filtering process we map the set of observation samples into [0, 1] with an affinity function $\mathscr{A}^{\mu,\gamma}$, where $\mu \in (-\infty, \infty)$ and $\gamma \in [0, \infty)$ control the location and spread of the affinity function, respectively. Thus,

$$\mathscr{A}^{\mu,\gamma} : x_i \mapsto \mathscr{A}_i^{\mu,\gamma} \in [0, 1] \tag{28}$$

for $i = 1, 2, \ldots, N$. The real number $\mathscr{A}_i^{\mu,\gamma}$ is a metric of the proximity of the sample x_i to the reference point μ as measured by $\mathscr{A}^{\gamma}$. If μ corresponds to one of the order statistics, $\mathscr{A}_i^{\mu,\gamma}$ is equal to the fuzzy time–rank relation between x_i and μ, as introduced in Ref. 15.

While many forms of affinity functions can be adopted, we impose the following restrictions:

1. The affinity function is unimodal with mode μ and such that $\mathscr{A}^{\mu,\gamma} : \mu \mapsto 1$.
2. The affinity function is a nondecreasing function of γ, i.e., $\mathscr{A}^{\mu,\gamma_1} : x \geq \mathscr{A}^{\mu,\gamma_2} : x$ for $\gamma_1 \geq \gamma_2$.
3. The affinity function reduces to a delta function at the mode, and is uniform for all inputs at the limits of γ, i.e.,

$$\lim_{\gamma \to 0} \mathscr{A}^{\mu,\gamma} : x_i \mapsto \delta(x_i - \mu) \quad \text{and} \quad \lim_{\gamma \to \infty} \mathscr{A}^{\mu,\gamma} : x_i \mapsto 1 \tag{29}$$

These restrictions have the following intuitive interpretations. The closer a sample is to the reference location μ, the higher is its affinity and hence its degree of reliability. The scale on which the transition from reliable to unreliable occurs is controlled by γ. Here we concentrate on a symmetric Gaussian affinity function,

$$\mathscr{A}^{\mu,\gamma} : x \mapsto e^{-(x-\mu)^2/\gamma} \tag{30}$$

The sensitivity of the Gaussian affinity function, as a function of γ, is shown in Fig. 12.

Median Affine Filters

Median affine filters use a Gaussian affinity function centered on the median observation sample to gage the reliability of observation samples and modify their influence on the estimate. Thus, let $\mu = \text{MED}[x_1, x_2, \ldots, x_N]$. Then the (normalized) median affine filter is defined as

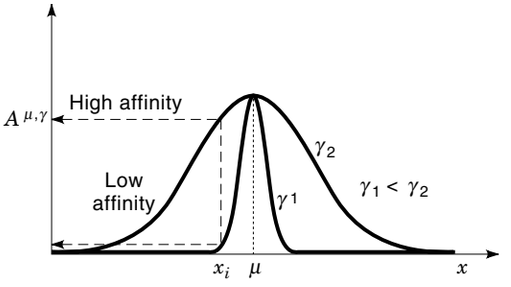

Figure 12. The affinity function $\mathscr{A}^{\mu,\gamma}$ assigns a low or high affinity to the sample x_i depending on the location and dispersion parameters μ and γ.

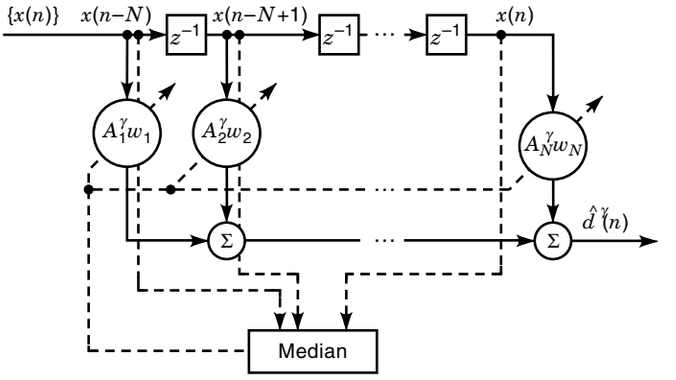

Figure 13. Structure of the median affine filter.

$$\hat{d}^{\gamma} = \frac{\sum_{i=1}^{N} w_i \mathscr{A}_i^{\mu,\gamma} x_i}{\sum_{i=1}^{N} |w_i| \mathscr{A}_i^{\mu,\gamma}} \tag{31}$$

where the w_i's are the filter weights and $\mathscr{A}_i^{\mu,\gamma}$ is the affinity of the ith observation with respect to the median observation sample. When the context is clear, we shall refer to $\mathscr{A}_i^{\mu,\gamma}$ as simply $\mathscr{A}_i^{\gamma}$. Figure 13 shows a schematic diagram of the (unnormalized) median affine filter.

The filter structure in Eq. (31) weights each observation twice: first according to its reliability, and second according to its natural order. Median affine estimates are therefore based on observations that are both reliable and favorable due to their natural order. Observations that fail to meet either or both criteria have only a limited influence on the estimate.

The flexibility provided by the tunable affinity function translates to the filter characteristics of the estimator. By varying the dispersion parameter certain properties of the median affine filter can be stressed: while a large value of γ emphasizes the linear properties of the filter, a small value puts more weight on its OS properties. Of special interest are the limiting cases. For $\gamma \to \infty$, the affinity function is constant on its entire domain. The estimator, therefore, weights all observations merely according to their natural order, i.e.,

$$\lim_{\gamma \to \infty} \hat{d}^{\gamma} = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i} \tag{32}$$

and the median affine estimator reduces to a normalized *linear filter*. For $\gamma \to 0$ the affinity function shrinks to a δ impulse at μ. Thus, the constant weights w_i are disregarded and the estimate is equal to the sample median,

$$\lim_{\gamma \to 0} \hat{d}^{\gamma} = \text{MED}[x_1, x_2, \ldots, x_N] \tag{33}$$

Design and Optimization

The median affine filter can be designed adaptively as well as through a process called medianization of linear FIR filters. Medianization, although suboptimal, is a very simple and intuitive design procedure that is derived from the fact that the median affine filter behaves like a linear filter for $\gamma \to \infty$. Setting γ to a large initial value, we can take advantage of

the multitude of linear filter design methods to find the coefficients w_i of the median affine filter. Holding the w_i's constant, the filter performance can, in general, be improved by gradually reducing the value of γ until a desired level of robustness is achieved. Since this process strengthens the medianlike properties of a linear filter, it is referred to as *medianization* of a linear FIR filter.

