POLYNOMIAL WEIGHTED MEDIAN FILTERING

Kenneth E. Barner and Tuncer Can Aysal

Department of Electrical and Computer Engineering, University of Delaware, Newark, 19716 USA

ABSTRACT

Polynomial filter theory methods are able to approximate important classes of nonlinear systems. In the impulsive noise environments, however outliers are amplified by the polynomial structure, yielding poor performance. Median filters are well-known for their impulsive noise suppression and fine-detail preservation. In this paper, we extend the weighted median (WM) filtering to a polynomial class, Polynomial Weighted Median (PWM) filters which exploit higher order statistics while being robust to outliers. The PWM filter class is motivated by a study of the statistics of cross and square terms. The breakdown probability for the PWM is given and weight optimization is presented. Finally, the effectiveness of the PWM filter is shown through simulations.

1. INTRODUCTION

In many digital signal processing problems it is necessary to introduce nonlinear systems. A possible way to describe the input-output relation in a nonlinear system that is amenable to characterization, analysis, and synthesis, is to use a discrete Volterra series representation [1]. In many cases, a nonlinear system can be presented by a truncated version of Volterra series, which results in a simpler representation and requires a limited knowledge of the higher order statistics.

If the input-output relation is restricted to the quadratic term of a Volterra processor, the system becomes a second-order polynomial filter [2]. Indeed, the second-order filters are successfully used to address many digital signal processing problems, such as the optimal detection of a signal in Gaussian noise as well as texture discrimination. The quadratic structure of the second-order polynomial filter poses a significant problem in impulsive noise environments, as the cross and square terms residing in the second-order Kernel amplifies the effects of outliers. This paper develops a Polynomial Weighted Median (PWM) filter structure that is robust to noise, thus overcoming the short-comings of the conventional polynomial filtering. The proposed filter structure is motivated by WM optimality , and the statistics of the square and cross terms. The study of the heavy tails of the square and cross terms demonstrate that robust methods for their sample combinations should be considered to avoid undue effect of outliers. We also addressed the PWM filter weights optimization.

The remainder of paper is organized as follows. The statistics of cross and square terms are examined through the tail analysis in Section 2. In Section 3, the PWM filter derivation is introduced, the weights optimization scheme is given, and the breakdown probability analysis is presented. The simulations evaluating the performance of the proposed filter are presented in Section 4. Finally, the conclusions are drawn in Section 5.

2. STATISTICS ANALYSIS

The most popular extensions of the Gaussian distribution are those characterized by the generalized Gaussian distribution:

$$f(x) = \frac{k}{2\sigma\Gamma(1/k)} \exp\left(-|x-\mu|/\sigma\right)^k \tag{1}$$

where $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is the Gamma function. In this representation, the scale of the distribution is determined by $\sigma > 0$ and the impulsiveness is determined by $k > 0$. The presentation in (1) includes the standard Gaussian distribution as a special case for $k = 2$. For $k < 2$, the tails decay slower than in the Gaussian case, resulting in a heavier tailed distribution. A second special case that is of particular interest is $k = 1$, which yields the Laplacian distribution. It is well known that the linear and median filters are the Maximum Likelihood (ML) estimates of location under Gaussian and Laplacian distributions, respectively. The filters are smoothers when the weights are restricted to positive values, and more general filter characteristics are obtained by relaxing this constraint and allowing real-valued (positive/negative) weights.

Consider now the statistics of the cross and square terms that reside in the quadratic structure of the second-order polynomial filter. The effects of the product and square operators on a random variable's (*rv*) distribution's tail are of particular interest. In the following analysis we utilize the zero-mean Laplacian distribution:

$$\Psi_{x_i}(t) = \frac{1}{2\lambda_{x_i}} e^{-|t|/\lambda_{x_i}} \tag{2}$$

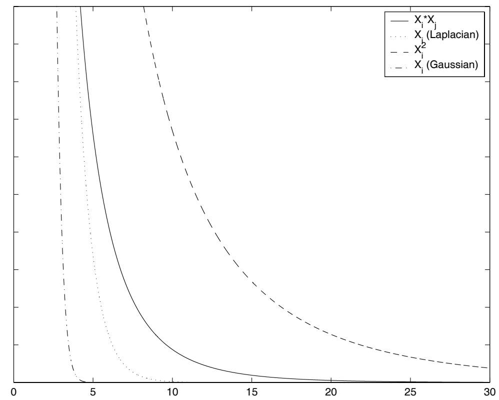

Fig. 1. The *pdf*s of x_i (Gaussian), x_i (Laplacian), $x_i x_j$ (both Laplacian) and x_i^2 (Laplacian) with zero-mean and $\sigma_{x_i} = \sqrt{2}$ and $\lambda_{x_i} = 1$ for Gaussian and Laplacian statistics, respectively

