Local transform-based image de-noising with adaptive window size selection

Karen Egiazarian*, Vladimir Katkovnik and Jaakko Astola
Tampere University of Technology, Signal Processing Lab.,
FIN-33101, P.O. Box 553, Tampere, Finland

ABSTRACT

An algorithm for image noise-removal based on local adaptive filtering is proposed in this paper. Three features to use into the local transform-domain filtering are suggested. First, filtering is performed on images corrupted not only by an additive white noise, but also by image-dependent (e.g. film-grain noise) or multiplicative noises. Second, a number of transforms is used instead of the single one, the resulting estimate is a linear combination of estimates from each of the transforms using local statistics. Third, these transforms are equipped with a varying adaptive window size for which we use the so-called intersection of confidence intervals (ICI) rule. Finally, we combine all the estimates for a pixel from neighboring windows by weighted averaging them. Comparison of the algorithm with known techniques for noise removal from images shows the advantage of the new approach, both quantitatively and visually.

Keywords: Image de-noising, local adaptive filters, window size selection.

1 INTRODUCTION

Transform domain signal denoising finds applications in restoration of different type of one and two dimensional signals such as medical signals, SAR and astronomical images. Depending on imaging systems different noise models were considered - starting from the additive white noise, data-dependent (e.g. film-grain type) noise, to multiplicative noise.

Image processing in a transform domain rather than in a 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 a spatial domain into a transform domain is especially useful if it is applied locally rather than globally.

Having an excellent performance in suppressing of 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 the blocking and ringing artifacts from images of JPEG and wavelet decoded images .[2,1,3]

*Correspondence: E-mail: karen@cs.tut.fi, Telephone: +358-3-365-3860, Fax: +358-3-365-3817.

Local adaptive filters [4] work in the domain of an orthogonal transform in a moving window and non-linearly modify the transform coefficients in order to obtain an estimate of the central pixel. Nonlinear filtering in the wavelet transform domain were introduced in terms of wavelet denoising by Donoho and Johnstone [5,6] and has been extended by several authors. In [7,1] translation invariant wavelet denoising algorithms were introduced and tested on different one dimensional signals and SAR images, respectively. In [8] wavelet transform domain denoising was combined with the empirical Wiener filtering for a better performance.

In [9] 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 window gives an estimate of the overall subimage within the window and not only at the central pixel. 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.

This sort of filters will be used in this paper as a starting "prototype" filter. We are going to equip it by several additional functions:

F1) It works on images corrupted not only by an additive white noise, but also on those with an additive image-dependent noise, as well as with a multiplicative noise.

F2) Several transforms in parallel are in use (multichannel filtering), the results of filtering from these channels are combined in order to get a final estimate.

F3) Transforms are used with a varying adaptive window size for which the intersection of confidence intervals (ICI) rule is applied.[10,11]

Extensive experiments confirm the expected good performance of the proposed filters in different noise models.

The rest of the paper is organized as follows. In Section 2, the noise model is described and the "prototype" transform domain denoising technique is presented. In Section 3, this algorithm is extended to get above-defined functions F1-F3. The experimental results and performance analysis of the proposed filter for restoration of images corrupted with different types of noises, are presented.

2 FILTERING IN TRANSFORM DOMAIN

2.1 Noise Model

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

$$Y(i,j) = X(i,j) + Q(i,j) = X(i,j) + X(i,j)^{\gamma} n(i,j), \tag{1}$$

where $X(i,j)$ is the noise-free image and $n(i,j)$ is a noise with mean 0 and 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, $\gamma = 1$.

2.2 Filtering in transform domain

Describing the observed image in a domain of an orthogonal transform, we obtain

$$\tilde{Y}(i,j) = \tilde{X}(i,j) + \tilde{Q}(i,j), \tag{2}$$

where $\tilde{X}(i,j)$ and $\tilde{Q}(i,j)$ are noise-free image and noisy term from the equation (1), respectively, in the transform domain.

The optimum filter coefficients $\eta(i,j)$ can be found by minimizing the expected value of the quadratic error between an estimated output in the transform domain $\hat{\tilde{X}}(i,j)$ and the transform of the noise-free image $\tilde{X}(i,j)$:

$$\eta(i,j) = \arg\min_{\eta} E\{\| \tilde{X} - \hat{\tilde{X}} \|^2\},$$

so, that $\hat{\tilde{X}}(i,j) = \eta(i,j)\tilde{Y}(i,j)$, where $\tilde{Y}(i,j)$ is the transform of the observed image within the transform mask starting from the indices (i,j).

