Adaptive weighted meridian nonlinear filter used for filtering of signal with impulsive noise

M. Stork

Abstract—The median, weighted median, myriad and weighted myriad filters have recently been proposed as a class of nonlinear filters for robust non-Gaussian signal processing in impulsive noise environments. A broad range of statistical processes is characterized by the generalized Gaussian statistics. For instance, the Gaussian and Laplacian probability density functions are special cases of generalized Gaussian statistics. Moreover, the linear and median filtering structures are statistically related to the maximum likelihood estimates of location under Gaussian and Laplacian statistics, respectively. Recently proposed myriad filtering is based on the maximum likelihood estimate of location under Cauchy and Laplacian statistics. An analogous relationship is formed here for the Laplacian statistics, as the ratio of Laplacian statistics yields the distribution referred here to as the Meridian. Therefore, the Meridian distribution is also a member of the generalized Cauchy family. Based on the maximum likelihood estimate under meridian statistics, the meridian filtering is proposed.

In this paper the basic properties concerning nonlinear filters and especially meridian filter are shown. The computer simulations are also presented.

Keywords—Median, myriad, nonlinear filter.

I. INTRODUCTION

The class of maximum-likelihood type estimators (M -estimators) of location, which were developed in the theory of robust statistics [1], has been of fundamental importance in the development of robust signal processing techniques [2]. Robust nonlinear estimators are critical for applications involving impulsive processes (e.g., communications systems, switching systems, biomedical signal processing radar clutter, ocean acoustic noise), where heavy-tailed non-Gaussian distributions are necessary in order to model the underlying signals [3], [4], [5]. Given a set of observations (input samples) \( \{ x_i \}_{i=1}^{N} \), an M - estimate of their common location is given by

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

(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 can be derived from maximum likelihood location estimator under Gaussian, Cauchy, and Laplacian statistics [6]–[7]. The cost functions to minimize, in these cases, are given by \( \rho(u) = u^2 \), \( \rho(u) = \log[K^2 + u^2] \), where \( K \) is the linearity parameter [8], [9] and \( \rho(u) = |u| \) for mean, myriad, and median estimators, respectively.

In is important to point that the Gaussian and Laplacian distributions are special cases of the generalized Gaussian distributions (GGD) (3) family corresponding to the \( k=2 \) and \( k=1 \), respectively, where \( k \) is the tail parameter [4]. Moreover, the well-established statistical relation between the Gaussian and Cauchy distributions, indicating that the ratio of two independent Gaussian random variables is Cauchy distributed and the Cauchy distribution is a special case of the generalized Cauchy distributions (GCD) family corresponding to \( p=2 \), where \( p \) is the tail parameter (13). An analogous statistical relationship is constructed here for the Laplacian distribution, where the distribution function of the random variable formed as the ratio of two independent Laplacian distributed random variables is derived. It is also 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 meridian statistics is analyzed, where the cost function, in this case, is given by

\[
\rho(u) = \log(\delta + |u|)
\]

(2)

with \( \delta \) controlling the robustness of the meridian estimator. Because the meridian estimator is likelihood-based, guarantees that the estimate is unbiased, consistent, and efficient in Meridian statistics.

II. STATISTICAL MODELS

A broad range of statistical processes can be characterized by the generalized Gaussian probability density function (PDF)

\[
f(x) = \frac{k\alpha}{2\Gamma(1/k)} \exp(\alpha |x|^k)
\]

(3)

where \( \Gamma(x) \) is the Gamma function (given by (4)), \( \alpha \) is a constant defined as \( \alpha = \sigma^{-1} \sqrt{\Gamma(3/k)(\Gamma(1/k))^{-1}} \) and \( \sigma \) is the standard deviation. The Gamma function is given by

\[
\Gamma(n) = \int_{0}^{\infty} t^{n-1} e^{-t} dt
\]
\[ \Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt \] (4)

The scale of the PDF is determined by the scale parameter \( \sigma > 0 \), and the impulsiveness is determined by the tail parameter \( k > 0 \). This representation includes the standard Gaussian PDF as a special case \((k=2)\). For \( k < 2 \), the PDF’s tail decays slower than in the Gaussian case, resulting in a heavier-tailed PDF. In a second case the Laplacian PDF is realized when \( k = 1 \).

