ROBUST MERIDIAN FILTERING

Tuncer C. Aysal and Kenneth E. Barner

Signal Processing and Communications Group
Electrical and Computer Engineering, University of Delaware

ABSTRACT

The linear, median, myriad filtering structures are statistically related to the maximum likelihood (ML) estimates of location under Gaussian, Laplacian, and Cauchy statistics, respectively. In this paper, we propose a filtering structure based on the ML estimate of the constructed and so–called meridian statistics. Analysis and simulations presented here indicate that the proposed filtering structure exhibits characteristics more robust than that of mean, median and myriad filtering structures.

***Index Terms*—** nonlinear estimation, maximum–likelihood estimation, filtering

1. INTRODUCTION

Given a set of observations (input samples), $\{x_i|_{i=1}^N\}$, an M–estimate of their common location, β, is given by [1]

$$\hat{\beta} = \arg\min_{\beta} \left[\sum_{i=1}^{N} \rho(x_i - \beta) \right] \tag{1}$$

where $\rho(\cdot)$ is the *cost function* of the M–estimators. Maximum likelihood location estimates form a special case of M–estimators, with the observations being independent and identically distributed and $\rho(u) = -\log f(u)$, where $f(u)$ is the common density function of the samples. The weighted mean (linear), weighted myriad and weighted median filter families are well–derived from Maximum Likelihood location estimator under Gaussian, Cauchy and Laplacian statistics, respectfully [2, 3]. The cost functions to minimize, in these cases, are given by $\rho(u) = u^2$, $\rho(u) = \log\{\gamma^2 + u^2\}$, where γ is the linearity parameter [3] and $\rho(u) = |u|$ for mean, myriad and median estimators, respectively.

In this paper, we focus on the well–established statistical relation between the Gaussian and Cauchy distributions, indicating that the ratio of two independent Gaussian RVs is Cauchy distributed. We note that the Cauchy distribution is a special case of the GCD family corresponding to $p = 2$, where p is the tail parameter. An analogous statistical relationship is constructed here for the Laplacian distribution, where the distribution function of the RV formed as the ratio of two independent Laplacian distributed RVs is derived. Interestingly, it is shown that the obtained statistics, referred to as the Meridian distribution, is also a member of the GCD with $p = 1$. Hence, a connection between the GGD and GCD families is formed. The maximum likelihood estimate under the new statistics is analyzed, where the cost function, in this case, is given by $\rho(u) = \log\{\delta + |u|\}$, with δ controlling the robustness of the meridian estimator. The fact that the meridian estimator is likelihood–based guarantees that the estimate is (at least asymptotically) unbiased, consistent and efficient in Meridian statistics.

2. STATISTICAL PROCESSES, MAXIMUM LIKELIHOOD ESTIMATION AND FILTERING

A broad range of statistical processes can be characterized by the generalized Gaussian probability density function (PDF). The Gaussian ($k = 2$) and Laplacian ($k = 1$) density cases, where k denotes the tail parameter, are of special interest. Consider the problem of estimating the constant amplitude signal β from the samples $x_i, x_{i+1}, \ldots, x_{i+N-1}$ of noisy observation data $\{x(i)\}$. Let $x_i = \beta + \eta_i$, where the η_i terms are independent and identically distributed zero–mean noise. The ML estimate of location under Gaussian and Laplacian statistics are the mean and median estimators [1–3]. These estimators can be viewed as mean and median filtering structures operating on window basis.

The mean and median filtering structure is extended to admit weights considering a set of N independent samples $\mathbf{x}_N = x_1, x_2, \ldots, x_N$ each obeying the Gaussian or Laplcian distribution with (possibly) *different* variances σ_i^2. The weight positivity constraining the filters to smoothers can, as in the FIR filter case, be relaxed to enable more general filtering characteristics utilizing the sign–coupling approach [3].

The generalized Cauchy distribution was proposed by Miller and Thomas in 1972 [4]. The generalized Cauchy distribution is used in several studies of impulsive radio noise [3, 4]. The generalized Cauchy PDF is given by

$$f(x) = a(\gamma^p + |x|^p)^{-2/p} \tag{2}$$

with $a = p\Gamma(2/p)\gamma/(2(\Gamma(1/p))^2)$. In this representation, γ is the scale parameter and p is the tail constant. This representation includes the standard *Cauchy* PDF as a special case

$(p = 2)$. For $p < 2$, the PDF's tail decays slower than in the Cauchy case, resulting in a heavier–tailed PDF. In the following, the well–established statistical relation between Gaussian and Cauchy distributions is discussed. Also, the recently proposed myriad filtering structure [3], which is based on the ML estimate under Cauchy statistics is discussed.

