Adaptive Weighted Vector Median Filter Using a Gradient Algorithm

Laurent Lucat and Pierre Siohan
France Telecom CNET/DMR/DDH,
4 rue du Clos Courtel, BP 59
35512 Cesson-Sévigné, FRANCE
pierre.siohan@cnet.francetelecom.fr

ABSTRACT

In this paper, a gradient based approach to adapt the parameters of the Weighted Vector Median Filter is presented. The validity of the method is inspected through a convergence test of the filter parameters and with results of noisy image filtering.

1 INTRODUCTION

Non-linear filters using order statistics are useful tools in digital image processing. Among them, the median filter is well known for its impulsive noise removing ability; however, its only tunable parameter is its sliding window, shape or size. Hence, more flexible filters, such as the weighted order statistics filters (WOSF) have been proposed; WOSF are parametrized by a set of N *weights*, where N is the number of samples included in the window, and a *rank* parameter. When the normalized rank is set to 0.5, the WOSF reduces to a weighted median filter. A judicious adjustment of the weights allows one for instance to reach a better preservation of desired patterns. This emphasizes the need to develop techniques for adequately choosing these parameters. A structural constraint approach [1] or an error gradient back-propagation algorithm [2] are efficient ways to adapt the WOSF parameters.
In order to process multi-component signals, standard and weighted median filters have been extended to the multi-variate case, leading to the vector median (VMF) and weighted vector median filters (WVMF), respectively [3]. As in the scalar case, WVMF are parametrized by a set of N weights. Adaptive WVMF have been proposed, for instance in [4]. However, no "optimal" WVMF has yet been developed, in the way of an optimization procedure, *e.g.* minimization of a cost function.

In this paper, we propose a design of the adaptive WVMF, based on the minimization of the quadratic or absolute error, using an error gradient back-propagation algorithm. This new technique can be used for the WVMF defined with both L_1 or L_2 metric.
In Section 2, we describe the proposed basic optimization procedure and a modified version. Validity of the method is inspected in Section 3 through the study of the parameters convergence, and filtering results using the proposed adaptive WVMF are presented in Section 4, for the case of noisy color textured image filtering.

2 GRADIENT BACK-PROPAGATION FOR WEIGHTED VECTOR MEDIAN FILTERS

2.1 Definition of the WVMF

The WVMF is defined by [3]

$$y = \arg\min_{x_j} \sum_{i=1}^{N} w_i \|x_j - x_i\|_{L_p}, \qquad (1)$$

where $\{x_i\}_{i=1..N}$ are the input vector samples included in the window, $\{w_i\}_{i=1..N}$ are the associated weights to be adapted, L_p is a norm (usually L_1 or L_2), and y is the filter's output.
When the *min* in eq.(1) is not unique, an additional rule is required to choose the WVMF output sample.

2.2 Optimization of the WVMF

The optimal WVMF is derived from the minimization of a cost function, which corresponds to the $\mathrm{ML_1E}$ or $\mathrm{ML_2E}$ criteria. They are respective extensions to vector data of the well-known MAE and MSE : $\mathrm{ML_1E} = \frac{1}{M}\sum_{i=1}^{M} \|d_i - y_i\|_{L_1}$ and $\mathrm{ML_2E} = \frac{1}{M}\sum_{i=1}^{M} \|d_i - y_i\|^2_{L_2}$, where $\{d_i\}_{i=1..M}$ are the noise-free vectors (desired output) and $\{y_i\}_{i=1..M}$ are the estimates given by the filter output. The optimization is conducted independently for each weight w_l; using the stochastic approximation, it can be summarized by

$$w_l^{k+1} = w_l^k + 2\mu(d-y)\frac{\partial y}{\partial w_l}, \quad 1 \le l \le N, \qquad (2)$$

for the quadratic error (LMS-type algorithm), and

$$w_l^{k+1} = w_l^k + \mu \; sgn_v(d-y)\frac{\partial y}{\partial w_l}, \quad 1 \le l \le N, \qquad (3)$$

for the absolute error case, where $sgn_v(a) = (sgn(a^1), ..., sgn(a^P))^T$, sgn denoting the sign function and $a = (a^1, ..., a^P)^T$, μ is the adaptation step, and w_l^k

is the value of the weight w_l at the k^{th} iteration.
Thus, the problem is to find a mathematical expression for the derivatives $\frac{\partial y}{\partial w_l}$, $1 \le l \le N$.