Consider the problem of estimating the constant amplitude signal \( \theta \) from the samples \( x_1, x_{i+1}, \ldots, x_{i+N} \) of noisy observation data \( \{x(i)\} \). Let

\[ x_i = \theta + \eta_i \] (5)

where the \( \eta_i \) terms are independent and identically distributed zero-mean noise. The ML estimate of \( \theta \) is given by

\[ \hat{\theta} = \arg \max_{\theta} \left[ \prod_{i=1}^{N} f_x(x_i) \right] \] (6)

In the next part, the ML estimate under Gaussian and Laplacian statistics are shortly presented.

**Sample mean**

Consider a set of \( N \) independent samples \( x_1, x_2, \ldots, x_N \) each obeying the Gaussian distribution with variance \( \sigma^2 \). The ML estimate of location is given by

\[ \hat{\theta} = \arg \min_{\theta} \left[ \sum_{i=1}^{N} (x_i - \theta)^2 \right] \]

\( = \frac{1}{N} \sum_{i=1}^{N} x_i = \text{mean}\left\{x_i \mid i = 1 \right\} \) (7)

**Sample weighted mean** (finite impulse response (FIR) filter)

Consider a set of \( N \) independent samples \( x_1, x_2, \ldots, x_N \) each obeying the Gaussian distribution with (possibly) different variance \( \sigma^2 \). The ML estimate of location is given by

\[ \hat{\theta} = \arg \min_{\theta} \left[ \sum_{i=1}^{N} \frac{1}{\sigma_i^2} (x_i - \theta)^2 \right] \]

\( = \frac{1}{\sum_{i=1}^{N} h_i} \sum_{i=1}^{N} h_i x_i = \text{mean}\left\{h_i x_i \mid i = 1 \right\} \) (8)

where \( h_i = 1/\sigma_i^2 > 0 \).

**Sample median**

Consider a set of \( N \) independent samples \( x_1, x_2, \ldots, x_N \) each obeying the Laplacian distribution with (possibly) different variance \( \sigma^2 \). The ML estimate of location is given by

\[ \hat{\theta} = \arg \min_{\theta} \left[ \sum_{i=1}^{N} |x_i - \theta| \right] = \text{median}\left\{x_i \mid i = 1 \right\} \] (9)

**Sample weighted median**

Consider a set of \( N \) independent samples \( x_1, x_2, \ldots, x_N \) each obeying the Laplacian distribution with (possibly) different variance \( \sigma^2 \). The ML estimate of location \([8], [10]\) is given by

\[ \hat{\theta} = \arg \min_{\theta} \left[ \sum_{i=1}^{N} \frac{1}{\sigma_i^2} |x_i - \theta| \right] \]

\( = \text{median}\left\{h_i \odot x_i \mid i = 1 \right\} \) (10)

where \( h_i = 1/\sigma_i^2 > 0 \) and \( \odot \) is the replication operator as

\[ h_i \odot x_i = x_i, x_i, \ldots, x_i \] (11)

**Sample myriad**

Consider a set of \( N \) independent samples \( x_1, x_2, \ldots, x_N \) each obeying the Cauchy distribution with common parameter \( K \). The ML estimate of location \([8], [10]\) is given by

\[ \hat{\theta} = \arg \min_{\theta} \left[ \sum_{i=1}^{N} \log \left\{K^2 + (x_i - \theta)^2 \right\} \right] \]

\( = \text{myriad}\left\{x_i \mid i = 1 \right\}, K \) (12)

where \( K \) is the linearity parameter. For \( K \to \infty \), the myriad filter limiting to mean filter. The PDF of generalized Cauchy distribution is given by

\[ f(x) = aK + \left|x \right|^{p-2} \]

with

\[ a = \frac{p\Gamma(2/p)}{2(\Gamma(1/p))^2} \]

(13)

(14)

where \( K \) is the scale parameter and \( p \) is the tail constant. For \( p < 2 \) PDF tail decays slower than normal Cauchy case.

**Sample weighted myriad**

Consider a set of \( N \) independent samples \( x_1, x_2, \ldots, x_N \) each obeying the Cauchy distribution with (possibly) varying scale factor \( s_i = K/\sqrt{h_i} \) \([5], [6], [8], [9], [10]\). The ML estimate of location is given by

\[ \hat{\theta} = \arg \min_{\theta} \left[ \sum_{i=1}^{N} \log \left\{K^2 + h_i (x_i - \theta)^2 \right\} \right] \]

\( = \text{myriad}\left\{h_i \odot x_i \mid i = 1 \right\}, K \} \) (15)

where \( \odot \) denotes the weighting operation in the minimization problems.
Sample meridian
Consider a set of $N$ independent samples $x_1, x_2, \ldots, x_N$ each obeying the Meridian distribution with common scale parameter $K$. The ML estimate of location is given by

$$
\hat{\theta} = \arg \min_{\theta} \left[ \sum_{i=1}^{N} \log \left( K + |x_i - \theta| \right) \right] \tag{16}
$$

where $K$ is referred as the medianity parameter [11]. The PDF of Meridian distribution is given by

$$
f_M(x) = \frac{b}{2} \left( K + |x| \right)^2 \tag{17}
$$

where $b$ is scale parameter.

