ADAPTIVE WINDOW SIZE IMAGE DENOISING BASED ON ICI RULE

Karen Egiazarian, Vladimir Katkovnik and Jaakko Astola

Signal Processing Laboratory, Tampere University of Technology, Tampere, Finland
e-mail: karen@cs.tut.fi

ABSTRACT

An algorithm for image noise-removal based on local adaptive window size filtering is developed in this paper. Two features to use into local spatial/transform-domain filtering are suggested. First, filtering is performed on images corrupted not only by additive white noise, but also by image-dependent (e.g. film-grain noise) or multiplicative noise. Second, used transforms are equipped with a varying adaptive window size obtained by the intersection of confidence intervals (ICI) rule. Finally, we combine all estimates available for each pixel from neighboring overlapping windows by weighted averaging these estimates. Comparison of the algorithm with the known techniques for noise removal from images shows the advantage of the new algorithm, both quantitatively and visually.

1. INTRODUCTION

Transform domain signal denoising finds applications in restoration of different type of one and two dimensional signals. Depending on imaging systems different noise models were considered - starting from additive white noise, data-dependent (e.g. film-grain type) noise, to multiplicative noise. Image processing in the transform domain rather than in the spatial domain has certain advantages of incorporating a priori knowledge on images into design of processing algorithms and in terms of computational expenses. The transfer from the spatial domain into the transform domain is especially useful if it is applied locally rather than globally. Having an excellent performance in suppressing of the Gaussian noise, transform based methods work fairly well also in several applications where the error is neither white nor Gaussian [1]. These applications are noise reduction (denoising) of synthetic aperture radar (SAR) signals, medical and geophysical signals, as well as removing blocking and ringing artifacts from images of JPEG and wavelet decoded images [1, 2, 3].

Nonlinear filtering in the wavelet transform domain was introduced in terms of wavelet denoising by Donoho and Johnstone [5] and has been extended by several authors. In [1, 6] translation invariant wavelet denoising algorithms were introduced and tested on different one dimensional signals and SAR images, respectively.

In [7] the local average transform domain denoising was presented. The difference between this filter and the one in [4] is that the nonlinear modification of the transform coefficients within a sliding window gives the estimate for the overall subimage within the window and not only at the central pixel as it is in [4]. Thus, it makes an overlap of the estimates in the neighboring windows and we obtain the multiple estimates for each pixel. All of the above estimates are averaged in order to obtain the final estimate for each pixel.

The filters presented in [4] and [7] are used in this paper as a starting "prototype" filter. We are going to equip it by two additional features:

1) Transforms are used with varying directional adaptive size windows. The intersection of confidence intervals (ICI) rule is applied for a window size selection [8, 9];

2) The algorithms work on images corrupted not only by additive white noise, but also by image-dependent noise, as well as by multiplicative noise.

Extensive experiments confirm the expected improved performance of the proposed filter for different noise models.

2. FILTERING IN TRANSFORM DOMAIN

Consider an observed noisy image $y(i,j)$ modeled as

$$y(i,j) = x(i,j) + e(i,j) = x(i,j) + x(i,j)^{\gamma} n(i,j), \tag{1}$$

where $x(i,j)$ is the noise-free image and $n(i,j)$ is zero-mean noise with the variance σ_n^2.

Note, that in the case of different values of γ this model will coincide with: an additive noise model, $\gamma = 0$, an image-dependent additive noise model, $0 < \gamma < 1$ (e.g. a film-grain noise, if $\frac{1}{3} \leq \gamma \leq \frac{1}{2}$), a multiplicative speckle noise model, $\gamma = 1$.

The main reason to make filtering in a transform domain rather than in a time (spatial) domain is due to decorrelating properties of some transforms.

Let an orthogonal transform be defined by $N \times N$ matrix H, $H^T H = I$. The observation model (1) in the transform domain and in matrix notation is of the form

$$Y = X + E, \tag{2}$$

where

$$Y = H^T y H, X = H^T x H, E = H^T e H.$$

The filtering into the transform domain can be done, say, by the hard shrinkage (rejecting filtering) defined as [5]

$$\hat{Y}(i,j) = \begin{cases} Y(i,j), \text{ if } |Y(i,j)| \geq \beta_n, \\ 0, \text{ otherwise,} \end{cases}. \tag{3}$$

where β_n is a threshold of the filter.

Inversion of the transform gives the estimates of the signals as

$$\hat{y} = H\hat{Y}H^T. \tag{4}$$

In the local transform based filtering the procedure is applied not to the transform coefficients of the entire image but to a block of the image in sliding filtering fashion (running, overlapping blocks). It is shown in [7] that keeping all filtered outputs for all windows improves the performance of de-noising if we combine all available estimates for each pixel coming from different windows it belongs to.