From this one can get

$$\eta(i,j) = \frac{E\{\tilde{X}(i,j)\tilde{Y}(i,j)\}}{E\{\tilde{Y}^2(i,j)\}}, \tag{3}$$

Using (1) the assumption that the transform size is small enough and $X(i,j)$ is smoothly varying (so that the subimage inside the transform window can be approximated by a constant, say, by the sample mean $m(i,j)$ of this subimage) we obtain[12]

$$\eta(i,j) = \frac{E\{\tilde{Y}^2(i,j)\} - \sigma_n^2 m^{2\gamma}(i,j)}{E\{\tilde{Y}^2(i,j)\}}. \tag{4}$$

Thus, the filtering procedure can be done by the so-called rejective filter,[12] defined by

$$\eta(i,j) = \begin{cases} 1, \text{ if } \tilde{Y}^2(i,j) \geq \sigma_n^2 m^{2\gamma}(i,j) \\ 0, \text{ otherwise} \end{cases}. \tag{5}$$

The noise reduction (de-noising) method by nonlinear thresholding in the wavelet transform domain proposed by Donoho and Johnstone,[5,7] consists of the following three steps:

1. Transform the noisy image $\mathbf{Y}$, applying a two-dimensional orthogonal transform, $\hat{\mathbf{X}} = \mathbf{W}\tilde{Y}$, $\tilde{Y} = \mathbf{W}^T\mathbf{Y}$,where $\mathbf{W}$ is the transform matrix, ^{T} stays for the transpose.

2. Apply thresholding to $\tilde{Y}(i,j)$ (the elements of the matrix $\tilde{Y}$) which results in suppression of the coefficients $\tilde{Y}(i,j)$ of a lower energy, using hard thresholding:

$$\hat{\mathbf{Z}}(i,j) = T_h(\tilde{Y}(i,j),t) = \begin{cases} \tilde{Y}(i,j), \text{ if } |\tilde{Y}(i,j)| \geq t, \\ 0, \text{ otherwise} \end{cases},$$

or soft thresholding:

$$\hat{\mathbf{Z}}(i,j) = T_s(\tilde{Y}(i,j),t) = \begin{cases} sign(\tilde{Y}(i,j))(|\tilde{Y}(i,j)| - t), \text{ if } |\tilde{Y}(i,j)| \geq t, \\ 0, \text{ otherwise} \end{cases}.$$

3. Transform back to the original domain, performing the inverse transform: $\hat{\mathbf{X}} = \mathbf{W}\hat{\mathbf{Z}}$.

It can be seen that the hard threshold operation coincides with the rejective filter (5) when $\gamma = 0$ (which is the case of the white additive noise corrupted image).

However, applying this classical wavelet de-noising scheme in practice, one may end up with some artifacts near singularities (the pseudo-Gibbs phenomena in a neighborhood of a discontinuity). One way to overcome this problem is to use undecimated (or shift-invariant) wavelet transforms,[7] which can be done according to the following simple strategy: wavelet de-noising is applied to all circular shifts of a signal and the average of all these results is calculated as a final estimate.

2.3 Local transform based de-noising

This method is basically similar to the one described above. The main difference is that the filtering procedure is applied not to the transform coefficients of entire image but to blocks of the image in sliding (running, overlapping blocks) filtering fashion. Here we distinguish two approaches. First one, the local transform domain filter [4] is done with a sliding window over the image, and the central pixel of the inverse transform of the filtered fragment is kept as the estimated value of the corresponding pixel.

It is shown in [9] that keeping all filtered outputs for each window improves the performance of de-noising. For a final estimate of a pixel $\hat{\mathbf{X}}(i,j)$ we combine all estimates of this pixel coming from different windows it belongs to. It defines the second approach.

Local transform based filters have certain advantages. They can use several transforms by changing one basis to another while moving from one window to another one. They can also use the varying window size.[13] Due to these reasons we have chosen these filters as a "prototype" in our new scheme.

3 WINDOW SIZE AND TRANSFORM SELECTION

3.1 The ICI rule for window size selection

Let us explain the idea of the intersection of confidence interval (ICI) rule on example of the simple sample mean estimate of one dimensional signal.[10,11]

Let for the observation model (1) $\hat{X}_N(i) = \frac{1}{2N+1}\sum_{u=-N}^{N} Y(i+u)$ be the estimate of the signal at the point i . Then, the estimation error is $e_i = X(i) - \hat{X}_N(i)$. It can be shown that the estimation bias is $|E\{e_i\}| \leq A\ |X^{(2)}(i)|(2N+1)^2$, where $X^{(2)}(i)$ is the second derivative of the signal X and A is a constant, and the standard deviation of the error is $std_i = \sqrt{\frac{B^2}{2N+1}\sigma^2}$, $B^2 = \frac{1}{2N+1}\sum_{u=-N}^{N} |X(i+u)|^{2\gamma}$.

