 20; (dotted:)Additional observation at 100; (dashed:)Additional observations at 800, -500 and 700.

2.1. The Myriad As A Location Parameter

Much like the sample mean and sample median are the estimates of the mean and median parameters, the limiting case of the sample myriad defines the *myriad* as a new location parameter. It turns out that the myriad of a probability distribution function is the value β_k which minimizes the expectation $E\{log\left[k^2 + (x - \beta_k)^2\right]\}$, where $k \in [0, \infty]$ is a tunable parameter. For the case $k = 0$, the mode-myriad parameter takes on the value that minimizes $E\{log|x - \beta_0|\}$. It can be shown that the myriad is always in the center of symmetry, whenever the underlying distribution is symmetric. Thus, for any k, β_k is an adequate indicator of location. For non-symmetric distributions, the value of the myriad depends on k as illustrated next. Figure 2(a) and (b) depict the location of the myriad for different values of k, in the case of a bimodal distribution. For $k = 0$, the myriad cautiously localizes the distribution close to 8.5, the center of the dominant mode. As k increases the myriad is pulled to the value 8. Notice, however, that at $k = 1$ the value of the myriad suddenly jumps to 4. This is due to the fact that k is large enough so that both modes of the distribution are considered jointly reliable. For large k, the myriad is confident of all data and the location approaches the mean of the density function, close to the middle of the two modes.

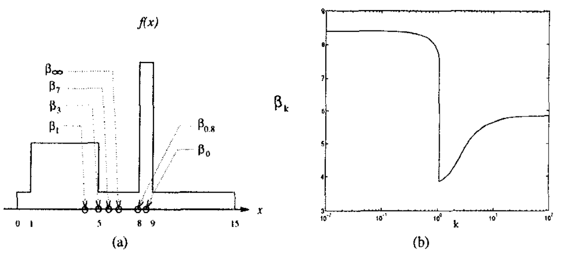

Figure 2. (a) A density function and the location of the myriad for several values of k. (b) The myriad as a function of k for this pdf.

2.2. Geometrical Interpretation

Myriad estimation, defined in (3), can be interpreted in a more intuitive manner. As depicted in Fig.3(a), it can be shown that the sample myriad, $\hat{\beta}_k$, is the value which minimizes the product of distances from point A to the sample points $x_1, x_2, \cdots, x_6$. Any other value, such as $x = \beta'$, produces a higher product of distances. As k is reduced, the myriad searches clusters as shown in Fig. 3(b). If k is made large, all distances become close and it can be shown that the myriad tends to the sample mean. This geometrical interpretation can lead to computationally efficient algorithms to find the value of the myriad.

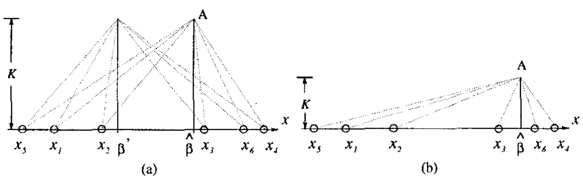

Figure 3. (a) The sample myriad, $\hat{\beta}$, minimizes the product of distances from point A to all samples. Any other value, such as $x = \beta'$, produces a higher product of distances; (b) the myriad as k is reduced.

3. WEIGHTED MYRIAD FILTERS

The formulation of weighted myriad filters is motivated by applications where certain samples used in forming an estimate are more reliable than others. Intuitively, an estimate should be modified taking into account the credibility level of each sample. The standard method to accomplish this is to, somehow, assign more weight to the most reliable samples. Thus a set of N weights, $W_1, \cdots, W_N$, must be defined, one for each input sample used in the estimate.

A very sound way to define these weights is intimately related to the maximum likelihood estimation procedure. Instead of considering the samples identically distributed, we can assume that our more reliable samples have a smaller dispersion around the center of the distribution. Thus, an observation assigned to a large weight W_i, can be related to a highly localized density function $W_i f[W_i(x_i - \beta)]$, where $f(.)$ is the primitive "unweighted" pdf. The limiting case in which $W_i = \infty$ relates the observation to an impulse density, which means that the sample is 100% reliable. On the other hand, a very small value of W_i indicates a large spread in the density function (almost flat), which implies a very poor chance of this observation to be close to the center of the distribution. The weighted likelihood function for independent samples with the same primitive pdf $f(.)$ results in $\prod_i W_i \; f[W_i(x_i - \beta)]$. For the Cauchy case, maximizing this function is equivalent to minimizing $\prod_{i=1}^{N}[k^2 + W_i^2(x_i - \beta)^2]$. Letting $w_i = W_i{}^2$ for convenience of notation, the *weighted myriad* is then defined as