3. VARYING WINDOW SIZE SELECTION

In this section we present a modification of ICI rule [8, 9, 10] corresponding to the observation model (1). Let an estimate of the signal is given as the sample average

$$\hat{x}_h(i,j) = \frac{1}{h^2}\sum_{u_1,u_2} \rho(u_1/h, u_2/h) y(i+u_1, j+u_2) \tag{5}$$

with a normalized mask $\rho > 0$, $\frac{1}{h^2}\sum_{u_1,u_2} \rho(u_1/h, u_2/h) = 1$, where h is a scale (size) parameter of the mask. Then, the estimation error is $e(i,j) = x(i,j) - \hat{x}_h(i,j)$. It can be shown that for the estimation bias $|E\{e(i,j)\}| \leq Ch \triangleq \omega_{i,j}$, $C = A \cdot max_{i,j}\{|\frac{\partial}{\partial i}x(i,j)| + |\frac{\partial}{\partial j}x(i,j)|\}$, where A is a constant, and the standard deviation of the error is of the form $std_{i,j} = B\sigma/h$, $B^2 = \frac{1}{h^2}\sum_{u_1,u_2} |x(i+u_1, j+u_2)|^{2\gamma}$.

Then we have for the mean squared error (MSE):

$$MSE \leq \omega_{i,j}^2 + std_{i,j}^2 = C^2h^2 + B^2\sigma^2/h^2.$$

Let us assume that B at least locally does not depend on h. Then, the minimum of MSE on h is achieved at $h = h^*$, $h^* \simeq \sqrt{\frac{B\sigma}{C}}$. It can be verified that the ratio of the bias to the standard deviation of the error is constant for $h = h^*$, $\omega_{i,j}/std_{i,j} = 1$.This h^* provides the optimal balance between the bias and the random error of estimation.

It can be seen in this analysis that

$$\omega_{i,j} \leq std_{i,j}, \text{ for } h \leq h^*. \tag{6}$$

Then $|e(i,j)| \leq \omega_{i,j} + |\xi_{i,j}|$, where asymptotically the random term $\xi_{i,j}$ is Gaussian and with the probability $p = 1 - \alpha$ the following inequality holds

$$|e(i,j)| \leq \omega_{i,j} + \chi_{1-\alpha/2} std_{i,j}, \tag{7}$$

where $\chi_{1-\alpha/2}$ is $(1-\alpha/2) - th$ quantile of the standard Gaussian distribution. Now we introduce a finite set of increasing window sizes $H = \{h_1 < h_2 < .... < h_J\}$, starting with quite a small h_1.

Then, according to (6) the inequality (7) can be weakened for $h \leq h^*$ to

$$|e(i,j)| \leq (1 + \chi_{1-\alpha/2}) std_{i,j}. \tag{8}$$

In what follows we use the inequalities (8) corresponding to different h in order to test the hypothesis $h \leq h^*$ and to find the value of h close to h^*. According to (8) determine a sequence of the confidence intervals $D(k)$ of the biased estimates as follows

$$D(k) = [\hat{x}_{h_k}(i,j) - \Gamma \cdot std_{i,j}(h_k),$$
$$\hat{x}_{h_k}(i,j) + \Gamma \cdot std_{i,j}(h_k)],$$

where $\Gamma = 1 + \chi_{1-\alpha/2}$ is a threshold of the confidence interval. Then the inequality (8) is of the form

$$\hat{x}_{h_k}(i,j) \in D(k), \tag{9}$$

and we can conclude from (7) that as long as the inequality $h \leq h^*$ holds for $h = h_k$, $1 \leq k \leq r$, all the intervals $D(k)$, $1 \leq k \leq r$, have a point in common, namely, $x(i,j)$.

The following is the ICI statistic, which is used in order to tests the very existence of this common point and in order to obtain the adaptive window size value:

Consider the intersection of the intervals $D(k)$, $1 \leq k \leq r$, with increasing r, and let r^+ be the largest of those r for which the intervals $D(k)$, $1 \leq k \leq r$, have a point in common. This r^+ defines the adaptive window size and the adaptive mean estimate as follows

$$\hat{x}^+(i,j) = \hat{x}_{h_{r^+}}(i,j). \tag{10}$$

The following algorithm implements the procedure (10). Determine the sequence of the upper and lower bounds of the confidence intervals $D(j)$ as follows

$$D(k) = [L_k,\ U_k],$$
$$U_k = \hat{x}_{h_k}(i,j) + \Gamma \cdot std_{i,j}(h_k),$$
$$L_k = \hat{x}_{h_k}(i,j) - \Gamma \cdot std_{i,j}(h_k).$$