Let us denote the upper bound of the bias as $\omega_i = A\ |X^{(2)}(i)|(2N+1)^2$. Then we have for the mean squared error (MSE):

$$MSE \leq \omega_i^2 + std_i^2 = A^2|X^{(2)}(i)|^2(2N+1)^4 + \frac{B^2}{2N+1}\sigma^2.$$

Let us assume that the mean B does not depend on N. In particular for a homogeneous area $B \simeq |X(i)|^{\gamma}$. Then, the minimum of MSE on N is achieved at $N = N^*$,

$$2N^* + 1 \simeq \left(\frac{B^2\sigma^2}{4A^2|X^{(2)}(i)|^2}\right)^{1/5}.$$

It can be verified that the ratio of the bias to the standard deviation of the error is constant for $N = N^*$:

$$\omega_i / std_i = \lambda,\ \lambda = 1/2.$$

This N^* provides the optimal balance between the bias and the random error of smoothing.

It can be seen also that

$$\omega_i \leq \lambda \cdot std_i, \text{ for } N \leq N^*. \tag{6}$$

Then $|e_i| = \left|X(i) - \hat{X}_N(i)\right| \leq \omega_i + |\xi_i|$, where asymptotically the random term ξ_i is Gaussian and with the probability $p = 1 - \alpha$ the following inequality holds

$$|e_i| \leq \omega_i + \chi_{1-\alpha/2} std_i, \tag{7}$$

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:

$$H = \{N_1 < N_2 < .... < N_J\},$$

starting with quite small N_1.

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

$$|e_i| \leq (\lambda + \chi_{1-\alpha/2}) std_i. \tag{8}$$

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

$$D(j) = [\hat{x}_{N_j}(i) - \Gamma \cdot std_i(N_j),\ \hat{x}_{N_j}(i) + \Gamma \cdot std_i(N_j)], \tag{9}$$

where

$$\Gamma = \lambda + \chi_{1-\alpha/2} \tag{10}$$

is the threshold of the confidence interval. Then for $N = N_j$ the inequality (8) is of the form

$$X(i) \in D(j), \tag{11}$$

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

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

$$\hat{X}^+(i) = \hat{X}_{N^+}(i),\ N^+ = N_{k^+}. \tag{12}$$

The following algorithm implements the procedure (12). Determine the sequence of the upper and lower bounds of the confidence intervals $D(j)$ as follows

$$\begin{aligned} D(j) &= [L_j,\ U_j], \\ U_j &= \hat{X}_{N_j}(i) + \Gamma \cdot std_i(N_j), \\ L_j &= \hat{X}_{N_j}(i) - \Gamma \cdot std_i(N_j). \end{aligned} \tag{13}$$

Let

$$\begin{aligned} \bar{L}_{j+1} &= \max[\bar{L}_j,\ L_{j+1}],\ \underline{U}_{j+1} = \min[\underline{U}_j,\ U_{j+1}], \\ j &= 1,2,...,J,\ \bar{L}_1 = L_1,\ \underline{U}_1 = U_1 \end{aligned} \tag{14}$$

then the optimal window length N^+ comes for the largest j, k^+, for which the inequality

$$\bar{L}_j \le \underline{U}_j \tag{15}$$

is still satisfied. This k^+ is the largest of those j for which the confidence intervals $D(j)$ have a point in common as discussed above. This ICI window size selection procedure requires knowledge of the estimate $\hat{X}_{N_j}(i)$ and its local variance only. The procedure described above is repeated for every i. Note that for the two dimensional (image) case the algorithm is different only in $\lambda = 1/\sqrt{2}$.

Let us illustrate how the ICI rule works for the multiplicative noise. We consider the step-wise signal $X(t) = 1$ for $0 \le t \le 1/2$ and $X(t) = 2$ for $1/2 \le t \le 1$, with multiplicative noise $\sigma = 0.1$ and $\gamma = 1$. Number of observation is equal to 256. This step-wise signal is simplest to demonstrate the ability of the algorithm to estimate a jump-wise singularity in a signal. In our algorithm implementation we assume that $std_i(N_j) = |\hat{X}_{N_j}(i)|\sigma/\sqrt{2N_j + 1}$. These experiments present qualitative results demonstrating what sort of improvements can be expected from the algorithm based on the ICI rule.