$$\hat{\beta}_{k,\mathbf{w}} = myriad \; \{k;\; w_1 \circ x_1, \cdots, w_N \circ x_N\} \quad (6)$$

$$= \arg\min_{\beta} \prod_{i=1}^{N} \left[k^2 + w_i(x_i - \beta)^2\right]. \quad (7)$$

where $w_1, \cdots, w_N \geq 0$, and $w_i \circ x_i$ represents the weighting operation in (6). Since $\hat{\beta}_{k,c^2\mathbf{w}} = \hat{\beta}_{k/c,\mathbf{w}}$, it is clear that finding the optimal myriad filter weights will implicitly find the best k. As in the unweighted case, it can be shown that the linear property also holds for the weighted myriad. This is, as k tends to infinity, the weighted myriad tends to the weighted mean $\sum_{i=1}^{N} w_i x_i / \sum_{i=1}^{N} w_i$. Thus, the myriad filtering framework includes that of linear filters.

The design of the optimal coefficients for a weighted myriad filter is an important problem that we do not address here due to space limitations. A suboptimal and simpler approach, which we call *myriadization*, is described and tested in the following section.

4. FILTER DESIGN: MYRIADIZATION

The linear property indicates that for very large values of k, the weighted myriad filter reduces to a linear FIR filter. Since robust behavior is obtained whenever the value of k is small, a very simple method to design robust myriad filters is to use the weights of a linear filter ($k = \infty$), designed for Gaussian or noiseless environments, and to subsequently reduce k to a finite value, attaining the level of robustness desired. We refer to this method as *"myriadization"*, in contrast to the well known *"linearization"* approaches used in engineering.

The *myriadization* idea was tested in the phase synchronization problem of the first order Phase-Locked Loop (PLL) depicted in Fig. 4.

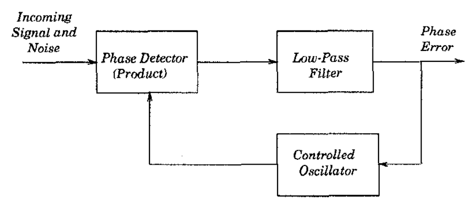

Figure 4. Block diagram of the Phase-Locked Loop analyzed in the paper.

Simulations were run in which the PLL had to track the carrier phase in an additive white Gaussian noise environment. The signal-to-noise ratio was set to 30 db, and the parameters of the system were adjusted so that the PLL was critically damped. A linear low-pass FIR filter was designed with 13 coefficients. Fig. 5 shows a typical phase error plot in which random noise bursts were present. During these short noisy intervals (from 4 to 10 sampling times), the signal-to-noise ratio was decreased to -10 db. It is evident from the figure that the system, when operating with the linear filter, is very likely to lose synchronism after one of these bursts. Fig. 6 shows the phase error under the same noise conditions for a PLL in which a weighted median filter has been optimally designed to imitate the low-pass characteristics of the original linear filter. Although the short noise bursts do not affect the estimate of the phase, the variance of the estimate is very large. Fig. 7 shows the phase error of the system with the same noise conditions, after the low-pass filter has been myriadized using a parameter k equal to half the carrier amplitude. Although phase error is increased during the bursts, the performance of the myriadized PLL is not degraded, and the system does not lose synchronism. More interesting is the fact that even with the normal low-amplitude Gaussian noise, the myriadized system shows a smaller steady-state variance, while maintaining the same synchronization responsivity.

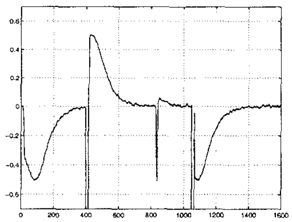

Figure 5. Phase error plot for the PLL with a linear FIR filter. Synchronism is lost due to small duration high amplitude noise bursts.