![Fig. 1. The meridian objective function (top) and derivation (bottom) for different values of parameter $K$. Left – $K=1$, right – $K=1000$ and samples $x_i=[-3; 5; -2; 3; 7]$.](image)

![Fig. 2. The meridian objective function for parameter $K=0; 0.5; 1; 3$ and samples $x_i=[-3; 5; -2; 3; 7]$.](image)

It is important to show that for $K>0$ and samples $x_i=[-3; 5; -2; 3; 7]$, sample meridian optimal values of $\theta$ are only 5 and 3. For $K \in (0; 0.605]$ is $\theta=5$ and for $K \in [0.606; \infty)$ is $\theta=3$. For $K \to \infty$ sample meridian converges to sample median.

The sample meridian function $F(\theta)$ holds following properties:

a) $F(\theta)$ is strictly decreasing for $\theta < x_1$ and strictly increasing for $\theta > x_N$.

b) All extremes of $F(\theta)$ are within the range of input samples $[x_1; x_N]$.

c) The meridian $\theta$ is one of the local minima, i.e. one of the input samples.

![Fig. 3. The optimal values of sample weighted meridian depending on $K$. Samples $x_i=[-3; 5; -2; 3; 7]$, weights $h_i=[1; 2; 3; 2; 1]$.](image)

Sample weighted meridian
Consider a set of $N$ independent samples $x_1, x_2, \ldots, x_N$ each obeying the Meridian distribution with (possibly) varying scale parameters $K/h_i$. The ML estimate of location is given by

$$
\hat{\theta} = \arg \min_{\theta} \left[ \sum_{i=1}^{N} \log \left( K + h_i |x_i - \theta| \right) \right] \tag{19}
$$

where $\bullet$ denotes the weighting operation in minimization task. The weighted meridian $\theta$ converges to the weighted median as $K \to \infty$. The optimal values of sample weighted meridian depending on $K$, for samples $x_i=[-3; 5; -2; 3; 7]$ and weights $h_i=[1; 2; 3; 2; 1]$ are shown in Fig. 3.

III. ROBUSTNESS AND INFLUENCE FUNCTION
Because the sample meridian belongs to the class of M-estimators [1], many robust statistics tools are available for evaluating its robustness [1], [12], [14]. M-estimators, also called generalized ML estimators, are formulated by (1) where $\rho(.)$ is an arbitrary objective function to be designed. Assuming that $\psi(x) = \partial \rho(x) / \partial x$ exists, the M-estimator is obtained by solving equation
\[ \sum_{i=1}^{N} \psi(x_i - \theta) = 0 \]  
\[ \text{(20)} \]

where \( \psi(.) \) is proportional to the influence function [12]. The influence function of an estimator determines the effect of contamination on the estimator. The influence function for sample mean is

\[ \psi(x) = 2x \]  
\[ \text{(21)} \]

for sample median

\[ \psi(x) = \text{sign}(x) \]  
\[ \text{(22)} \]

for sample myriad

\[ \psi(x) = 2x / (K^2 + x^2) \]  
\[ \text{(23)} \]

and for sample meridian

\[ \psi(x) = \text{sign}(x) / (K + |x|) \]  
\[ \text{(24)} \]

The influence functions for all privies estimators are shown in Fig. 4.

**B-robustness** - An estimator is B-robust if the supremum of the absolute value of the influence function is finite.

**Rejection Point** - The rejection point, defined as the distance from the center of the influence function to the point where the influence function becomes negligible and should be finite. Rejection point measures whether the estimator rejects outliers and, if so, at what distance.

The filtering error, in estimating a desired signal \( d(n) \), is:

\[ e(n) = y(n) - d(n) \]  
\[ \text{(26)} \]

The optimal filter weights minimize the MSE cost function:

\[ J(w) = \mathbb{E}\{e^2\} \]  
\[ \text{(27)} \]

where \( \mathbb{E}\{.\} \) denotes statistical expectation. In an environment of unknown or changing signal statistics, the LMS algorithm [15] attempts to minimize the MSE by continually updating the weights as

\[ h(n+1) = h(n) - \mu e(n) x(n), \]  
\[ \text{(28)} \]

where \( \mu > 0 \) is the so-called step-size of the update.