Let

$$\bar{L}_{k+1} = \max[\bar{L}_k,\ L_{k+1}], \tag{11}$$
$$\underline{U}_{k+1} = \min[\underline{U}_k,\ U_{k+1}],$$
$$k = 1, 2, ..., J,\ \bar{L}_1 = L_1,\ \underline{U}_1 = U_1$$

then the optimal window length comes for the largest r^+, for which the inequalities $\bar{L}_k \leq \underline{U}_k$, $k \leq r$, is still satisfied. This r^+ is the largest of those k for which the confidence intervals $D(k)$ have a point in common as discussed above. This ICI window size selection procedure requires knowledge of the estimate $\hat{x}_{h_k}(i,j)$ and its local variance only. The procedure described above is repeated for every pixel i, j.

4. ALGORITHMS AND EXPERIMENTS

The developed algorithm comprises two parts. The first part is used for an image segmentation. This segmentation assumes that the ICI rule is used for every pixel in order to find the adaptive sizes of four directional rectangular windows as shown in Figure 1. As a result every pixel can be an entry of many different estimates obtained for varying size windows

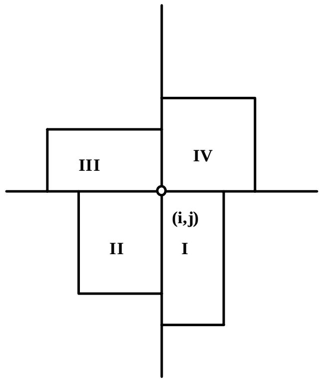

Figure 1: Four quadrant windows I,II,III and IV used for directional window size selection by the ICI rule.

with different centers. The second part assumes the DCT transform filtering for every of these adaptive size windows. All obtained estimates are accumulated in a buffer and averaged in order to produce the final estimate for every pixel.

Experiments were performed on the test images "House" and "Cameraman" (8 bit gray-scale 256×256 image) corrupted by different types of noise. The results are compared with the wavelet transform based (Haar, Symmlet, Coiflet, Translation Invariant) and Kuan filters. The new algorithm showed a valuable SNR improvement (more than 4-5 dB) for most cases. Some illustrative images are given in Figure 2. Figure 2a,b show the original and noisy image (additive Gaussian noise), while the DCT estimate described above is given in Figure 2d. The $RMSE$ values show a valuable original noise reduction. The visual quality is quite acceptable for this level of the noise. In Figure 2c we show as an intermediate results the filtering obtained from the LPA (5). The estimates obtained for four adaptive varying windows are averaged with the weights reciprocal to the variances of these estimates [10]. Figure 3 shows the varying adaptive window sizes obtained respectively for the windows I,II,III and IV (Figure 1, $h_k = 2^k$, $k = 0, 1, ..., 6$). Here black and white correspond respectively to small and large window sizes. Actually the adaptive window sizes delineate contours of the image and demonstrate a very reasonable performance of the ICI rule as a window size selector.

5. REFERENCES

[1] M. Lang, H. Guo, J.E. Odegard, C.S. Burrus, and R.O. Wells, "Nonlinear processing of a shift invariant DWT for noise reduction," *Proc. of SPIE*, vol. 2491, 1995, pp. 640-651.

[2] K. Egiazarian, M.Helsingius, P.Kuosmanen, and J.Astola, "Removal of blocking and ringing artifacts in transform domain de-noising," *ISCAS'99*, Orlando, Florida, May 30 - June 2, 1999.

[3] P. Moulin, "A wavelet regularization method for diffuse radar-target imaging and speckle-noise reduction," *J. of Math. Imaging and Vision*, vol. 3, No. 1, 1993, pp. 123-134.

[4] L. Yarovslavsky, "Local Adaptive Image Restoration and Enhancement with the Use of DFT and DCT in a Running Window," *Proceedings, Wavelet Applications in Signal and Image Processing IV*, SPIE Proc. Series, v. 2825, pp. 1-13, 6-9 August 1996, Denver, Colorado.

[5] D.L. Donoho, "De-noising by soft thresholding," *IEEE Trans. Inform. Theory*, vol. 41, 1995, pp. 613-627.

[6] R.R. Coifman and D.L. Donoho, "Translation-invariant de-noising," in *Wavelets and statistics*, pp. 125-150, Springer-Verlag, 1995.

[7] R. Oktem, L. Yarovslavsky, K. Egiazarian, "Signal and image de-noising in transform domain and wavelet shrinkage: A comparative study," *EUSIPCO'98*, Island of Rhodes, Greece, 8-11 Sept. 1998, v. 4, pp. 2269-2272.