2.3 Mathematical expression for $\frac{\partial y}{\partial w_l}$

When the WVMF output is uniquely defined (i.e. there is a single sample corresponding to the *min* in eq.(1)), it can be demonstrated that the filter's output is invariant to an infinitesimal variation of its weights. Hence, the mathematical derivative $\frac{\partial y}{\partial w_l}$ is null for each l, therefore there is no need to look for an exact expression of the derivatives. Thus, we consider the following approximation

$$\frac{\partial y}{\partial w_l} = \lim_{\delta w_l \to 0} \frac{\delta y}{\delta w_l} \approx \lim_{\substack{|\delta w_l| \to 0 \\ \delta y \neq 0}} \frac{\delta y}{\delta w_l}, \quad 1 \le l \le N. \quad (4)$$

Let us look for the smallest increment δw_l of the weight w_l such that δy does not become null. Hence, the output of the WVMF switches from the sample x_{j_0} to a sample $x_{j_1(l)}$ due to a weight increment equal to $\delta w_l^{j_1}$. Using these notations, the derivative approximation can be written as

$$\frac{\partial y}{\partial w_l} = \frac{x_{j_1(l)} - x_{j_0}}{\delta w_l^{j_1}}, \quad 1 \le l \le N. \quad (5)$$

The problem is now to find the index $j_1(l)$ and the weight increment $\delta w_l^{j_1}$. Let us define $d_j = d(x_j) = \sum_{i=1}^{N} w_i \|x_j - x_i\|$. The WVMF expression in eq.(1) can be rewritten as $y = \arg\min_{x_j} d(x_j)$. The partial derivative of d_j with respect to w_l is

$$\frac{\partial d_j}{\partial w_l} = \|x_j - x_l\|. \quad (6)$$

Denoting by d_j' the aggregate weighted distance d_j after a weight incrementation δw_l, we get

$$d_j' = d_j + \int_{w_l}^{w_l + \delta w_l} \|x_j - x_l\| \, dw_l = d_j + \delta w_l \, \|x_j - x_l\|. \quad (7)$$

Hence, it can be seen that $\{d_j'\}_{j=1..N}$ are affine functions of the weight increment δw_l. In Fig. 1, we represent the $\{d_j'\}_{j=1..N}$ affine functions through the set of N straight lines $\{(D_j)\}_{j=1..N}$. The intersection points of each (D_j) with (D_{j_0}) are the key issue of the problem; they are computed for $1 \le j \le N$ (and for each l) through

$$\begin{aligned} d_j' = d_{j_0}' &\Leftrightarrow d_j + \delta w_l^j \, \|x_l - x_j\| = d_{j_0} + \delta w_l^j \, \|x_l - x_{j_0}\| \\ &\Leftrightarrow \delta w_l^j = \frac{d_j - d_{j_0}}{\|x_{j_0} - x_l\| - \|x_j - x_l\|}. \end{aligned} \quad (8)$$

Thus, we know each minimal weight increment δw_l^j of the weight w_l implying $d(x_j)$ becomes lower than $d(x_{j_0})$. Hence, for each l, $\delta w_l^{j_1} = \min_j |\delta w_l^j|$ corresponds to the *minimal* weight increment (or decrement) of w_l which allows the WVMF output to switch from the sample x_{j_0} to another sample, namely $x_{j_1(l)}$. In the example presented in Fig. 1, $j_0 = 4$ and $j_1(l) = 2$.

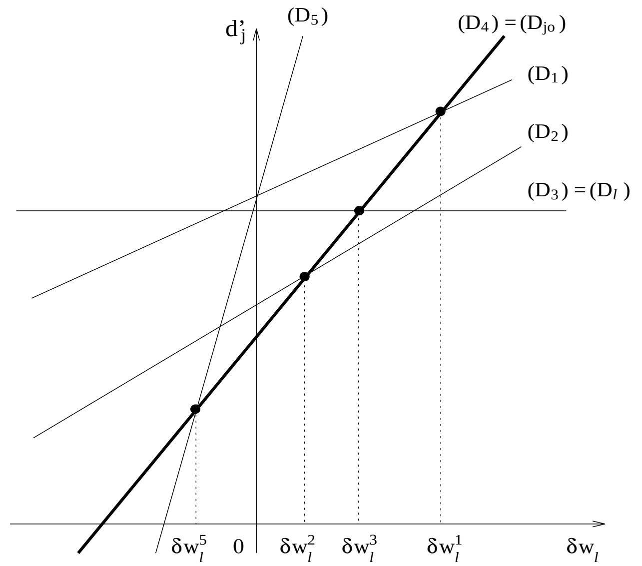

Figure 1: Evolution of the aggregate weighted distances according to the weight increment.