The median affine filter can also be optimized adaptively through an LMS-type algorithm. The development is similar to that for FIR linear filters and the WM optimization presented in the previous section. In this case, however, a dispersion parameter γ must be optimized in addition to the filter weights. Consider first the optimization of γ. Utilizing a gradient-based method, it can be shown that the optimization update for γ reduces to

$$\gamma(n+1) = \gamma(n) + \nu_\gamma (d - \hat{d})(n) \times \left(\sum_{i=1}^{N} (x_i - \hat{d}) w_i g\big((x_i - x_{(\text{med})})^2\big) \right)(n) \tag{34}$$

where ν_γ is the step size and $g(y) = (y/\gamma^2)\exp(-y/\gamma)$. In many applications, it is desirable to optimize the filter coefficients w_i as well. A similar development leads to the following algorithm for the adaptive optimization of the w_i's:

$$w_i(n+1) = w_i(n) + \nu_w (d - \hat{d})(n) \left(\mathscr{A}_i^\gamma \sum_{k=1}^{N} w_k \mathscr{A}_k^\gamma (x_i - x_k) \right)(n) \tag{35}$$

Applications

The median affine filter is a very flexible filtering framework and can thus be applied to a wide range of applications. As an illustrative example, the problem of processing inverse synthetic aperture radar (ISAR) for feature extraction and noise smoothing is presented. ISAR has attracted increasing interest in the context of target classification due to the high resolution that results from the mapping of the reflectivity density function of the target onto the range–Doppler plane (16). Difficulties in target identification arise from the fact that radar backscatters from the target are typically embedded in heavy clutter noise.

Affinity measures can be incorporated into the filtering process in a way that is particularly useful in the extraction of ISAR features and smoothing of noise (14). To illustrate the filter performance, a 128×128, 8 bit/pixel intensity image of a B-727 has been processed. Figure 14 shows the L–l, WOS, and affine filter outputs and errors. Note that the WOS filter eliminates the noise well, but blurs plane details. In contrast the L–l filter preserves the plane much better but is not very effective in removing the clutter noise. The affine filter removes the background noise to a large extent while preserving the plane in all its details.

WEIGHTED MYRIAD FILTERS

Alpha-Stable Processes

Although the generalized Gaussian model provides a flexible framework for impulsive processes, it lacks the theoretical foundation that can explain how signals with generalized Gaussian distributions may arise in practice. It turns out that the so-called α-stable distributions do have this theoretical foundation, as they satisfy an important generalization of the central limit theorem. As mentioned in the introduction, a wide variety of impulsive processes found in signal-processing applications arise as the superposition of many small independent effects. While Gaussian models are clearly inappropriate, the α-stable distributions have been proven to accurately model impulsive-type processes (5,6). Such models are appealing in that the generalization of the central limit theorem explains the apparent contradictions of its "ordinary" version, as well as the presence of non-Gaussian, infinite-variance processes.

A random variable that can be the limit of a normalized superposition according to the generalized central limit theorem is usually called *α-stable.* On a first-order analysis, symmetric α-stable processes are characterized by their distribution having a characteristic function

$$\phi(\omega) = e^{-\gamma|\omega|^\alpha} \tag{36}$$

The parameter γ, usually called the *dispersion,* is a positive constant related to the scale of the distribution. ($\gamma^{1/\alpha}$ is a scale parameter of the distribution.) The parameter α is referred to as the characteristic exponent. In order for Eq. (36) to define a characteristic function, the values of α must be restricted to the interval (0, 2]. Conceptually speaking, α determines the impulsiveness or tail heaviness of the distribution (smaller values of α indicate increased levels of impulsiveness). The limit case, $\alpha = 2$, corresponds to the zero-mean Gaussian distribution with variance 2γ. All other values of α correspond to heavy-tailed distributions.

The case $\alpha = 1$ corresponds to the zero-centered Cauchy distribution, which has density

$$f(x) = \frac{\gamma}{\pi} \frac{1}{\gamma^2 + x^2} \tag{37}$$

When $\alpha \neq 1, 2$, no closed expressions are known for the density functions, making it necessary to resort to series expansions or integral transforms to describe them. A commonly used characterization of the general symmetric α-stable density with unitary dispersion is given by (6)

$$f_\alpha(x) = \begin{cases} \dfrac{1}{\pi} \displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \Gamma(k\alpha + 1) \sin\left(\frac{\pi k \alpha}{2}\right) |x|^{-k\alpha - 1} & \\ \qquad\qquad \text{for} \quad 0 < \alpha < 1, \quad x \neq 0 & \\ \dfrac{1}{\pi(x^2 + 1)} \qquad \text{for} \quad \alpha = 1 & \\ \dfrac{1}{\pi\alpha} \displaystyle\sum_{k=0}^{\infty} \frac{(-1)^k +}{(2k)!} \Gamma\left(\frac{2k+1}{\alpha}\right) x^{2k} \quad \text{for} \quad 1 < \alpha < 2 & \\ \dfrac{1}{2\sqrt{\pi}} \exp\left(-\dfrac{x^2}{4}\right) \qquad \text{for} \quad \alpha = 2 & \end{cases} \tag{38}$$

Figure 15 shows plots of normalized unitary-dispersion α-stable densities. Note that lower values of α correspond to densities with heavier tails. Symmetric α-stable densities maintain many of the features of the Gaussian density. They are