Proposition 1 *The RV formed as the ratio of two independent zero–mean* Gaussian *distributed RVs, U and V, with variances σ_U^2 and σ_V^2, respectively, is* Cauchy *distributed (generalized Cauchy distribution with $p = 2$) with $\gamma = \sigma_U/\sigma_V$.*

Remark 1 *It should be noted that the original authors derived the myriad filter starting from α–Stable distributions [5]. They noted that there is only two closed–form expression for α–Stable distributions, $\alpha = 2$ and $\alpha = 1$ corresponding to Gaussian and Cauchy distributions, respectively. This original development [3] did not mention or utilize the statistical relation between Gaussian and Cauchy distributions.*

As in the previous cases, the sample myriad is generalized to admit weights [3]. Also, the weight positivity constraint restricting the filters to be smoothers is relaxed to enable more general filtering characteristics [3]. Next, a relationship similar to that between the Gaussian and Cauchy statistics is constructed here for the Laplacian case. That is, the distribution of the RV formed as the ratio of two independent Laplacian RVs is considered.

Proposition 2 *The RV formed as the ratio of two independent zero–mean* Laplacian *distributed RVs, U and V, with scale parameters λ_U and λ_V, respectively, is a member of the GCD family, with $p = 1$ and $\gamma = \lambda_U/\lambda_V$, and is referred to as the* meridian *distribution.*

Proof 1 *Let X be the RV formed as the ratio of two RVs, U and V: $X = U/V$. The PDF of the RV X, $f_X(\cdot)$, is given by: $f_X(x) = \int_{-\infty}^{\infty} |v| f_{U,V}(xv, v)dv$, where $f_{U,V}(\cdot,\cdot)$ is the joint PDF of U and V. Considering the independent case and solving for RVs U and V with Laplacian PDFs, $f_U(x) = 1/(2\lambda_U)\exp\{-|x|/\lambda_U\}$ and $f_V(x) = 1/(2\lambda_V)\exp\{-|x|/\lambda_V\}$, performing some manipulations and setting $\delta = \lambda_U/\lambda_V$ yields:*

$$f_M(x) = \frac{\delta}{2}\frac{1}{(\delta + |x|)^2}. \tag{3}$$

A careful inspection of the observed distribution shows that $f_M(\cdot)$ belongs to the generalized Cauchy distribution family, corresponding specifically to the $p = 1$ case. We refer to $f_M(\cdot)$ as the Meridian distribution.

Remark 2 *It is surprising to note that the ratio of two generalized Gaussian distributions, with $k = 2$, yields the generalized Cauchy distribution, with $p = 2$. Also, the ratio of two generalized Gaussian distributions, with $k = 1$, yields the generalized Cauchy distribution, with $p = 1$[1].*

[1] We recently established the generalization of this statistical relation for any k

3. MERIDIAN FILTERING

In the following, location estimation from observed samples under *Meridian* statistics is considered and the filtering problem is related to ML estimation in an analogous fashion to the previous Gaussian, Laplacian and Cauchy cases.

Theorem 1 *Consider a set of N independent samples $\mathbf{x}_N$ each obeying the* Meridian *distribution with common scale parameter δ. The ML estimate of location, β, or* sample meridian*, is given by*

$$\hat{\beta} = \arg\min_{\beta}\left[\sum_{i=1}^{N}\log\{\delta + |x_i - \beta|\}\right] \triangleq \textit{meridian}\{\mathbf{x}_N; \delta\} \tag{4}$$

where δ is referred to as the medianity *parameter.*

Proof Sketch 1 *Replacing the Meridian distribution for each sample in the ML formulation, utilizing basic properties of the* $\arg\max$ *function and noting that maximizing the fraction is equivalent to minimizing the denominator, and taking the natural* $\log$ *of the above yields the final result.*