where λ_{x_i} is the scale parameter. The probability density function (*pdf*) of a *rv* generated by squaring a Laplacian distributed *rv* x_i is given by:

$$\Psi_{x_i^2}(t) = \frac{1}{2\lambda_{x_i}\sqrt{t}} e^{-\sqrt{t}/\lambda_{x_i}} \qquad (3)$$

Also, the *pdf* of a *rv* generated by the product of two independent Laplacian distributed *rv*s x_i and x_j, with scale parameters λ_{x_i} and λ_{x_j}, respectively, is given by:

$$\Psi_{x_i x_j}(t) = \frac{K_0\left(2\sqrt{\frac{|t|}{\lambda_{x_i}\lambda_{x_j}}}\right)}{2\lambda_{x_i}\lambda_{x_j}}, \qquad (4)$$

where $K_n(\cdot)$ is the modified Bessel function of the second kind of order n. For large values of t, $K_0(t)$ behaves like $1/(\sqrt{t})e^{-t}$ [3]. Thus $\Psi_{x_i x_j}(t)$ can be approximated as $1/(2\sqrt{2t^{1/2}})e^{-2\sqrt{t}}$ for $\lambda_{x_i} = \lambda_{x_j} = 1$. Also, Laplacian distribution (2) and equation (3) become $\Psi_{x_i}(t) = 1/2e^{-|t|}$ and $\Psi_{x_i^2}(t) = 1/(2\sqrt{t})e^{-\sqrt{t}}$, respectively, for $\lambda_{x_i} = 1$. For large t, the tail decay rate order is $\gamma_{x_i} > \gamma_{x_i x_j} > \gamma_{x_i^2}$, where γ_x denote the tail decay rate of $\Psi_x(t)$. The tails of Laplacian distribution in (2) and *pdf*s given in (3) and (4) are shown in Figure 1 for $\lambda_{x_i} = \lambda_{x_j} = 1$. As expected, the tail of $\Psi_{x_i x_j}(t)$ is heavier then that of $\Psi_{x_i}(t)$, and the tail of $\Psi_{x_i^2}(t)$ is heavier then that of $\Psi_{x_i x_j}(t)$. The heavy tails of the cross and square terms indicate that robust methods for their sample combinations should be considered to avoid undue influence of outliers.

3. MEDIAN-TYPE POLYNOMIAL FILTERING

It is known that polynomial models are capable of approximating a large class of nonlinear systems with a finite number of coefficients [2]. Consider the class of nonlinear, shift invariant systems with memory based on the discrete-time Volterra series. The input-output relation, in the case of a finite support, is given by:

$$y = \sum_{k=1}^{\infty} \bar{h}_k[x_i] \qquad (5)$$

where y is the output, $x_i, x_{i-1}, ..., x_{i-N+1}$ are the input samples, and $\bar{h}_k$ is defined by:

$$\bar{h}_k[x_i] = \sum_{i_1=0}^{N-1} \cdots \sum_{i_k=0}^{N-1} h_k(i_1, ..., i_k) x_{i-i_1} \cdots x_{i-i_k} \qquad (6)$$

Note that this is a causal formulation and that causality can be relaxed without conflict. Also, for *k*=1 in (6), the term $h_1(i_1)$ is the usual linear impulse response, and the term $h_k(i_1, ..., i_k)$ can be considered as the finite extent k-th order impulse response that characterizes the nonlinear behavior of the filter.