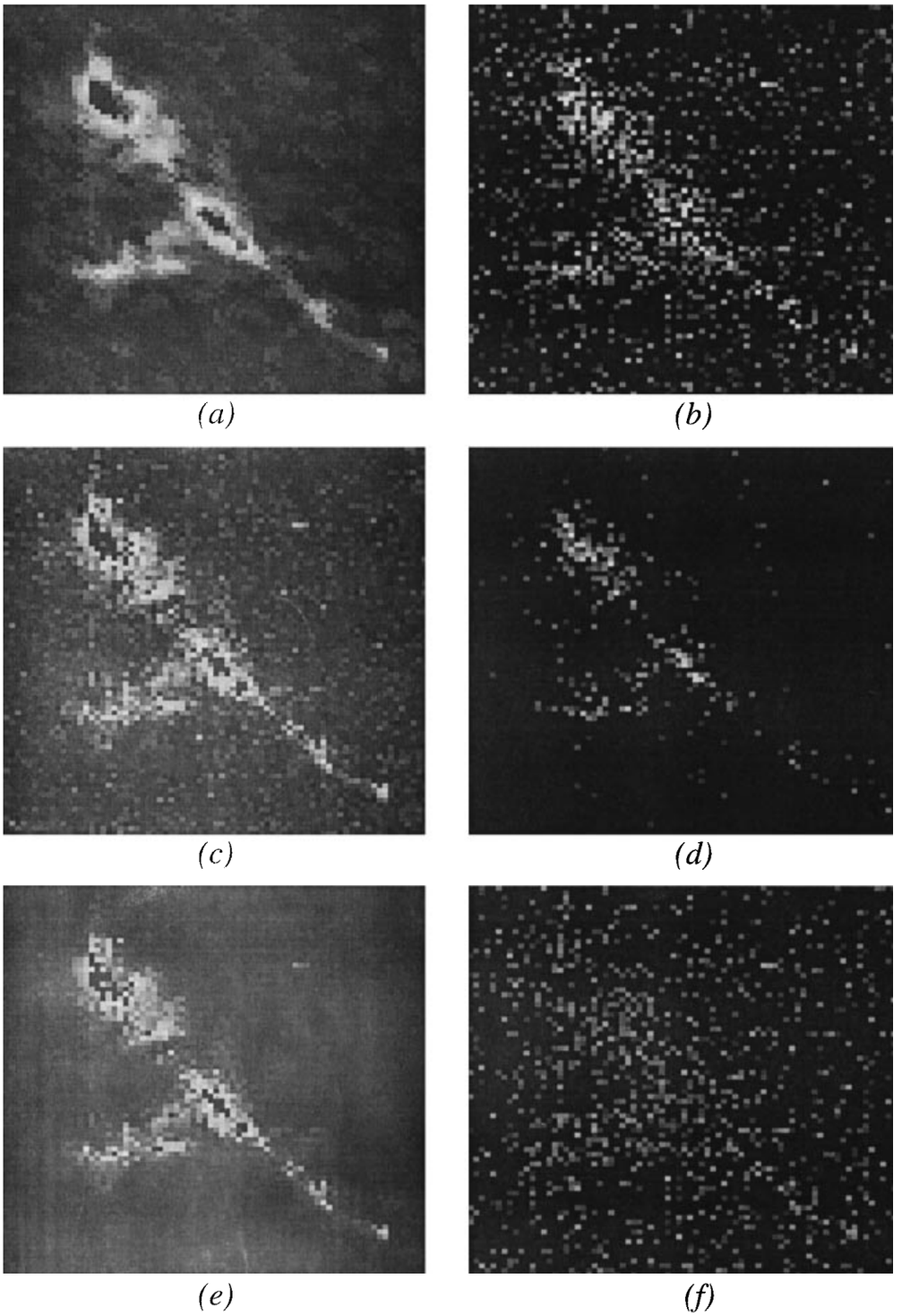

Figure 14. Feature enhancing: (a) WOS filter, (b) absolute difference between original and WOS filter, (c) L–l filter, (d) absolute difference between original and L–l filter, (e) affine filter, and (f) absolute difference between original and affine filter.

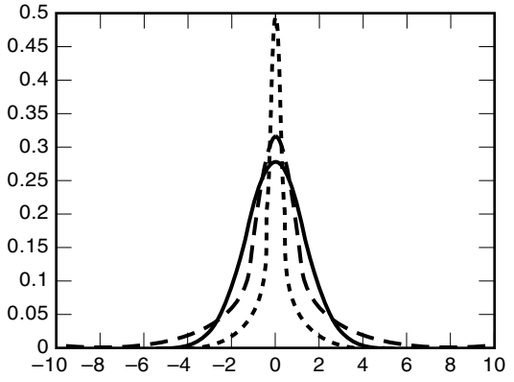

Figure 15. Symmetric α-stable density functions for different values of the tail constant α. Solid: $\alpha = 2$ or Gaussian case; dashed: $\alpha = 1$ or Cauchy case; dotted: $\alpha = 0.3$. According to the meaning of the tail constant, smaller values of α correspond to heavier-tailed density functions.

smooth, unimodal, symmetric with respect to the mode, and bell-shaped.

The Myriad and the Sample Myriad

Having the theoretical framework of α-stable processes, it is thus logical to develop new estimation algorithms for impulsive environments, which can overcome many of the limitations of WM and linear filters. To this end, we consider the maximum likelihood 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 $x_1, x_2, \ldots, x_N$ obeying the Cauchy distribution with scaling factor k,

$$f_\beta(x) = \frac{k}{\pi}\frac{1}{k^2 + (x-\beta)^2} \tag{39}$$

the location parameter β is to be estimated from the data samples as the values $\hat{\beta}_k$ that maximizes the likelihood function

$$\begin{aligned}\hat{\beta}_k &= \arg\max_\beta \prod_{i=1}^{N} f_\beta(x_i) \\ &= \arg\max_\beta \left(\frac{k}{\pi}\right)^N \prod_{i=1}^{N} \frac{1}{k^2 + (x_i-\beta)^2}\end{aligned} \tag{40}$$

This is equivalent to minimizing $G_k(\beta) = \prod_{i=1}^{N}[k^2 + (x_i - \beta)^2]$. Thus given $k > 0$, the ML location estimate is known as the sample *myriad* and is given by (17)

$$\begin{aligned}\hat{\beta}_k &= \text{MYRIAD}\{k; x_1, x_2, \ldots, x_N\} \\ &= \arg\min_\beta \prod_{i=1}^{N}[k^2 + (x_i-\beta)^2]\end{aligned} \tag{41}$$

The myriad is the ML location estimate for Cauchy random variables. However, it can be used as a robust location estimator in general, where the samples may not obey a Cauchy distribution and may not even be symmetrically distributed. Much as the robustness of the sample median is explained by the heavy tails of the Laplacian distribution, the myriad is highly robust due to the very heavy tails of the Cauchy distribution.