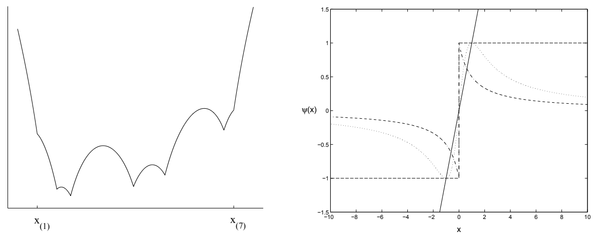
The performance of the meridian filtering is directly related to the objective function that arises naturally from the PDF. The following proposition presents several key properties of the meridian objective function. The properties described below are illustrated by Fig. 1 (a) and proved in [6], which illustrates the objective function, $Q(\beta)$, that results from a set of observed samples in the window size of $N = 7$ case.

Proposition 3 *[6] Let $\{x_{(i)}|_{i=1}^{N}\}$ denote the order statistics of the input vector $\mathbf{x}_N$, with $x_{(1)}$ the smallest and $x_{(N)}$ the largest. Also, define $Q(\beta) \triangleq \sum_{i=1}^{N}\log\{\delta + |x_i - \beta|\}$. The following statements hold:*

1. *$Q'(\beta) > 0$ for $\beta > x_{(N)}$, and $Q'(\beta) < 0$ for $\beta < x_{(1)}$.*
2. *The objective function $Q(\beta)$ is concave in $x_{(i)} < \beta < x_{(i+1)}$ for $i = 1, 2, \ldots, N-1$.*
3. *The objective function $Q(\beta)$ has a finite number of local minima [input samples].*
4. *The meridian $\hat{\beta}$ is one of the local minima of $Q(\beta)$, i.e., one of the input samples.*

The meridian estimator output is hence the input sample that yields the smallest $Q(\beta)$ function value. The selective nature of the meridian estimator, shared with the median estimator, facilitates the filter output computation which is formulated as $\hat{\beta} = \arg\min_{\beta\in\mathbf{x}_N} Q(\beta)$.

Property 1 (Median Property) *Given a set of samples $\mathbf{x}_N$, the sample meridian $\hat{\beta}$ converges to the sample median as $\delta \to \infty$. This is, $\lim_{\delta\to\infty}\hat{\beta} = \textit{median}\{\mathbf{x}_N\}$.*



(a) (b)

Fig. 1. (a) Sketch of a typical meridian objective function: Input samples are $\mathbf{x}_7 = [4.9, 0.0, 6.5, 10.0, 9.5, 1.7, 1]$ and $\delta = 1$. (b) IFs for (solid:) the mean, (dashed:) the median, (dotted:) the myriad, and (dash–dotted:) the meridian.

Proof 2 *Using the properties of the* $\arg\min$ *function, the estimator can be expressed as*

$$\hat{\beta} = \arg\min_{\beta} \left[\sum_{i=1}^{N} \log \left\{ 1 + \frac{|x_i - \beta|}{\delta} \right\}^{\delta} \right]. \tag{5}$$

Since $\lim_{\delta\to\infty} \log\left\{1 + |x_i - \beta|/\delta\right\}^{\delta} = \exp\left\{|x_i - \beta|\right\}$*, and the exponential function* $\exp\{\cdot\}$ *is monotonically increasing, it follows that* $\hat{\beta} = \arg\min_{\beta}\left[\sum_{i=1}^{N} |x_i - \beta|\right]$.

It is important to emphasize that the family of meridian estimators subsumes the sample median as a limiting case. This simple fact makes the meridian filter class inherently more efficient than (or at least equally efficient to) median filters over all noise distribution, including the Laplacian.

As the meridian moves, in function, away from the median region (large values of δ) to lower medianity values, the estimator becomes more robust to the presence of impulsive noise. In the limit, when δ tends to zero, the meridian estimator treats every observation as a possible outlier, assigning more credibility to the most repeated values in the sample set.

Property 2 (Mode Property) *[6] Given a set of samples* $\mathbf{x}_N$*, the sample meridian* $\hat{\beta}$ *converges to a mode–type estimator as* $\delta \to 0$*. This is,*

$$\lim_{\delta\to 0} \hat{\beta} = \arg\min_{x_j \in \mathcal{M}} \left[\prod_{i=1, x_i \neq x_j}^{N} |x_i - x_j| \right] \tag{6}$$

where $\mathcal{M}$ *is the set of most repeated values.*

The influence function (IF) of an estimator determines the effect of contamination on the estimator. To further characterize M–estimates, it is useful to list the desirable features of a robust IF [1]: 1) *B–robustness*. An estimator is *B–robust* if the supremum of the absolute value of the IF is finite, 2) *Rejection Point*. The *rejection point*, defined as the distance from the center of the IF to the point where the IF becomes negligible, should be finite.

