Non-Linear Local Harmonic Filters For Edge-Preserving Image Denoising

Miroslaw Pawlak  
Dept. of Electrical and Computer Eng.  
The University of Manitoba  
Winnipeg, Manitoba, Canada

Ewaryst Rafajlowicz  
Institute of Engineering Cybernetics  
Wroclaw University Of Technology  
Wyb. Wyspianskiego 27, Wroclaw, Poland

Abstract

We propose a new class of non-linear image recovery operators based on local smoothing applied both in the spatial domain (horizontal smoothing) and in the gray-level domain (vertical smoothing). The commonly used horizontal smoothing gives the way of reducing the noise level (noise filtering), whereas the vertical smoothing yields the edge preservation property. The method is indexed by a real parameter yielding a generalized class of local harmonic filters. In particular the classical Min/Max filters as well as all linear locally weighted filters can be obtained as special cases.

1. Introduction

The problem of designing an image filtering and recovering algorithms which are able to combat noise and in the same time to preserve edges has a long history in the field of image analysis. It is well known that linear filters are able to reduce noise but at the expense of smearing edges. Nonlinear filters [11] on the other hand are much more useful in this respect and the most widely used filters of this type are order statistics filters and sigma-filters [11], [6]. These are powerful methods but they are not flexible enough to be able to adapt to complex structures and complex noise models. In fact they do not possess any regularization parameters which (if selected properly) allow a designer to tune a method. It should be also noted that non-linear filtering techniques are unable to inform us what order of the accuracy we can expect when the image resolution increases. There are other methods which are based on the idea of image segmentation, i.e., either finding approximately homogeneous regions in the image or estimating edges [5], [1]. In this paper we present an approach which estimates an image directly without estimating edges but which still pays special attention to the quality of estimation near discontinuities. Relevant approaches concern with jump preserving regression estimation have been proposed in [2], [4], [7], [12], [13], [9], [10]. The aim of this paper is to propose a new family of image edge preserving filters, which stem from the concept of weighted regression proposed in [9], [10]. The weighting in both horizontal and vertical directions, the former leading to image dependent weighting, are applied. The proposed family of filters is parametrized in such a way that it covers a wide range of filter operators including the classical from Min/Max filters, sigma filters, and all linear locally weighted smoothers [11], [6], [3].
2. Preliminaries and Harmonic Filters

Let \( x \in X \subset R^2 \) denote a position of a pixel in a gray scale image. Function \( m^*(x) \) represents unknown ("true") image intensity. The observation model relating \( m^*(x) \) to measured image intensity \( Y \) is as follows

\[
Y = m^*(x) + \varepsilon, \quad (1)
\]

where \( \varepsilon \) is a random error with \( E(\varepsilon) = 0, E(\varepsilon^2) < \infty \). The density function of \( \varepsilon \) is denoted by \( f_\varepsilon \) and is unknown. For simplicity of derivations we have assumed the additive noise model, but our methodology can be easily carried over to various signal dependent noise models like multiplicative noise, film-grain noise, Poisson noise etc. From (1) it is seen that for fixed \( x \), the density of \( Y \), in the sequel denoted by \( f(y;x) \), is given by \( f(y;x) = f_\varepsilon(y - m^*(x)) \). Further all the expectations are calculated with respect to \( f(y;x) \) and are denoted by \( E_x \).

Let \( w(x,y) > 0 \) be a weight function defined by the experimenter such that \( E_x(Yw(x,Y)) < \infty, E_x(w(x,Y)) < \infty \) for every \( x \in X \). For a given pixel \( x \) let us define the vertically weighted mean squared error

\[
Q(m;x) = E_{w}(w(x,Y) \cdot (Y - m)^2). \quad (2)
\]

The minimizer \( m_w \) of \( Q(m;x) \) with respect to \( m \) is given by

\[
m_w(x) = E_{x}(Yw(x,Y)) / E_{x}(w(x,Y)), \quad x \in X. \quad (3)
\]

It is natural to name \( m_w \) as the vertically weighted regression. This term was coined and examined in [9], [10] in the context of jump preserving signal estimation. Let us note that if one sets \( w(x,y) \equiv 1 \) then \( m_w(x) = E_{x}Y \) - the ordinary regression function of \( Y \) on \( x \). An interesting class of weight functions \( w(x,y) \) which will be utilized in this paper is of the following form \( w(x,y) = y^p \cdot k(y - m^0(x)) \), where \( p \) is the real valued parameter and the functions \( m^0(x), k(y) \) are specified by the experimenter. The function \( k(y) \) plays role of the smoothing kernel in the vertical direction and can be selected as a compactly supported function over a certain interval. The interval length will play role of the amount smoothing applied in the \( y \) direction. On the other hand the function \( m^0(x) \) represents the amount of prior information one has about the unknown image \( m^*(x) \). It is often realized in the form a pilot (preliminary) estimate of \( m^*(x) \).