Note that, unlike the sample mean or median, the definition of the sample myriad involves the free parameter k. For reasons that will become apparent shortly, we will refer to k as the *linearity parameter* of the myriad. The behavior of the myriad estimator is markedly dependent on the value of its linearity parameter. It can be shown that for large values of k, the sample myriad is equivalent to the sample mean. Thus, given a set of samples $x_1, x_2, \ldots, x_N$, the sample myriad $\hat{\beta}_k$ converges to the sample average as $k \to \infty$ (17):

$$\lim_{k\to\infty} \hat{\beta}_k = \lim_{k\to\infty} \text{MYRIAD}\{k; x_1, \ldots, x_N\} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{42}$$

Plainly, an infinite value of k converts the myriad into the sample average. This behavior explains our choice "linearity" for the name of the parameter: the larger the value of k, the closer the behavior of the myriad to a linear estimator. From our experience, we have found that values of k on the order of the data range,

$$k \sim x_{(N)} - x_{(1)} \tag{43}$$

often drives the myriad to an acceptable approximation of the sample average. As the myriad moves away from the linear region (large values of k) to lower linearity values, the estimator becomes more resistant to the presence of impulsive noise. In the limit when k tends to zero, the myriad acquires its maximum resistance to impulsive noise.

It is important to note that the availability of k as a tunable parameter allows a myriad estimator to acquire some "intelligence," in the sense that the degree of linear or robust behavior can be inferred from the data by estimating an adequate value for k. Figure 16(a) depicts the sample myriad for the data set {0, 1, 3, 6, 7, 8, 9} as k is varied from 0 to ∞. 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 sample myriad favors the value $\beta = 7$, which indicates the location of the cluster formed by the samples {6, 7, 8, 9}. This is the typical behavior of the sample 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 estimator. The dotted line shows how the sample myriad is affected if an additional observation of value 100 is included. 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 an outlier and does not influence the value of the sample myriad. More interestingly, if the additional observations are taken to be {800, −500, 700}, the sample myriad is practically unchanged for moderate values of k (dashed curve).

Notably, the sample myriad presents important optimality properties compelling the use of myriad-based methods in the α-stable framework (17). First, the optimality of the sample myriad in the Cauchy distribution follows from its definition. Secondly, since the sample mean is the optimal location estimator at the Gaussian model, by assigning large values to the linearity parameter, the linear property guarantees the optimality of the sample myriad in the Gaussian distribution ($\alpha = 2$). Finally, it has been shown that the sample myriad with $k = 0$ is the optimal location estimator in the ML sense for stable processes with $\alpha \to 0$.

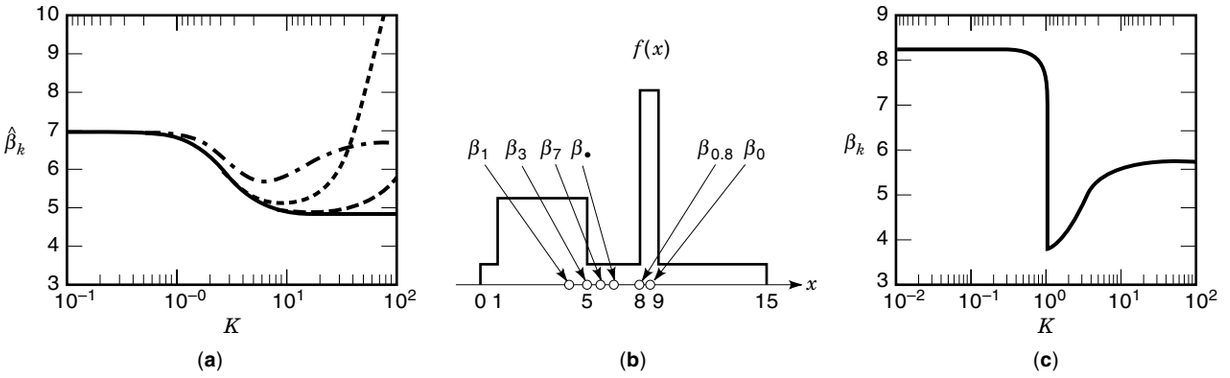

Figure 16. (a) Sample myriad as a function of k for the following data sets: solid: original data set {0, 1, 3, 6, 7, 8, 9}; dash–dot: original set plus an additional sample at 20; dotted: additional observation at 100; dashed: additional observations at 800, −500, and 700. (b) A probability density function, and (c) the myriad as a function of k for this pdf.

Much as the sample mean and sample median are the estimates of the mean and median parameters, 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 that minimizes the expectation $E\{\log[k^2 + (x - \beta_k)^2]\}$, where $k \in [0, \infty]$ is a tunable parameter. For $k = 0$ the myriad of a distribution function takes on the value that minimizes $E\{\log|x - \beta_0|\}$. It can be shown that the myriad is always at the center of symmetry whenever the underlying distribution is symmetric. Thus, for any k, β_k is an adequate indicator of location. For nonsymmetric distributions, the value of the myriad depends on k, as illustrated next. Figure 16(b) and (c) depict the mean and myriad (for $k = 0$, 0.8, 1, 3, 7, ∞) of a bimodal distribution. For $k = 0$, the myriad cautiously localizes the distribution at 8.5, which is 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 because 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.