The IF for the sample mean, median and myriad can be shown to be $\psi(x) = 2x$, $\psi(x) = \mathrm{sgn}(x)$ and $\psi(x) = 2x/(\gamma^2 + x^2)$, respectively. The IF of the sample meridian is discussed in the following.

Proposition 4 *The IF of the meridian estimator is given by* $\psi(x) = sgn(x)/(\delta + |x|)$.

The IFs for the sample mean, median, myriad and meridian are depicted in Fig. 1 (b). The mean is clearly not B–robust and its rejection point is infinite. On the other hand, a gross error has a limited effect on the median estimate. While the median is B–robust, its rejection point, like the mean, is not finite. Thus the median estimate is always affected by outliers. The myriad estimate is clearly B–robust and the effect of the errors decreases as the error increases. The meridian estimate is also B–robust, and in addition, the rejection point is smaller than that of myriad as it has a higher IF decay rate. This indicates that the meridian is more robust than the myriad.

In addition to desirable IF features, the meridian possesses the followings properties important in signal processing applications [6]: Outlier rejection, i.e. $\lim_{x_N\to\pm\infty} \hat{\beta}(\mathbf{x}_N) = \hat{\beta}(\mathbf{x}_{N-1})$, no overshoot/undershoot, and, shift and sign invariance.

Theorem 2 *[6] Given a set of* N *independent samples* $\mathbf{x}_N$*, each obeying the Meridian distribution with varying scale parameters* $v_i = \delta/h_i$*, the ML estimate of location, or* weighted meridian*, is given by*

$$\hat{\beta} = \arg\min_{\beta} \left[\sum_{i=1}^{N} \log\{\delta + h_i|x_i - \beta|\} \right] \triangleq \textit{meridian}\{\mathbf{h} \star \mathbf{x}_N\} \tag{7}$$

where $\star$ *denotes the weighting operation in the minimization problem.*

All the properties given for the meridian estimator are easily extended to the weighted meridian estimator case. The weighted meridian filter is also extended to admit real–valued weights utilizing the sign–coupling approach [6]. Table 1 summarizes the M–smoothers (weighted M–estimators) for existing and proposed filter families discussed in the paper. The (weighted) meridian filtering structure hence completes the missing link of the estimator/smoother quadruplet which is now composed of: (weighted) mean, (weighted) median, (weighted) myriad and (weighted) meridian.

4. NUMERICAL RESULTS

Recall that the meridian operator derives its optimality from the algebraic–tailed Generalized Cauchy distribution for $p = 1$, a distribution referred to as the Meridian. Thus in such environments, the meridian is the optimal estimator. Experiments validating this expected result have been carried out,

Table 1. M–smoother (weighted M–estimator) objective functions and outputs for various filter families [$\diamond$: Replication in median, $\circ$: weigthing in myriad operators.]

Cost Function	Filter Output
$\sum_{i=1}^{N} h_i(x_i-\beta)^2$	mean$\{\mathbf{h}\cdot\mathbf{x}_N\}$
$\sum_{i=1}^{N} h_i\vert x_i-\beta\vert$	median$\{\mathbf{h}\diamond\mathbf{x}_N\}$
$\sum_{i=1}^{N} \log\{\gamma^2+h_i(x_i-\beta)^2\}$	myriad$\{\mathbf{h}\circ\mathbf{x}_N;\gamma\}$
$\sum_{i=1}^{N} \log\{\delta+h_i\vert x_i-\beta\vert\}$	meridian$\{\mathbf{h}\star\mathbf{x}_N;\delta\}$

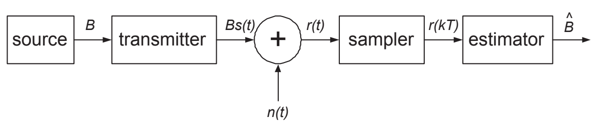


Fig. 2. The baseband communication model.

but are not presented due to their expectation nature and space constraints. Rather, results are presented for the commonly utilized α–Stable density family, which is also algebraic–tailed providing a fairer comparison.