It is worth noting the use of the quadratic term in addition to the linear term is often sufficient to yield performance improvements [2]. The second-order polynomial filter is given by setting the upper limit in the above equation to K=2:

$$y = C_1 \sum_{i_1=0}^{N-1} h_1(i_1) x_{i-i_1} + C_2 \sum_{i_1=0}^{N-1}\sum_{i_2=0}^{N-1} h_2(i_1, i_2) x_{i-i_1} x_{i-i_2} \qquad (7)$$

where $h_1(i_1)$ is a $N \times 1$ vector representing the first-order, $h_2(i_1, i_2)$ is a $N \times N$ matrix representing the second-order Volterra Kernel, and C_1 and C_2 are constants [2]. Note that although the overall filtering operation is (polynomial) nonlinear, the filter output is linear with respect to the filter coefficients and observation sample, their squares, and cross terms. The linear combination of samples is the ML estimate of location only in Gaussian noise case. For heavier tailed distributions, more robust methods must be developed. Though it is mathematically intractable to derive the ML estimate for cross terms and squared observation samples under Laplacian statistics case, by analogy we can approximate their performance with median operators since they have heavier tails closer to the median optimal Laplacian then to the linear optimal Gaussian distribution. Thus, we define a powerful nonlinear second-order filter by separating all the linear, square, and cross terms contributions in (7), replacing the summation operators with median operators, and incorporating *sign coupling* accordingly [4]. The newly formulated filter is referred to as the Polynomial Weighted Median (PWM) filter:

$$\begin{aligned} y = {} & C_1' MED(|\mathbf{h}_1'| \diamond sgn(\mathbf{h}_1')\mathbf{x}_1) \\ & + C_{2,2}' MED(|\mathbf{h}_{2,2}'| \diamond sgn(\mathbf{h}_{2,2}')\mathbf{x}_{2,2}) \\ & + C_{2,1}' MED(|\mathbf{h}_{2,1}'| \diamond sgn(\mathbf{h}_{2,1}')\mathbf{x}_{2,1}) \end{aligned} \qquad (8)$$

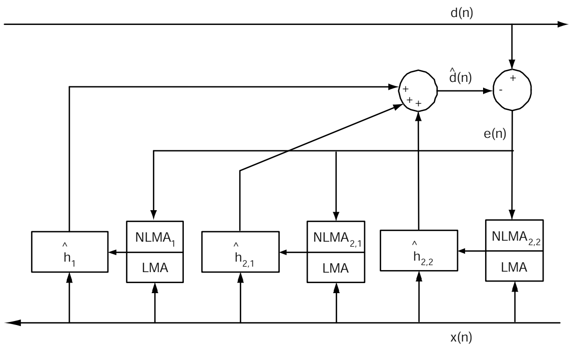

Fig. 2. The PWM filter coefficients optimization with the Normalized fast LMA

where $\mathbf{x}_1 = [x_i, x_{i-1}, ..., x_{i-N+1}]$, $\mathbf{x}_{2,2} = [x_i^2, x_{i-1}^2, ..., x_{i-N+1}^2]$,and $\mathbf{x}_{2,1} = [x_i x_{i-1}, x_i x_{i-2}, ..., x_i x_{i-N+1}, x_{i-1} x_{i-2}, x_{i-1}x_{i-3}, ..., x_{i-N+2}x_{i-N+1}]$, are the input sample, square, and cross term vectors, respectively, and $\mathbf{h}_1'$, $\mathbf{h}_{2,2}'$ and $\mathbf{h}_{2,1}'$ are first-order, square, and cross term filter Kernels, respectively. It is clear that if we set the new constants as $C_{2,1}'$=0, $C_{2,2}'$=0 and C_1'=1, then the PWM filter reduces to the traditional WM filter.

Consider next, the breakdown probability (*bdp*) of the PWM filter. The *bdp* is used as a measure of filter robustness [5], and is the probability of an impulse occurring at the output of the filter. It is related to the probability p of an impulse in the input. The WM filters in the PWM filter are nonlinearly coupled, to simplify the analysis we assume that the filters are independent, and given that an impulse occurring at the output of at least one of the WM operators will result a breakdown, yields the approximated $\beta(p)$ and defined as:

$$\beta(p) \approx 1 - (1 - \beta_{x_i}(p))(1 - \beta_{x_i^2}(p))(1 - \beta_{x_i x_j}(2p - p^2)) \tag{9}$$

where $\beta(p)$ denotes the *bdp* of the PWM filter and $\beta_{x_i}, \beta_{x_i^2}$, and $\beta_{x_i x_j}$, represent the *bdp* for first-order, square term, and cross term WM filters, respectively.