Myriad estimation, defined in Eq. (41), can be interpreted in a more intuitive manner. As depicted in Fig. 17(a), it can be shown that the sample myriad, $\hat{\beta}_k$, is the value that minimizes the product of distances from point A to the sample points $x_1, x_2, \ldots, x_6$. Any other value, such as $x = \beta'$, produces a larger product of distances. As k is reduced, the myriad searches clusters as shown in Fig. 17(b). If k is made large, all distances become close and it can be shown that the myriad tends to the sample mean. The geometrical interpretation provides an approach to define the myriad in a complex or multidimensional space (17). Figure 17(c)) illustrates an approach to define the vector myriad as the point in the two-dimensional space that minimizes the product of distances from each sample point to the point A whose height is determined by the parameter k.

Weighted Myriad Filters

Much as linear FIR and weighted median filters are important extensions of the sample mean and median, weighted myriad filters are important generalizations of the sample myriad. As with FIR filters, N weights are defined, one for each input sample used in the myriad estimate. Filters with nonnegative weights have inherent limitations and are referred to as *smoothers*. Filters whose weights are not constrained to be nonnegative are, in general, more flexible and powerful than their constrained counterparts. The definitions of *weighted myriad smoothers* and *weighted myriad filters* are given next.

To define the weighted myriad smoother, we invoke the same ML estimation principles used in defining the weighted median filter. Recall that the mean emerges as the ML estimate of equally likely Gaussian observations ($w_1 = w_2 = \cdots = w_N = 1$). In the same fashion, the weighted mean can be seen as the value $\hat{\beta}_w$ that maximizes the likelihood function

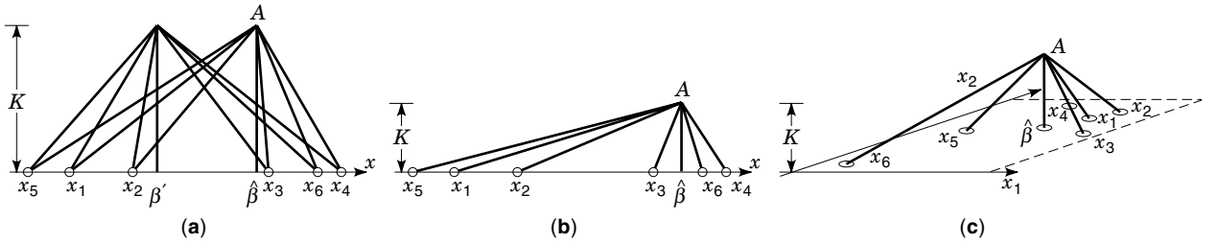

Figure 17. (a) The sample myriad $\hat{\beta}$ minimizes the product of distances from point A to all samples. Any other value, such as $x = \beta'$, produces a higher product of distances. (b) The myriad as k is reduced. (c) Minimum product of distances for a two-dimensional vector sample set.

when the density functions are defined as

$$f_i(x_i; \beta, w_i) = \frac{|w_i|}{\sqrt{2\pi}\sigma} \exp\left(\frac{-w_i^2(x_i - \beta)^2}{2\sigma^2}\right) \tag{44}$$

Clearly, $f_i(x_i; \beta, w_i)$ is a scaled version of $f_i(x_i; \beta)$, where the spread of the distribution is inversely proportional to w_i. Thus, an observation assigned to a large weight can be related to a highly localized density function. 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, which implies a very poor chance of this observation to be close to the center of the distribution. If the Cauchy density model is used, maximizing the weighted likelihood function is equivalent to minimizing $\prod_{i=1}^{N}[k^2 + w_i(x_i - \beta)^2]$. The *weighted myriad smoother* is then defined as (17)

$$\hat{\beta}_{k,\boldsymbol{w}} = \text{MYRIAD}\{k; w_1^2 \circ x_1, w_2^2 \circ x_2, \ldots, w_N^2 \circ x_N\} \tag{45}$$

$$= \arg\min_{\beta} \prod_{i=1}^{N}[k^2 + w_i(x_i - \beta)^2] \tag{46}$$

where $w_i^2 \circ x_i$ represents the weighting operation in Eq. (46). Since $\hat{\beta}_{k,c\boldsymbol{w}} = \hat{\beta}_{k/c,\boldsymbol{w}}$, it is clear that finding the optimal myriad smoother weights will implicitly find the best k. As with the sample myriad, it can be shown that as $k \to \infty$, the weighted myriad smoother tends to the weighted mean smoother.

Weighted myriad filters admitting real-valued weights are more flexible and powerful filter structures. These are analogous to linear FIR filters, whereas weighted myriad smoothers (and weighted median filters) are analogous to constrained linear FIR smoothers. Many linear FIR filters used in practice, such as band-pass and high-pass filters, do in fact require negative weights. In communications technology, modulated signals are more effectively processed by filters rather than smoothers. It is thus important to determine the structure of the general myriad filter—one that admits real-valued weights. Following the approach in Ref. 12, it is shown in Ref. 18 that the weighted myriad filter can be defined as

$$\hat{\beta}_{k,\boldsymbol{w}} = \text{MYRIAD}\{k; |w_1| \circ \text{sgn}(w_1)\, x_1, |w_2| \circ \text{sgn}(w_2)\, x_2, \ldots, |w_N| \circ \text{sgn}(w_N)\, x_N\} \tag{47}$$

$$= \arg\min_{\beta} \prod_{i=1}^{N}\{k^2 + |w_i|[\text{sgn}(w_i)\, x_i - \beta]^2\} \tag{48}$$

where $|w_i| \circ \text{sgn}(w_i)\ x_i$ again represents the weighting operation in Eq. (48).