The observations and the true signal are shown in Figure 1a,b. Curves in Figure 1c, e, f provide the adaptive varying values N^+ obtained for the left, right and symmetric mean estimates by the ICI rule. We wish to emphasize that the results given by these curves are in the accurate agreement with the window sizes that could be selected provided that we knew in advance that the signal is piece-wise constant with a single jump and the location of this jump is known exactly. In particular, Figure 1c shows the adaptive N^+ for the left mean estimate. In the first part of this curve, $0 \le t \le 1/2$, N^+ is a linear increasing function of t. This happens because the left mean uses for estimation only the observations which are on the left-hand side from the estimation point t and all these observations are used for estimation of the constant signal value $X = 1$. Immediately after $t = 1/2$ the value of the window size drops as the signal value is changed to $X = 2$ and only a small number of the observations can be used for this estimates for t close to $1/2$. While t is increasing towards the value 1 the window size N^+ is also increasing as a larger number of observations can be used for estimation of $X = 2$. The left mean estimate given into Figure 1d corresponds to the window sizes presented in Figure 1c. We can see that the estimate delineates accurately the jump in the signal from its left-hand side perfectly suppressing the noise. Naturally the picture is not so perfect on the right-hand side of the jump because only small number of the observations can be used by the left mean for estimation of $X(t)$ which are nearly the point $t = 1/2$. A similar behavior can be seen for the right mean estimate in Figure 1f and the corresponding N^+ Figure 1e. Figure 1g,h are given for the adaptive window size and estimates by the symmetric mean estimator. Figure 2b,c show the combined estimates obtained by fusing two left-right ($L\&R$) and three left-right-symmetric ($L\&R\&S$) estimates. The combined estimate gives a nearly perfect reconstruction of the jump. Figure 2d shows the ideal invariant window size symmetric estimate with the best invariant window size obtained provided that the signal is known. A comparison of estimates shows a clear advantage of the combined estimates with respect to the non-combined adaptive estimates as well as to the ideal invariant window size estimate.

Concerning the accuracy of estimation we wish to note that the ICI rule has a design parameter Γ which actually controls the smoothness of the adaptive window size given by the ICI rule. In general, larger values of Γ tend to result in a smoother N^+ as a function of t, while smaller values of Γ result in a larger variability in N^+. We use in Figures 1 and 2 quite large values $\Gamma = 3.5$ in order to expose a basic tendencies in the varying adaptive window size. However, the $RMSE$ accuracy improvement usually requires smaller values of Γ and results in a quite strong irregular component in N^+ (see[10,11]). For image denoising with an additive noise, $\gamma = 0$, a similar algorithm was developed in.[13]

Finally we wish to emphasize that the ICI rule compares the estimate of the signal versus the standard deviation of this estimate and in this way works similar to the usual for SAR image homogeneity analysis methods (see[14]) but based on the different statistical rule.

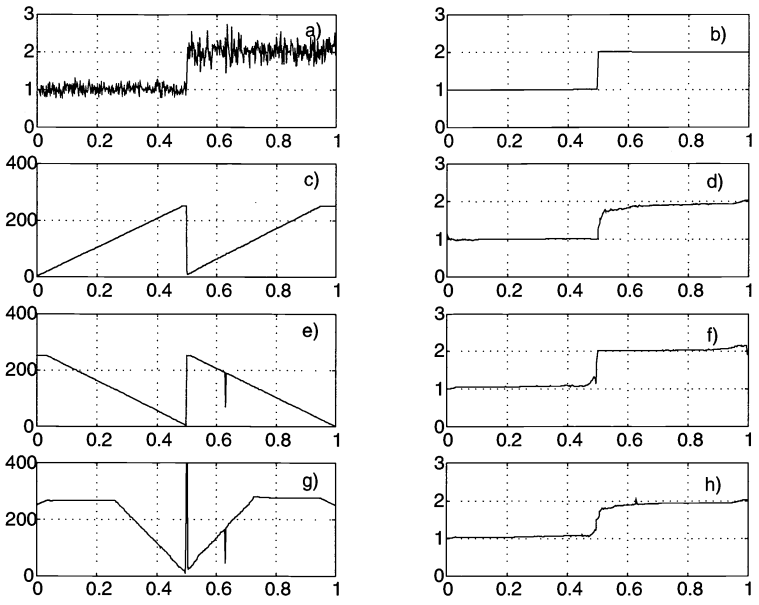

Figure 1: Performance of the mean sliding window estimates with ICI varying window size: a) Observations, b) True signal, c)-d) Adaptive window size and estimate by the left windowed mean, e)-f) Adaptive window size and estimate by the right windowed mean, g)-h) Adaptive window size and estimate by the symmetric windowed mean.