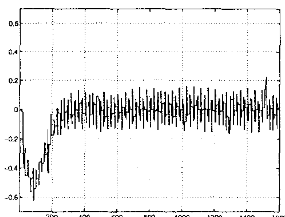

Figure 6. Phase error plot for the PLL with an optimal weighted median filter. Although the system is immune to the noise bursts, the error variance is significantly large.

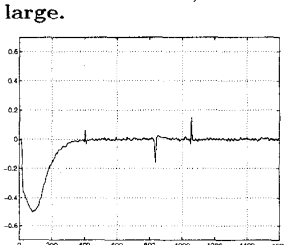

Figure 7. Phase error after the linear filter has been *myriadized*. Noise bursts do not degrade the system performance.

5. CONCLUSIONS

We have introduced weighted myriad filtering as a nonlinear framework that derives important robustness properties from the impulsive characteristics of symmetric α-stable distributions. In order to develop computationally tractable algorithms, the sample myriad has been introduced as the maximum likelihood location estimator for the Cauchy distribution, (the only non-Gaussian symmetric α-stable distribution for which a closed solution is known). The definition of the sample myriad as a location estimator implies the discovery of the myriad as a location parameter of statistical distributions.

When weights are considered in the definition, the weighted myriad filter appears as a very rich and flexible class of filters that can range, by only varying a tuning parameter k, from highly robust mode-like operations to simple and efficient linear FIR filters.

We have shown in a single example applied to the problem of robust synchronization, that weighted myriad filters have the potential to perform significantly better than both, linear and median filters in Gaussian and non-Gaussian environments.

This new field of research opens many interesting and important questions. Some of them, such as optimal filter optimization, adaptive design, as well as the extension of the myriad filters to bandpass types (negative weights) and vector operations, are subject of current research by the authors, and results are expected to be published soon.

A PROOF OF THE LINEAR PROPERTY

First note that $\hat{\beta}_k \leq x_{(N)} = \max\{x_1, \cdots, x_N\}$ by checking that for $\beta > x_{(N)}, k^2 + (x_i - \beta)^2 > k^2 + (x_i - x_{(N)})^2$. In the same way, $\hat{\beta}_k \geq x_{(1)} = \min\{x_1, \cdots, x_N\}$. Hence,

$$\begin{aligned}
\hat{\beta}_k &= \arg \min_{x_{(1)} \leq \beta \leq x_{(N)}} \prod_{i=1}^{N} [k^2 + (x_i - \beta)^2] \\
&= \arg \min_{x_{(1)} \leq \beta \leq x_{(N)}} k^{2N} + k^{2N-2} \sum_{i=1}^{N} (x_i - \beta)^2 + O(k^{2N-4}) \\
&= \arg \min_{x_{(1)} \leq \beta \leq x_{(N)}} \sum_{i=1}^{N} (x_i - \beta)^2 + \frac{O(k^{2N-4})}{k^{2N-2}}.
\end{aligned}$$

Letting $k \to \infty$, the term $\frac{O(k^{2N-4})}{k^{2N-2}}$ becomes negligible, and $\hat{\beta}_k \to \arg\min_\beta \sum_{i=1}^{N} (x_i - \beta)^2 = \sum_{i=1}^{N} x_i / N$ □

REFERENCES

[1] S. Ambike and D. Hatzinakos. A new filter for highly impulsive α-stable noise. In *Proc. of the 1995 IEEE Int. Workshop on Nonlinear Signal and Image Processing*, Halkidiki, Greece, June 1995.

[2] F. Hampel, E. Ronchetti, P. Rousseeuw, and W. Stahel. *Robust Statistics: The Approach Based on Influence Functions.* Wiley, New York, 1981.

[3] M. Shao and C. L. Nikias. Signal processing with fractional lower order moments: stable processes and their applications. *Proceedings of the IEEE*, 81(7), July 1993.

[4] P. Tsakalides and C. L. Nikias. Maximum likelihood localization of sources in noise modeled as stable processes. *IEEE Transactions on Signal Processing*, 43(11), November 1995.

[5] L. Yin, R. Yang, M. Gabbouj, and Y. Neuvo. Weighted median filters: a tutorial. *IEEE Transactions on Circuits and Systems*, 1995. To appear.

[6] P. Zurbach, J. G. Gonzalez, and G. R. Arce. Weighted myriad filters for image processing. In *Proceedings of the 1996 IEEE International Symposium on Circuits and Systems*, Atlanta, Georgia, 1996. Submitted.