In general, the computation of the weighted myriad does not admit closed-form solution, making it necessary to resort to iterative minimization procedures. In Ref. 19 it is shown that the myriad can be seen as a fixed point of

$$T(\beta) = \frac{\sum_i h_i(\beta) x_i}{\sum_i h_i(\beta)} \tag{49}$$

where

$$h_i(\beta) = \frac{w_i}{k^2 + w_i(x_i - \beta)^2} \tag{50}$$

which can be searched for by taking an initial value β_0 and then using the recursion

$$\beta_{n+1} = T(\beta_n) \tag{51}$$

Although it could happen that such recursion converges to a fixed point of T different from the weighted myriad, using one (or several) good initial value(s) can give very satisfactory results. This algorithm is very reliable and flexible, providing a convenient platform to manage the tradeoff between speed and reliability. The interested reader is referred to Ref. 19 for further information on this algorithm.

Design and Optimization

It was described that for very large k the weighted myriad filter reduces to a linear FIR filter. Likewise it was noted that robustness is achieved by decreasing the value of k. Thus, a very simple method to design weighted myriad filters is to the use the weights of a linear filter ($k = \infty$), designed for Gaussian or noiseless environments, and to subsequently reduce the value of k, attaining the level of robustness desired, before the weighted myriad is computed. We refer to this method as "myriadization," in contrast to the well-known "linearization" approaches used in engineering (17). The following example clearly illustrates the effectiveness of myriadizing a linear low-pass FIR filter.

Consider a phase-locked-loop (PLL) synchronization problem (17). It is well known that automobile FM radios do not use PLLs because a linear PLL cannot operate with impulsive noise generated by ignitions and other interference signals. Our goal is to design myriad PLLs that can withstand severe interference and still provide satisfactory performance. The commercial applications of such robust PLLs are many, including commercial FM and Loran-C receivers (20). The myriadization concept can be easily used in the design of a first-order PLL. Simulations were run in which the PLL had to track the carrier phase in additive Gaussian noise. The signal-to-noise ratio was set at 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. Figure 18(a) shows a typical phase error plot of a linear PLL in which random noise bursts are present. During these short noisy intervals (from 4 to 10 sampling times), the signal-to-noise ratio decreases to -10 dB.

It is evident that the system with the linear filter is very likely to lose synchronism after a noise burst. Figure 18(b) shows the phase error of the optimal median-based PLL (21) with the same noise conditions. Although the short noise bursts do not affect the estimate of the phase, the variance of the estimate is very large. Figure 18(c) 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 a burst, 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 low-amplitude Gaussian noise, the myriadized system shows a smaller steady-state variance, while maintaining the same synchronization response.

Myriadization provides a simple method to design the proposed filter class. However, significant gains can be attained

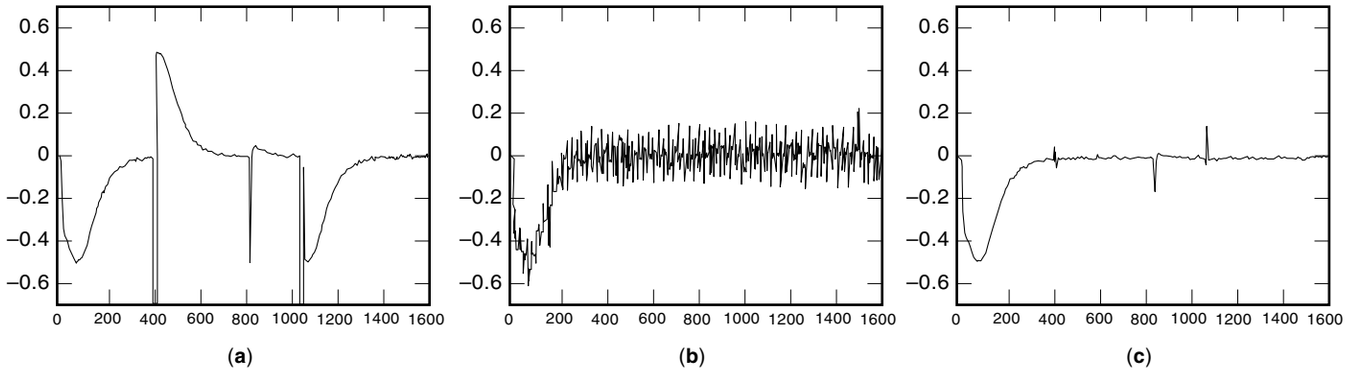

Figure 18. (a) Phase error plot for the PLL with a linear FIR filter; (b) phase error plot for the PLL with a median filter; (c) phase error after the low-pass filter has been myriadized.

if we optimize the filter coefficients. Let the input to the myriad filter be $\{x(n)\}$, which is assumed to be statistically related to some desired process $D(n)$. A window of width N slides across the input process, spanning the samples $x(n)$, $x(n-1)$, . . ., $x(n-N+1)$. The estimate of $Y(n)$ is given by the myriad filter output

$$Y(n) = \text{MYRIAD}[k; w_1 \circ x(n), w_2 \circ x(n-1), \ldots, w_N \circ x(n-N+1)] \tag{52}$$

The goal is to find the optimal set of weights $w_1, \ldots, w_N$ such that the mean squared error between the desired signal and the filter's output is minimized. This problem is analogous to finding the well-known normal equation of linear filters. To this end, conditions for the optimal weighted myriad filter have been derived (18).

A simpler yet effective method to optimally design the filter weights has been derived through filter algorithms whose complexities are comparable to Widrow's LMS algorithm. Given an N-long input vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_N]^T$, a weight vector $\boldsymbol{w} = [w_1, w_2, \ldots, w_N]^T$ of real-valued weights, and the *linearity parameter* $k > 0$, we first define the set *normalized weights* h $\boldsymbol{w}/k^2$. The adaptive LMS-type adaptive algorithm updates the weights of the myriad filter in order to converge to the optimal filter weights under the MSE criterion (18):

$$h_i(n+1) = h_i(n) + \mu e(n) \frac{v_i(n)}{[1 + |h_i(n)| v_i^2(n)]^2} \tag{53}$$

for $i = 1, 2, \ldots, N$, where $v_i = \text{sgn}(h_i) \cdot Y - x_i$, and $\text{sgn}(\cdot)$ is the sign function.