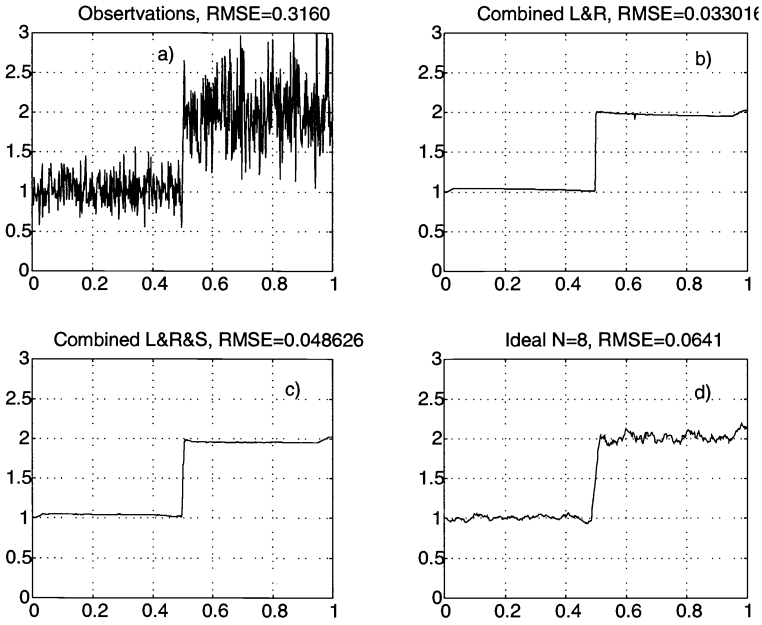

Figure 2: Performance of the combined sliding window estimates with ICI varying window sizes: a) Observations, b) Left-right combined estimate, c) Left-right-symmetric combined estimate, d) The symmetric window sliding mean with invariant ideal window size.

3.2 Multi-base local transform de-noising

In the previous section we gave a method to set a size of transform for a local transform-based filtering. Now we face a problem: how to choose a proper transform (or transforms)? One should think about a transform which will have an efficient computational algorithm. This problem is somehow similar to the selection of "best basis" among the library of possible bases in the wavelet packet construction[15] . Nevertheless, the local processing is a more specific procedure, the goal is not to find "globally" the best transform (as in the best wavelet packet basis selection[15]), but the "best transform" for each image block.

One could separate the whole image into statistically similar groups of subimages, and find the Karhunen-Loeve transforms (KLTs) corresponding to these groups. This would be quite a time consuming procedure, requiring preliminary analysis of the image for a proper clustering. Moreover, a direct implementation of these KLTs would be necessary because fast computational algorithms for KLTs do not exist in general.

One could still use the "best basis" algorithm[15] and apply it to each window. However, a selection among transforms having different features and which cannot be combined in the library of bases for the "sub-optimal" search (like the Discrete Cosine Transform and the Discrete Haar bases) may give better result.

Here we use k transforms in parallel and take a weighted average of the results from each channel. As a simple criterium for the choice of the preferable transform we have chosen the number ν of remaining non-zero transform coefficients after thresholding the spectra. Thus, the overall estimate $\hat{\mathbf{x}}$ for the image block is formed from the corresponding estimates $\hat{\mathbf{x}}^{(j)}$ from each channel $j = 1, ..., k$ by

$$\hat{\mathbf{x}} = \frac{\sum_{j=1}^{k} \frac{\hat{\mathbf{x}}^{(j)}}{\nu_j}}{\sum_{j=1}^{k} \frac{1}{\nu_j}}, \tag{16}$$

where ν_j shows the number of remaining coefficients in each channel after making thresholding in transform domains.

Let us show the performance of the algorithm on the examples of de-noising of signals and images corrupted by different types of noises from the model (1). In our experiments we have been using some test signals and routines from D. Donoho's Matlab software package "WaveLab".[16] Numerical performances have been estimated by the signal-to-noise ratio (SNR).