[8] Katkovnik V. "On adaptive local polynomial approximation with varying bandwidth," *Proceedings of IEEE Int Conference of Acoustics, Speech & Signal Processing (ICASSP'98)*, v.4, pp. 2321-2324, Seattle, Washington, May 12-15, 1998, USA.

[9] Katkovnik V. "A new method for varying adaptive bandwidth selection," *IEEE Trans. on Signal Processing*, vol.47, $N^0$9, pp. 2567-2571, 1999.

[10] V. Katkovnik, H. Oktem, and K. Egiazarian, "Filtering heavy noised images using ICI rule for adaptive varying bandwidth selection," *ISCAS'99*, Orlando, Florida, May 30 - June 2, 1999.

[11] J.M. Park, W.J. Song and W.A. Pearlman "Speckle filtering of SAR images based on adaptive windowing," *IEE Proc.-Vis. Image Signal Processing*, vol. 146, $N^0$4, pp. 191-197, 1999.

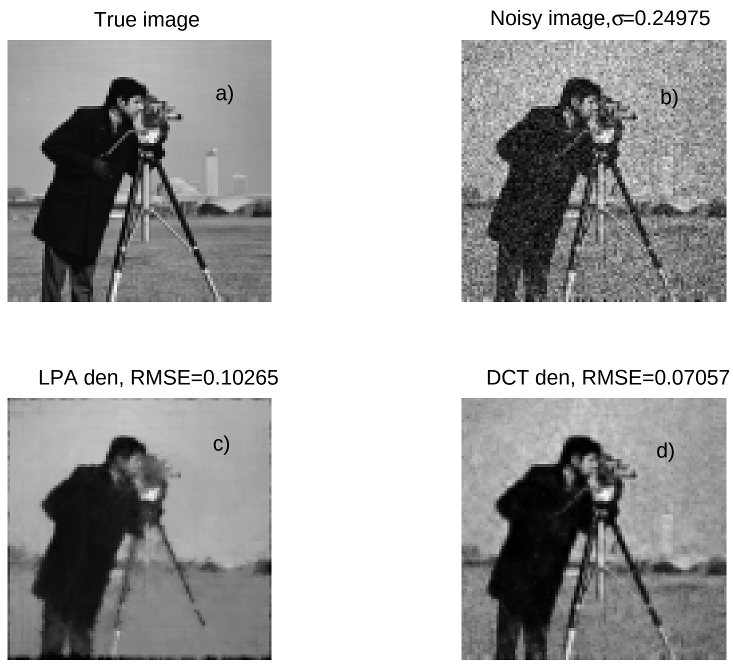

Figure 2: a) True image, b) Noisy image, c) LPA denoising, d) DCT denoisng with ICI adaptive window sizes.

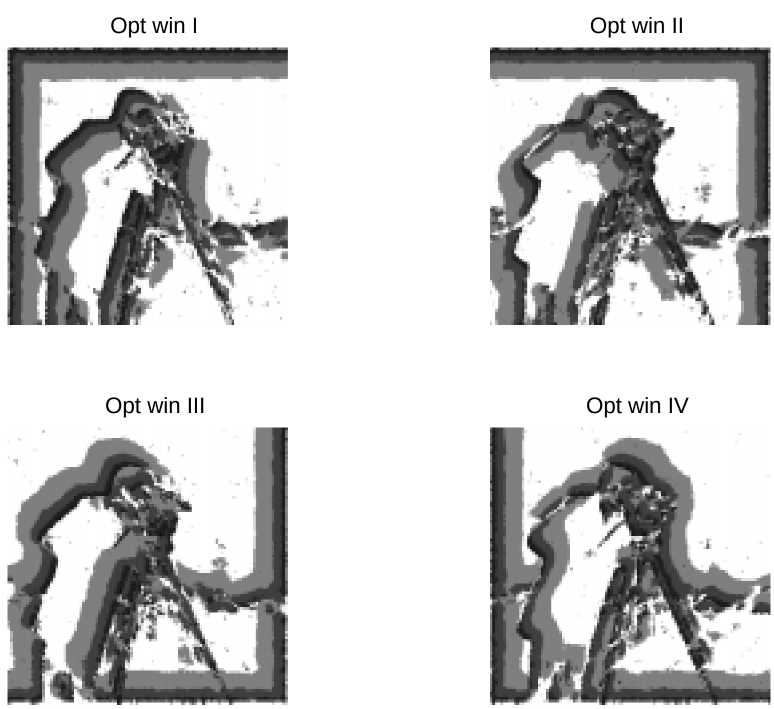

Figure 3: Adaptive window sizes obtained by ICI with Γ=2.5.