Consider the baseband communication model [7] given in Fig. 2. Suppose that β (real) is to be communicated over the channel. Denoted as $s(t)$ is the combined impulse response of the transmitter and channel, and take the pulse $\beta s(t)$ to be corrupted by additive white Meridian noise. The received pulse is then given by [7]: $r(t) = \beta s(t) + n(t)$, which after sampling at rate $1/T$ corresponds to the sequence $r(kT) = \beta s(kT) + n(kT)$. Taking the common case assumption that $s(kT) \neq 0$ only for $k \in K$, the communications goal is to estimate β using the samples $\beta s(kT) + n(kT)$. Note that for a fixed k, $r(kT)$ is meridian distributed centered around the $\beta s(kT)$. By the whiteness assumption, the random variables $r(kT) - \beta s(kT)$ are independent with identical meridian distributions. This implies that the ML estimate for β is given by $\hat{\beta} = \arg\min_\beta \left[\prod_{k\in K}(\delta + |r(kT) - \beta s(kT)|)\right]$. Taking the natural log of the above and rewriting the sum yields

$$\hat{\beta} = \arg\min_{\beta}\left[\sum_{k\in K}\log\left\{\delta + |s(kT)|\left|\frac{r(kT)}{s(kT)} - \beta\right|\right\}\right] \quad (8)$$

from which it can be seen that the ML estimate is the weighted meridian of normalized received signal values, where pulse shape determines the weights. Thus, we define the *matched meridian filter*: $\hat{\beta} = \text{meridian}\{|s(kT)| \star r(kT)/s(kT); \delta\}$, that is matched to the pulse shape $s(kT)$, $k \in K$, as the value β minimizing (8).

Table 2. Matched Linear (MLi), Median (MMe), Myriad (MMy) and Meridian (MMer) filters output MAEs and MSEs

	MLi	MMe	MMy	MMer
MAE	5.9215×10^3	0.2080	0.1605	0.1121
MSE	3.2949×10^7	0.1687	0.0520	0.0380

To evaluate and compare the matched meridian filter to its linear, median, and myriad counterparts, 10000 Gaussian distributed $\{\beta(i) : i = 1, 2, \ldots, 10000\}$ parameters are generated, sent through the baseband communication channel, sampled with $K = 21$ and filtered with matched linear, median, myriad and meridian filters to obtain the estimates $\{\hat{\beta}(i) : i = 1, 2, \ldots, 10000\}$. The corrupting channel noise is α–Stable distributed with $\alpha = 0.4$. The pulse carrying the symbol is taken to be rectangular. The mean absolute and squared errors of the matched filter outputs are tabulated in Table 2. It is clear from the results that the matched meridian filter provides the best performance under both the mean absolute (MAE) and mean squared error (MSE) criterions. The performance improvement provided by the matched meridian filter especially stands out in the MSE case since this criteria is sensitive to outliers.

The numerical examples showing the performance improvements of meridian structure over mean, median and myriad counterparts in $\alpha < 1$ environments, involving a motivating powerline communications and a multi–tone signal processing example [6], are excluded here due to space constraints.

5. REFERENCES

[1] Huber, *Robust Statistics*. John Wiley & Sons, Inc., 1981.

[2] L. Yin, R. Yang, M. Gabbouj, and Y. Neuvo, “Weighted median filters: A tutorial,” *IEEE Transactions on Circuits and Systems*, vol. 41, May 1996.

[3] G. R. Arce, *Nonlinear Signal Processing: A Statistical Approach*. John Wiley & Sons, Inc., 2005.

[4] J. H. Miller and J. B. Thomas, “Detectors for discrete–time signals in non–Gaussian noise,” *IEEE Transactions on Information Theory*, vol. 18, pp. 241–250, Mar. 1972.

[5] C. L. Nikias and M. Shao, *Signal Processing with Alpha–Stable Distributions and Applications*. New York: Wiley, 1995.

[6] T. C. Aysal and K. E. Barner, “Meridian filtering for robust signal processing,” *IEEE Transactions on Signal Processing*, 2006, submitted for publication.

[7] J. Astola and Y. Neuvo, “Matched median fitlering,” *IEEE Transactions on Communications*, vol. 40, no. 4, pp. 722–729, Apr. 1992.