The filter weights optimization is realized through the fast LMA [4], which is modified to a Normalized fast LMA due to quadratic structure of PWM filter. Also, the coefficients C_1', $C_{2,1}'$, and $C_{2,2}'$ are adapted with the standard LMA. The complete optimization block diagram is shown in Figure 2.

4. SIMULATIONS

In this section, the proposed PWM filter is evaluated through several investigations. The breakdown probability of the PWM filter is implemented and compared with the second-order Volterra filter. Next, simulations in noise-free environment are presented comparing the Volterra and PWM filter outputs. Volterra and PWM filters are then tested utilizing heavy-tailed noisy signals. The power spectrum analysis is also applied to the filter outputs. Finally, the L_1 norm error for filter outputs, under different types of noise distributions, including Gaussian, Laplacian, and α-Stable, are given.

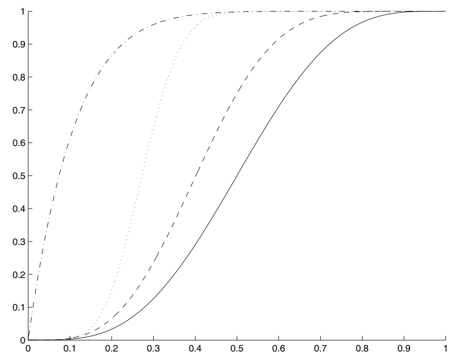

Fig. 3. The *bdp* of a single WM with x_i inputs, second-order PWM filter with x_i and x_i^2, second-order PWM filter with x_i, x_i^2, and $x_i x_j$, and standard second-order Volterra filter, are represented in solid line, dashed line, dotted line and dash-dotted line respectively, for N=7

In the implementation of the *bdp* for the filters, the PWM filter weights are set to unity, which yields the most robust case for each sub-filter. Fig. 3 gives the *bdp* of a standard WM filter, second-order PWM filter with linear and square terms, second-order PWM filter with linear, square, and cross terms, and second-order Volterra filter, all for N=7. The Fig. shows that the PWM filter is much more robust than the Volterra filter, with a *bdp* slightly larger than the robust WM. Thus the PWM filter offers far greater impulse rejection, and more robust performance than the standard Volterra filter.

The input signal used in the following experiments, is $sin(w_1 n)$+$sin(w_2 n)$, where w_1=$\pi/10$ and w_2=$\pi/50$. The PWM filter weights are adapted in a system identification configuration using the second-order Volterra filter output as the desired signal. Fig. 4 shows the outputs of the second-order Volterra filter and the PWM filter in noise-free and noisy environments. It is clear that the PWM filter performs very similarly to the Volterra filter, with absolute mean error of 0.1581 for the noise-free signal case. It can also be noted that in the Laplacian ($\mathcal{L}$), σ^2=0.25 noisy input case, that the PWM filter suppress the impulses, for instance at n=25, n=125, and n=225, while preserving the fine details. The information at these times however, is completely lost in the standard second-order Volterra output.

The input signal used in the following output power spectrum density (*psd*) analysis, is $sin(w_1 n)$+$sin(w_2 n)$, where w_1=$\pi/5$ and w_2=$\pi/10$. Fig. 5 shows the filters output *psd*s in Laplacian $\mathcal{L}$, σ^2=1 noisy input case. In this case, two ad-