The vertically weighted regression for the above class of \( w(x,y) \) takes the following form

\[
m_w(x) = \frac{E_{x}[Y^{p+1} k(Y - m^0(x))]}{E_{x}[Y^p k(Y - m^0(x))]}, \quad x \in X. \quad (4)
\]

To make the interpretation of (4) more clear, set \( k(y) \equiv 1 \) for a while. Then, for \( p = 0 \), \( m_w(x) = E_{x}Y \) is the already mentioned regression function, setting \( p = -1 \) we obtain \( m_w(x) = 1/E_{x}(Y^{-1}), \) i.e., the conditional version of the harmonic mean. Generally we will refer to (4) as the generalized harmonic regression. It is worth mentioning that \( m_w(x) \) in (4) becomes the Max operation if \( p \to -\infty \). This case is further called a super-harmonic mean. Analogously, if \( p \to +\infty \), then (4) plays the role of the Min operation and it is called a sub-harmonic mean. Conerning \( k(y) \) we assume that it is a positive and symmetric function with compact support.

Let \( x_i, i = 1,2, \ldots, n \) denote positions of pixels at which noisy samples of the image are available. It is assumed that \( x_i \)'s are placed in the nodes of an equidistant grid in \( X \). Assume that we are given observations \( (x_i, Y_i), i = 1,2, \ldots, n \), which satisfy

\[
Y_i = m^*(x_i) + \varepsilon_i, \quad i = 1,2, \ldots, n, \quad (5)
\]

where \( \varepsilon_i, i = 1,2, \ldots, n \) is a sequence of uncorrelated random variables with the same distribution as \( \varepsilon \).

Based on the observed data given in (5) we would like to construct an empirical counterpart of the generalized harmonic regression function introduced in (4). This parallels the developments in the well established field of nonparametric regression estimation [3] and the references cited therein. In particular we propose to use the kernel regression estimate, being a popular smoothing technique. It should be stressed, however, that the classical kernel estimate has no edge-preserving properties. The kernel estimate requires a kernel function \( K(x) \) which provides weighting in the \( x \) domain. We assume that \( K(x) \) is a positive and symmetric function such that \( \int K(x)dx = 1 \). Replacing in (4) the expectations by their local weighted empirical means, we obtain the following estimator \( \hat{m}_w(x) \) of \( m_w(x) \)

\[
\hat{m}_w(x) = \frac{\sum_{i=1}^{n} Y_{i}^{p+1} k \left( Y_{i} - m^0(x) \right) K ((x - x_i)/h)}{\sum_{i=1}^{n} Y_{i}^{p} k \left( Y_{i} - m^0(x) \right) K ((x - x_i)/h)}, \quad (6)
\]

where \( h \) is the bandwidth (or smoothing) parameter, which allows to localize the estimator to a region in the \( x \) domain. It should be noted that we have in (6) both the vertical kernel \( k(y) \) and the horizontal kernel \( K(x) \). The width of the kernel \( K(x) \) is controlled by the smoothing parameter \( h \). Similarly the kernel \( k(y) \) can have a smoothing parameter, i.e., we can replace \( k \left( Y_{i} - m^0(x) \right) \) in (6) by \( k \left( (Y_{i} - m^0(x))/H \right) \), where \( H \) is the vertical smoothing parameter. The results on the consistency of the kernel regression estimate, which are proved in [8] can be useful in establishing the convergence properties for \( \hat{m}_w(x) \) in regions far from sharp
edges. This requires some conditions on the smoothing parameters \( h, H \).

The estimate \( \hat{m}_{w}(x) \) in (6) has a number of free "parameters" allowing to derive a large class of filters by specializing the kernels \( k, K \) and the function \( m^0(x) \). Good reconstruction results can be obtained even for a simple choice of the aforementioned parameters. Hence let \( R(t) = 1 \), if \(|t| \leq 1\) and \( R(t) = 0 \) for \(|t| > 1\). Then let us set \( k(t) = R(t/H) \), where \( H > 0 \) is the height of the vertical window. Hence we have the rectangular vertical kernel with the height \( H > 0 \). We also define the rectangular kernel in the horizontal direction, i.e., let \( K(x) = R(x(1)) \cdot R(x(2)) \). For filtering at the pixel \( x_j \) let us define a set of indexes \( A_j(h) = \{ i : \| x_j - x_i \| \leq h \} \).