RESEARCH TOPICS IN NONLINEAR FILTERING

The need for nonlinear processing methods arises naturally in many applications. In this article we have focused on the classes of filters that arise naturally from two heavy-tailed distribution families. Specifically, we have focused on the generalized Gaussian distribution and α-stable distributions. Both distribution families include the standard Gaussian distribution as a special case. Importantly, both distributions more accurately model the impulsive nature of signals often observed in practice. The maximum likelihood estimation criterion can be applied to each of these distributions and was shown to result in the classes of weighted median and weighted myriad filters. As a direct result of the heavy-tailed nature of the underlying distribution models, both filter classes are robust and perform significantly better than linear filters in many applications.

The space constraints of this article allow for only a brief overview of the covered topics. To probe further, the interested reader is referred to the cited articles, as well as numerous other works in this area. Additionally, there are many other areas of research in nonlinear methods that are being actively investigated. Research areas of importance include: (1) OS-based signal processing, (2) mathematical morphology, (3) higher-order statistics, (4) radial basis functions, and (5) emerging nonlinear methods.

OS methods are a large class of nonlinear systems based on robust statistics. Included in this class are the discussed median and weighted median filters. Other OS-based methods include stack filters (21–26), L–l filters (27–30), permutation filters (31–36), and RCRS filters (37), as well as numerous hybrid and generalization methods. Each of these methods exploits, in some fashion, the rank order, or joint temporal rank order, of the observation samples. Recent efforts have included fuzzy time–rank relations (14,15). Each of these methods has proved advantageous over linear filters in the processing of nonstationary processes with heavy-tailed distributions.

Related to OS methods is the class of morphological processes (38–40), which stems historically from set–lattice theory and stochastic–integral geometry. Many signal analysis and computer vision tasks such as feature extraction, motion detection, segmentation, and object recognition often need a multiscale formulation, where features/objects are more easily detected at coarse scales rather then at their original resolution. Although early approaches in computer vision used linear low-pass filters (e.g. Gaussian convolutions) for multiscale analysis, the linear scale space suffers from its shifting and blurring important features across scales. In contrast, morphological and OS smoothing filters have recently been used to create a nonlinear scale space for multiscale image analysis that has as rich a theory as the Gaussian scale-space methods and can exactly preserve vertical edges and the outline and location of object shapes up to the maximum scale at which they exist.

Higher-order statistics (HOS) is another important approach to the nonlinear processing. HOS can offer significant

advantages over conventional second-order statistics due to the fact that most processes are non-Gaussian in nature. Thus, system identification, noise suppression, and signal synthesis are all signal-processing tasks that can gain from the use of HOS. Related to HOS systems are polynomial and Volterra filters, which use polynomial transformations of the input signal space. Thus, this class of nonlinear filters exploits the higher-order statistics of an underlying signal to provide improved estimates.

In applications where nonlinear processes must be modeled, radial basis function (RBF) methods can be employed. RBFs offer local approximations to fit local regions using a parametric function. The fit is in regions where data exist, and therefore it is immune to modeling errors that may occur in regions other than the one being approximated. In addition to the mentioned methods, there are numerous emerging nonlinear methods. While many of these methods presently lack a strong theoretical framework, their evolution will continue and new theories will arise.

In closing, we encourage the interested reader to investigate further the growing body of knowledge in nonlinear signal processing. While the field of digital signal processing has matured within the framework of linear systems, novel areas of nonlinear signal processing continue to appear. The response of the scientific community to nonlinear signal processing has been tremendous in recent years. We expect that novel research areas and applications will continue to appear within that broad domain.