2.4 Alternative expression for $\frac{\partial y}{\partial w_l}$

The objective of this alternative is to estimate a more global behaviour of the filter when a weight is modified. This global analysis is motivated by the fact that a straight gradient approach, such as the one studied in section 2.3, is confined to the search of a local minimum of the cost function ; however, a local variation of the WVMF output due to a small weight increment may not be representative of the filter behaviour for larger fluctuations.
The alternative method principle is to take into account the increments δw_l^j, $j \neq j_0$, of the weight w_l which are necessary to make the WVMF output switch from the sample x_{j_0} to *each* sample $x_{j \neq j_0}$, when these switches are possible. These required weight increments can be seen as *critical* values in the filter output evolution, in the way that each one indicates a new direction (in the vector component space) of the WVMF output variation. Then, the global behaviour may be illustrated by the mean WVMF output increasing rate, computed from all critical points. This leads to the following expression of the filter output derivative

$$\frac{\partial y}{\partial w_l} \approx \left(\frac{\delta y}{\delta w_l}\right)_m,$$

$$\text{with } \left(\frac{\delta y}{\delta w_l}\right)_m = \frac{1}{card(S_l)} \sum_{x_j \in S_l} \frac{x_j - x_{j_0}}{\delta w_l^j}, \quad 1 \le l \le N, \quad (9)$$

where S_l denotes the whole set of potential output vectors involved when increasing or decreasing the weight w_l. This second method is expected to be less reactive than the first one but the weights evolution may be more regular because of the *mean* behaviour.
The aim is now to find the location of the critical points. These points correspond to each weight increment implying that a given d_j becomes lower that all other d_i (the straight line (D_j) is then under the other lines (D_i)). This leads to the search of a convex curve, such

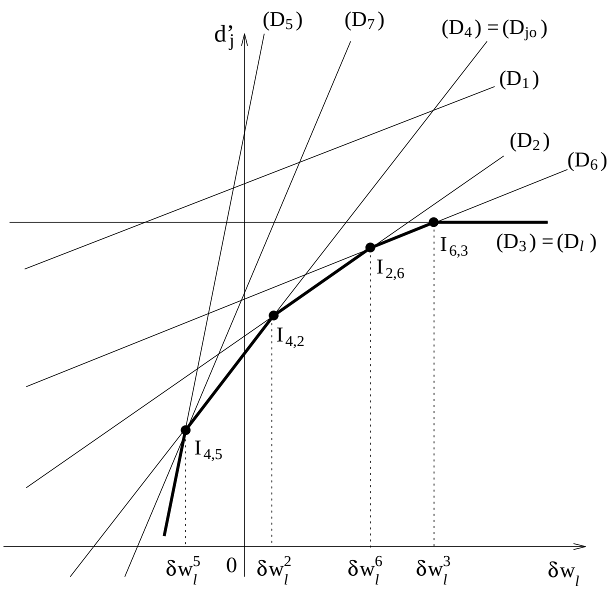

Figure 2: Iterative search of the intersection points for the alternative method.

as the one illustrated in Fig. 2. The intersection points are iteratively computed ; this search starts with the straight line (D_{j_0}) and is independently conducted for negative *vs* positive weight increments. At each step, the computation of the intersections between the current reference line (D_{j^*}) and the others (D_j) is obtained in the same way as in §2.3, using equation (8) where the index j_0 is then replaced with the current reference j^*. In the example of Fig. 2, the mean WVMF output increasing rate is then given by

$$\left(\frac{\delta y}{\delta w_l}\right)_m = \frac{1}{4}\left(\frac{\overrightarrow{X_4X_5}}{\delta w_l^5} + \frac{\overrightarrow{X_4X_2}}{\delta w_l^2} + \frac{\overrightarrow{X_4X_6}}{\delta w_l^6} + \frac{\overrightarrow{X_4X_3}}{\delta w_l^3}\right),$$