The computational simplicity of the LMS algorithm has made it an attractive choice for several applications in linear signal processing. However, it suffers from a slow rate of convergence. Further, its implementation requires the choice of an appropriate step-size \( \mu \) which affects the stability, steady-state MSE and convergence speed of the algorithm. The stability region for mean-square convergence [15], [16], [17] of the LMS algorithm is given by:

\[ 0 < \mu < (2/\text{trace}(R)), \]  
\[ \text{(29)} \]

\[ R = \mathbb{E}\{x(n)x^T(n)\}, \]  
\[ \text{(30)} \]

is the autocorrelation matrix of the input vector \( x(n) \). When the signal statistics are unknown or time-varying, it is difficult to choose a step-size that is guaranteed to lie within the stability region. The so-called Normalized LMS (NLMS) algorithm...
addresses the problem of the step-size design in [15], [18] by choosing a time-varying step-size $\mu(n)$ that minimizes the next-step MSE, $J_{n+1} = E[e^{(n+1)^2}]$. After incorporating an auxiliary fixed step-size $\mu_a > 0$, the NLMS algorithm [20] is written as

$$h(n+1) = h(n) - \mu_a e(n) \frac{x(n)}{||x(n)||^2}$$ (31)

where

$$||x(n)||^2 = \sum_{i=1}^{N} x_i^2(n)$$ (32)

is the squared Euclidean norm of the input vector $x(n)$. The theoretical bounds on the stability of NLSM algorithm are given by $0 < \mu_a < 2$ [15], [18], [20].

The described adaptive algorithms were evaluated through a computer simulation example involving sinus signal corrupted by impulse noise. Fig. 5 shows the block diagram representing our simulation example. The clean signal is corrupted by an additive noise process to yield the input or observed signal. The objective of the adaptive filtering algorithms is to train the meridian filter to converge to filter parameters $h$ (weights) that minimize the absolute value of the error signal between the filter output signal and the desired signal. In this section, we present the results of this training process, using learning curves and filter weight trajectories to demonstrate the convergence of the various values of adaptive parameters. In simulation example, the clean signal was chosen to be a sinus signal given by

$$x(n) = 2\sin(2\pi n/16666)$$ (33)

and max. amplitude of noise was also 2. The clean signal and signal + noise is shown in Fig. 6. The weight trajectories for adaptive algorithm according (31), during the training are shown in Fig. 7. The signal + noise and filtered signal after training are displayed in Fig. 8.

Fig. 8. Input and output signal, adaptive algorithm (31), $K=1$.

Fig. 9. Weight trajectories $h$ for adaptive weighted meridian filter.

$h = [0.1958 \; 0.1769 \; 0.1581 \; 0.1394 \; 0.1206 \; 0.1019 \; 0.083]$

Fig. 10. Input and output signal. ($K=1$)
In next simulation, the amplitude of noise was 4 times increased and NLMS algorithm was changed (for better stability [19]) to

$$h(n + 1) = h(n) - \mu_e e(n) \frac{x(n)}{0.1 + ||x(n)||^2}$$  \hspace{1cm} (34)

Weight trajectories, input and output signals are shown in Fig. 9 and Fig. 10 respectively.

V. CONCLUSIONS

In this paper a nonlinear meridian filter based on the maximum likelihood estimate under Meridian statistics was presented. The robustness of the meridian filter is shown through influence function and results of simulation examples. The meridian estimate is robust, and in addition, the rejection point is smaller than that of myriad estimation. This indicates that the meridian filter is more robust than the other filters. The adaptive meridian filter was also simulated and results were shown.

ACKNOWLEDGMENT

Milan Stork thanks of Department of Applied Electronics and Telecommunications University of West Bohemia, Plzen, Czech Republic for support in preparing of this paper.

REFERENCES


Mr. Milan Stork received the M. Sc degree in Electrical Engineering from the Technical University of Plzen, Czech Republic at the Department of Electronics in 1974 and Ph. D degree in automatic control systems at the Czech Technical University in Prague in 1985. In 1997, he became an Associate Professor at the Department of Applied Electronics and Telecommunication, faculty of Electrical Engineering on University of West Bohemia in Plzen, Czech Republic and full professor in 2007. He has numerous journal and conference publications. He is member of editorial board magazine “Physician and Technology”. His research interests include analog/digital systems, signal processing and biomedical engineering.