BIBLIOGRAPHY

1. H. M. Hall, A new model for "impulsive" phenomena: Application to atmospheric-noise communication channels, Technical Reports 3412-8 and 7050-7, Stanford Electronics Laboratories, Stanford University, Stanford, CA, August 1966.
2. J. Ilow and D. Hatzinakos, Analytic alpha-stable noise modeling in a Poisson field of interferers or scatterers, *IEEE Trans. Signal Process.,* submitted for publication.
3. B. Mandelbrot, Long-run linearity, locally Gaussian processes, H-spectra, and infinite variances, *Int. Econ. Rev.,* **10**: 82–111, 1969.
4. J. H. Miller and J. B. Thomas, Detectors for discrete-time signals in non-Gaussian noise, *IEEE Trans. Inf. Theory,* **IT-18**: 241–250, 1972.
5. C. L. Nikias and M. Shao, *Signal Processing with Alpha-Stable Distributions and Applications,* New York: Wiley-Interscience, 1995.
6. V. Zolotarev, *One-Dimensional Stable Distributions,* Providence, RI: Amer. Math. Soc., 1986.
7. N. C. Gallagher, Jr. and G. L. Wise, A theoretical analysis of the properties of median filters, *IEEE Trans. Acoust. Speech Signal Process.,* **29**: 1136, December 1981.
8. F. Y. Edgeworth, A new method of reducing observations relating to several quantities, *Phil. Mag. (Fifth Series),* **24**: 1887.
9. O. Yli-Harja, J. Astola, and Y. Neuvo, Analysis of the properties of median and weighted median filters using threshold logic and stack filter representation, *IEEE Trans. Acoust. Speech Signal Process.,* **39**: 395–410, 1991.
10. R. C. Hardie and C. G. Boncelet, Jr., LUM filters: A class of rank order based filters for smoothing and sharpening, *IEEE Trans. Signal Process.,* **41**: 1061–1076, 1993.
11. S.-J. Ko and Y. H. Lee, Center weighted median filters and their applications to image enhancement, *IEEE Trans. Circuits Syst.,* **38**: 984–993, 1991.
12. G. R. Arce, A generalized weighted median filter structure admitting real-valued weights, *IEEE Trans. Signal Process.,* submitted for publication.
13. L. Yin and Y. Neuvo, Fast adaptation and performance characteristics of fir-wos hybrid filters, *IEEE Trans. Signal Process.,* **42**: 1610–1628, 1994.
14. A. Flaig, G. R. Arce, and K. E. Barner, Affine order statistic filters: A data-adaptive filtering framework for nonstationary signals, *IEEE Trans. Signal Process.,* to be published.
15. K. E. Barner, A. Flaig, and G. R. Arce, Fuzzy time–rank relations and order statistics, *IEEE Signal Process. Lett.,* submitted for publication.
16. A. Zyweck and R .E. Bogner, High-resolution radar imagery of mirage III aircraft, *IEEE Trans. Antennas Propag.,* **42**: 1356–1360, 1994.
17. J. G. Gonzalez and G. R. Arce, Weighted myriad filters: A robust filtering framework derived from alpha-stable distributions, *IEEE Trans. Signal Process.,* submitted for publication.
18. S. Kalluri and G .R. Arce, Robust frequency-selective filtering using generalized weighted myriad filters admitting real-valued weights, *IEEE Trans. Signal Process.,* submitted for publication.
19. S. Kalluri and G. R. Arce, Fast weighted myriad computation using fixed point searches, *IEEE Trans. Signal Process.,* submitted for publication.
20. W. B. McCain and C. D. McGillem, Performance improvement of dpll's in non-gaussian noise using robust estimators, *IEEE Trans. Commun.,* **35**: 1207–1216, 1987.
21. R. W. Hawley, N. C. Gallagher, and M. P. Fitz, Stack filter phase lock loops, *IEEE Trans. Signal Process.,* **38**: 317–329, 1994.
22. E. J. Coyle and J.-H. Lin, Stack filters and the mean absolute error criterion, *IEEE Trans. Acoust. Speech Signal Process.,* **36**: 1244–1254, 1988.
23. E. J. Coyle, J.-H. Lin, and M. Gabbouj, Optimal stack filtering and the estimation and structural approaches to image processing, *IEEE Trans. Acoust. Speech Signal Process.,* **37**: 2037–2066, 1989.
24. P. Wendt, E. J. Coyle, and N. C. Gallagher, Jr., Stack filters, *IEEE Trans. Acoust. Speech Signal Process.,* **34**: August 1986.
25. I. Tabus, D. Petrescu, and M. Gabbouj, A training framework for stack and Boolean filtering—fast optimal procedures and robustness case study, *IEEE Trans. Image Process.,* **5**: 809–826, 1996.
26. P.-T. Yu and R.-C. Chen, Fuzzy stack filters—their definitions, fundamental properties, and application in image processing, *IEEE Trans. Image Process.,* **5**: 838–854, 1996.
27. F. Palmieri and C. G. Boncelet, Jr., *Ll*-filters—a new class of order statistic filters, *IEEE Trans. Acoust. Speech Signal Process.,* **37**: 691–701, 1989.
28. F. Palmieri and C. G. Boncelet, Jr., Frequency analysis and synthesis of a class of nonlinear filters, *IEEE Trans. Acoust. Speech Signal Process.,* **38**: 1363–1372, 1990.
29. P. Ghandi and S. A. Kassam, Design and performance of combination filters, *IEEE Trans. Signal Process.,* **39**: 1524–1540, 1991.
30. K. E. Barner, Colored *L–l* filters with applications to speech pitch detection, in *Proc. 1997 IEEE EURASIP Workshop Nonlinear Signal Image Process.,* Mackinac Island, MI, September 1997.
31. G. R. Arce, Y. T. Kim, and K. E. Barner, Order-statistic filtering and smoothing of time series: Part 1, in P. R. Krishnaiah (ed.), *Handbook of Statistics—vol. 16: Order Statistics and Their Applications,* Amsterdam: North-Holland, 1998.

32. K. E. Barner and G. R. Arce, Order-statistic filtering and smoothing of time series: Part 2, in P. R. Krishnaiah (ed.), *Handbook of Statistics—vol. 16: Order Statistics and Their Applications,* Amsterdam: North-Holland, 1998.
33. G. R. Arce, T. A. Hall, and K. E. Barner, Permutation weighted order statistic filter lattices, *IEEE Trans. Image Process.,* **4**: 1070–1083, 1995.
34. K. E. Barner and G. R. Arce, Permutation filters: A class of nonlinear filters based on set permutations, *IEEE Trans. Signal Proces.,* **42**: 782–798, 1994.
35. K. E. Barner and G. R. Arce, Design of permutation order statistic filters through group colorings, *IEEE Trans. Circuits Syst.-II,* **44**: 531–548, 1997.
36. R. C. Hardie and K. E. Barner, Extended permutation filters and their application to edge enhancement, *IEEE Trans. Image Process.,* **5**: 855–867, 1996.
37. R. C. Hardie and K. E. Barner, Rank conditioned rank selection filters for signal restoration, *IEEE Trans. Image Process.,* **3**: 192–206, 1994.
38. P. A. Maragos and R. W. Schafer, Morphological filters—part I: Their set theoretic analysis and relations to linear shift invariant filters; part II: Their relations to median, order-statistic, and stack filters, *IEEE Trans. Acoust. Speech Signal Process.,* **35**: 1153–1169, 1170–1184, 1987.
39. P. Salembier et al., Morphological operators for image and video compression, *IEEE Trans. Image Process.,* **5**: 881–898, 1996.
40. P. Maragos, Differential morphology in image processing, *IEEE Trans. Image Process.,* **5**: 922–937, 1996.

GONZALO R. ARCE
KENNETH E. BARNER
University of Delaware

NONLINEAR SYSTEMS. See STATISTICAL SIGNAL PROCESSING, HIGHER ORDER TOOLS.

NONLINEAR SYSTEMS REPRESENTATION. See VOLTERRA SERIES.

NONLINEAR SYSTEMS, STABILITY. See STABILITY THEORY, NONLINEAR.

NONSINUSOIDAL OSCILLATORS. See MULTIVIBRATORS.