Concerning the choice of \( m^0(x_j) \), it would be ideally to set \( m^0(x_j) \), which is however not available. One may use various pilot estimates of \( m^0(x_j) \), e.g., linear combinations of \( \{ Y_i, i \in A_j(h) \} \) or their ordered counterparts like a median filter. The simplest strategy for \( m^0(x_j) \) is \( Y_j \), i.e. we set \( m^0(x_j) = Y_j \). Denote by \( \hat{m}_j(p) \) the filtered value of the image at the pixel \( x_j \). The above leads to the following specific class of harmonic filters: for \( j = 1, 2, \ldots n \)

\[
\hat{m}_j(p) = \sum_{i \in A_j(h)} Y_{i}^{p+1} \cdot R \left( \frac{Y_j - Y_i}{H} \right) \cdot \sum_{i \in A_j(h)} Y_{i}^{p} \cdot R \left( \frac{Y_j - Y_i}{H} \right), \tag{7}
\]

Let us note again that the factor \( R \left( \frac{Y_j - Y_i}{H} \right) \) provides the vertical weighting which attempts to preserve edges in the image. For our particular choice of \( R(y) \) this factor takes value 1 if \(|Y_i - Y_j| \leq H \) and zero otherwise. Thus, the nominator of \( \hat{m}_j(p) \) is the sum of such \( Y_{i}^{p+1}, i \in A_j(h) \) which belong to relatively homogeneous region in the image. For large positive \( p \) this sum is dominated by \( \max_{i \in A_j(h)} \left\{ Y_i^{p+1} \cdot R \left( \frac{Y_j - Y_i}{H} \right) \right\} \).

Analogously, the denominator of (7) is dominated by \( \max_{i \in A_j(h)} \left\{ Y_i^{p} \cdot R \left( \frac{Y_j - Y_i}{H} \right) \right\} \) and in the limit we obtain the vertically weighted local Max filter

\[
\lim_{p \to \infty} \hat{m}_j(p) = \max_{i \in A_j(h)} \left\{ Y_i \cdot R \left( \frac{(Y_j - Y_i) / H} \right) \right\}.
\]

Similar considerations lead to the conclusion that \( \hat{m}_j(p) \) converges to the smallest among those \( Y_i, i \in A_j(h) \), which are sufficiently close to \( Y_j \), i.e., when \(|Y_i - Y_j| \leq H \).

Our filtering algorithm in (7) requires the specification of the parameters \( p, h, H \). The universal method for finding these parameters would rely on the cross-validation strategy, which needs to be modified for our purposes. Details of various possible methods of selection of \( p, h, H \) will be reported elsewhere. The rule of thumb in selecting \( p \) is to use larger positive \( p \), if the noise has intensive dark gray component and to apply more negative \( p \), if a light gray component is dominating in the noise.

A real image (see Fig. 1) was corrupted by adding errors not only to the present pixel but also to its 8 neighbors. This image was filtered by sub-harmonic filter \((p = -12)\) with \( h \) chosen so as to capture \( 3 \times 3 \) pixels in the horizontal direction and \( H = 0.5 \) in the vertical direction. For comparison the classical \( 3 \times 3 \) averaging filter was also applied to the same image, but it occurred to be quite useless in this case.

Analogously, in Fig. 2 and Fig. 3 the performance of sub- and superharmonic filters is compared with the performance of the median and the sigma filters, which are much less efficient (or even degrading) in filtering underexposed and overexposed noisy images.

References

Figure 2. Heavily corrupted image (upper left panel) filtered by superharmonic filter with $p = 16$, $H = 0.6$ (lower left panel). For comparison $5 \times 5$ median filter and sigma-filter (upper right and lower left panels, respectively).

Figure 3. Heavily corrupted image (upper left panel) filtered by subharmonic filter with $p = -16$, $H = 0.5$ (lower left panel). For comparison $5 \times 5$ median filter and sigma-filter (upper right and lower left panels, respectively).

Figure 4. Artificial image (upper left panel) corrupted by noise with the mean squared error $Q = 0.11$ (upper right panel) was filtered by the superharmonic $3 \times 3$ filter with $H = 0.4$ and $p = 10$ (lower right panel), providing noise reduction to $Q = 0.028$. For comparison, $3 \times 3$ moving average filter was applied leading to increase of error to $Q = 0.13$ (lower left panel).