where $\overrightarrow{X_iX_j} = \overrightarrow{OX_j} - \overrightarrow{OX_i} = x_j - x_i$. As a matter of comparison, the method of §2.3 would result in $\frac{\delta y}{\delta w_l} = \frac{\overrightarrow{X_4X_2}}{\delta w_l^2}$, δw_l^2 being the smallest increment of the weight w_l implying a WVMF output switch.
Note that when a multi-intersection occurs, such as the point $I_{4,5}$ in Fig. 2, the new reference line is the one having the larger slope for negative δw_l and the lower one for positive δw_l, in order to keep the curve convexity.

3 PARAMETER CONVERGENCE TEST

The aim of the test is to study the evolution of the parameters, and more particularly to assess whether the parameters can converge to desired values. This test is illustrated in Fig. 3. The input signal consists of a one-

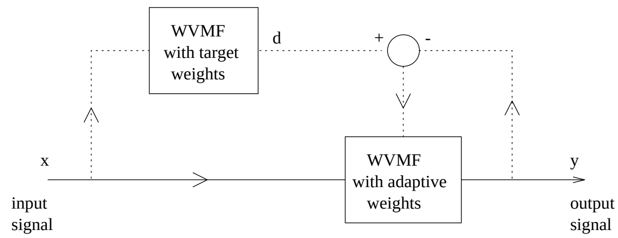

Figure 3: Schematic illustration of the convergence test.

dimensional (1D) fragment (70 pixels) extracted from a textured color image, and periodically duplicated. The 1-D WVMF having a size of 5 samples is defined with the L_1 metric and is optimized with the method presented in §2.3. The set of target weights is $\{1; 2; 3; 4; 5\}$ and the step μ is set to 10^{-7}. In Fig. 4, we give the results of the convergence test. It can be observed that the weights converge to values close to the normalized target weights. At the convergence, which is reached in this example after 620 iterations, the output y is equal to the desired output d. Other simulations with different sets of target weights, as well as using the alternative method of §2.4, lead to a weight convergence, for length pattern of up to 70 pixels. This "positive" test illustrates the validity of our optimization approach.

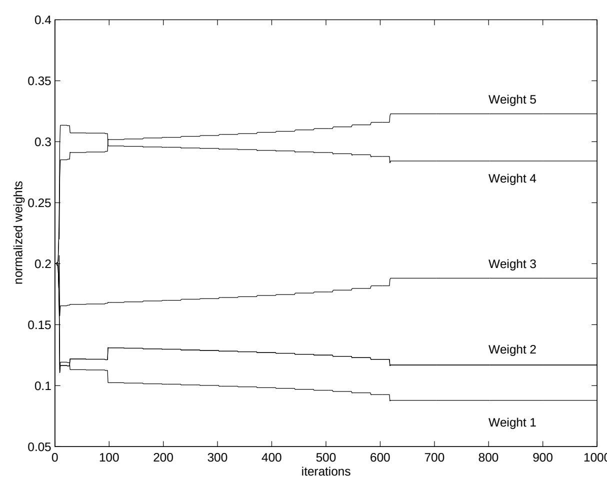

Figure 4: Evolution of the weights.

4 RESULTS OF COLOR IMAGE FILTERING

In this section, we evaluate the performance of the adaptive WVMF for removing impulsive noise. The input image is a texture embedded in a R-G-B channelindependent impulsive noise. The objective of the filtering is to remove the impulses while preserving the noise-free patterns. In Fig. 5, we present a black and white illustration of the original noisy image and the respective results of the adaptive WVMF and of the VMF (which is equivalent to an unoptimized WVMF whose weights are equal to 1), using a $5*5$ pixels mask, the L_1 metric and the optimization procedure of §2.4. This example emphasizes the benefit of using adaptive weights for a better preservation of fine details. We have to mention that the difference between the VMF and WVMF results is more sensitive when looking at the color images on the display or after color printing; visual disturbance due to the R-G-B impulses in the original image is also reduced by the black and white conversion.