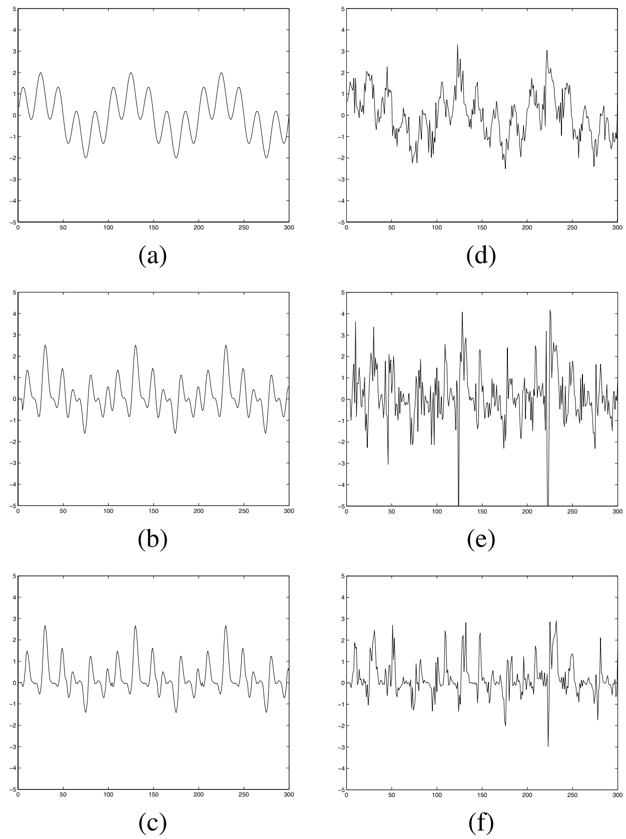

Fig. 4. (a) is noise-free input signal, (b) standard Volterra and (c) PWM filter output, (d) noisy input signal ($\mathcal{L}$ with σ^2=0.25), (e) standard Volterra and (f) PWM filter output for noisy input

ditional frequency components are expected in the *psd* estimation due to the quadratic structure of the Volterra system. The output frequency information is completely recovered in PWM filter case. The standard Volterra, however fails to recover the output frequency information due to the impulsive noise.

As a measure of the effectiveness of the filters in the noisy environments, the average L_1 norm error of the filter outputs and PWM filter performance gain, denoted as $\mathcal{K}$, are given in Table 1 for different noise distributions, including Gaussian ($\mathcal{N}$), Laplacian ($\mathcal{L}$), and α-Stable distributions, and powers. The statistics in Table 1 show that the PWM filter outperforms the standard second-order Volterra filter in heavy-tailed noise cases, and that the filters perform similarly in Gaussian ($\mathcal{N}$) noise case.

5. CONCLUSIONS

The PWM filter, novel polynomial filtering approach is proposed for heavy-tailed noise environments. The higher-order statistics of Laplacian distribution is analyzed. The breakdown probability of PWM is given, with the weights optimization scheme. It is shown, through the simulations, that the proposed PWM filtering overcomes the drawbacks of the standard Volterra filtering in impulsive noise.

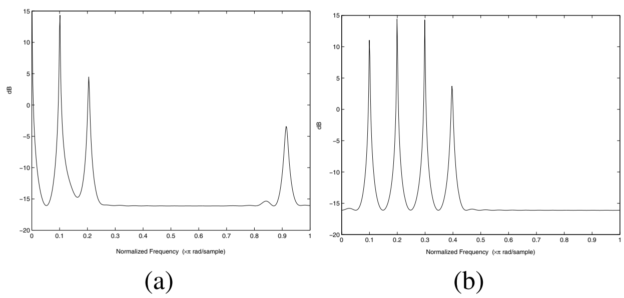

Fig. 5. (a) The standard Volterra filter output *psd* estimation and (b) PWM filter output *psd* estimation

Table 1. Mean Absolute Filtering Errors

Noise Power	*Volterra*	*PWM*	$\mathcal{K}$
Noise Free	-	0.1581	-
($\mathcal{N}$) σ^2=0.25	0.6139	0.6582	- 0.07 %
($\mathcal{L}$) σ^2=0.25	0.7839	0.4557	41.87 %
σ^2=1	1.8476	0.9397	51.75 %
σ^2=4	4.5489	1.6917	62.14 %
(α-Stable) α=1.4	22.1276	2.3063	89.58 %

6. REFERENCES

[1] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, Wiley, 1980.

[2] E. Biglieri, "Theory of volterra processors and some applications," in *Proc. IEEE ICASSP-82*, Paris, 1982, pp. 294–297.

[3] A. Erdelyi, W. Magnus, F. Oberhettinger, and F.C. Tricomi, *Higher Transcendental Functions, Vol. I, II and II, The Bateman Manuscript Project*, McGraw-Hill, 1953.

[4] G. Arce, "A general weighted median filter structure admitting negative weights," *IEEE Trans. on Signal Processing*, vol. 46, pp. 3195–3205, December 1998.

[5] G.R. Arce and R.E. Foster, "Detail preserving ranked-order based filters for image processing," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 37, no. 1, pp. 83–98, 1989.