Other simulations using the method of §2.3, the L_2 metric and other input textured images have also shown a significant improvement over the unoptimized filter. According to our simulations, the WVML$_2$ seems to work better with the optimization of §2.3 while the WVML$_1$ performs better using the method of §2.4 ; the latter leads to the best results among the possible configurations generated by the two norms and the two optimization versions. Comparative ML$_2$E criteria are presented in Table 1.

	VMF	WVMF initial method	WVMF alternative method
L_1	1123	995	**865**
L_2	1162	880	923

Table 1: Comparative ML$_2$E for the VMF and the optimized adaptive WVMF defined with L_1 or L_2.

The optimization can be conducted without any reference to the noise free image ; the corrupted image is then considered as the reference. Further simulations have shown that the generated degradation of the performances is weak, with regards to the gain due to the adaptive filtering.

Finally, we have to note that the central pixel of the mask has been ignored here ($w_c = 0$). The specification of w_c should further improve the filtering ; for instance, w_c should be chosen according to the occurrence probability of impulses in the case of impulse noise filtering.

References

[1] R. Yang, L. Yin, M. Gabbouj, J. Astola, and Y. Neuvo. "Optimal Weighted Median Filters under Structural Constraints". *Proc. of IEEE ISCAS*, pp. 942–945, 1993.

[2] M. Ropert and D. Pelé. "Synthesis of Adaptive Weighted Order Statistics Filters with Gradient Algorithms and Application to Image Processing". *Proc. of the IEEE ICIP*, volume 2, pages 512–516, November 1994.

[3] J. Astola, P. Haavisto, and Y. Neuvo. "Vector Median Filters". *Proc. of the IEEE*, vol. 78-4, pp. 678–689, 1990.

[4] L. Alparone *et al.* "Adaptively Weighted Vector-Median Filters for Motion-Fields Smoothing". *Proc. of IEEE ICASSP*, volume 4, pages 2267–2270, April 1996.

Original noisy textured image.

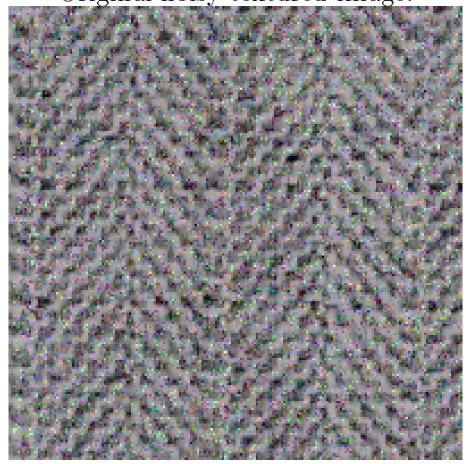

Image filtered by a 5*5 VMF.

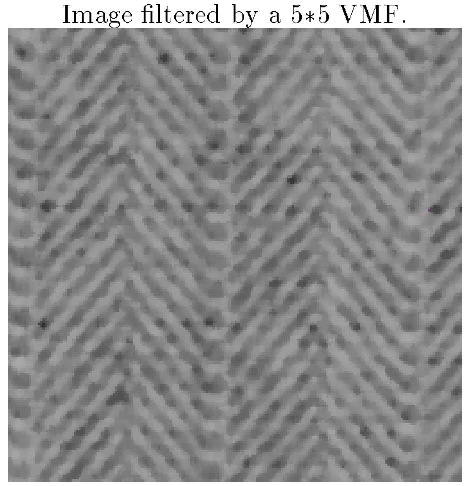

Image filtered by a 5*5 adaptive WVML$_1$.

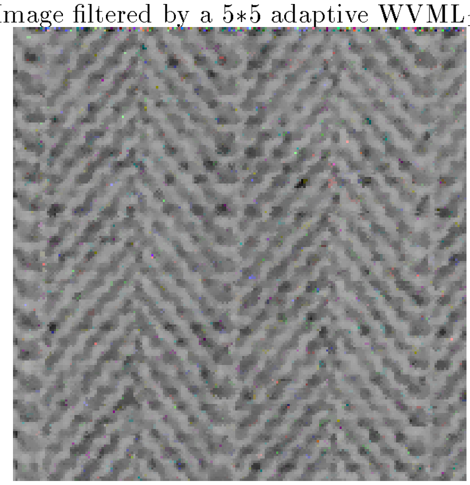

Figure 5: Comparison of the VMF and of the adaptive WVMF for impulsive noise filtering.