We start our experiments from one-dimensional signals corrupted by the Gaussian noise. As the test signals we have chosen a signal of the length 4096 which is a concatenation of two signals: "Blocks" and "Doppler" (from[16]). The noise was added (additive Gaussian noise) using routine "NoiseMaker"[16] with desired SNR $\rho = 2$. The test signal was selected with the purpose to check a performance of multi-base local transform methods.

The following methods of de-noising have been compared in our experiments:

- wavelet de-noising (shrinkage)[5] based on the Haar, Symmlet- and Coiflet wavelets,

- fully translation-invariant wavelet de-noising[7] using the stationary wavelet transforms, using the Haar (TI Haar), Symmlet (TI Sym4) and Coiflet (TI Coif3) wavelets.

Settings for our algorithm were: 3 different window sizes: N_1 =512, N_2 =1024 and N_3 =2048 (and selection between them using ICI rule at each moving window position), 3 different transforms (TI Haar, TI Sym4 and TI Coif3) of each above mentioned size were performed. Overall estimate for each window location is computed by (16).

The results of the TI de-noising schemes and the new algorithm, called local TI (Loc.TI) are much better (more than 4-5 dB) than those of the classical wavelet denoising.

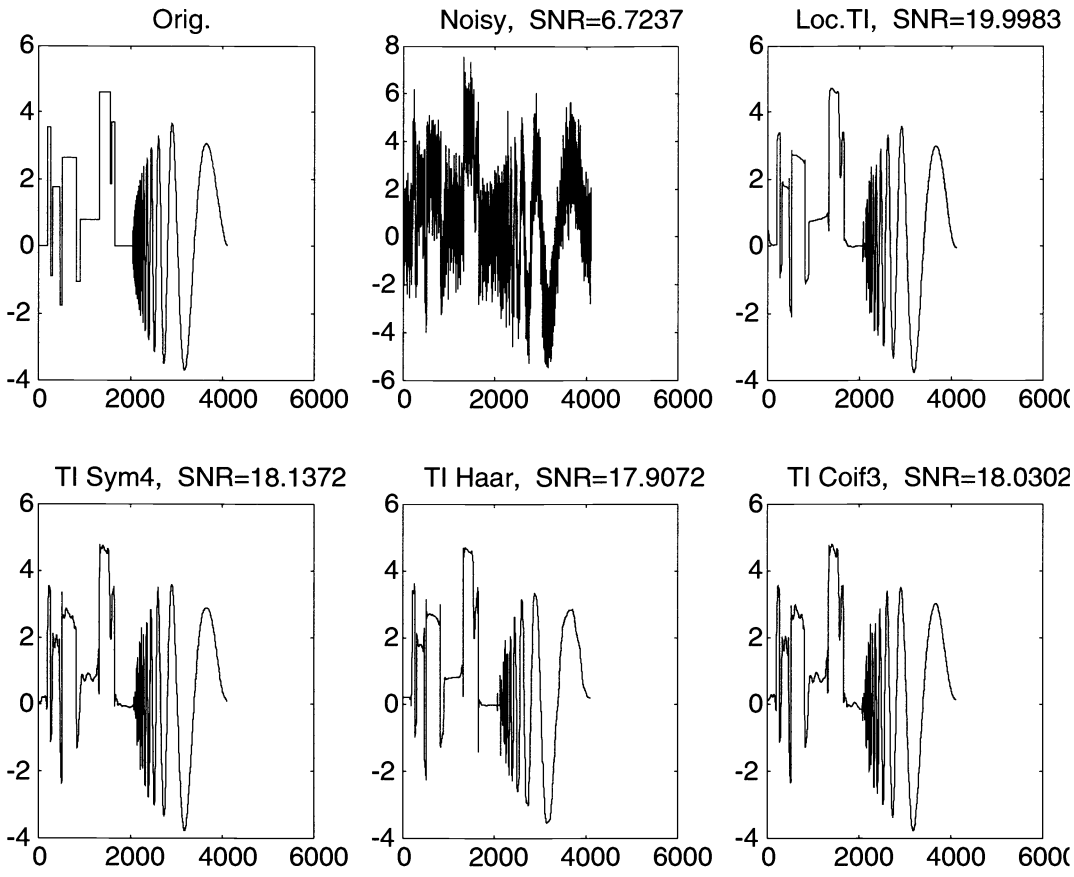

Figure 3: Composite signal de-noising: original, noisy (normalized Gaussian noise, SNR=2), and de-noised signals

In Figure 3 we show for illustration the results of filtering comparing with the best wavelet techniques.

Next set of experiments were performed on the test image "House" (8 bit gray-scale 256×256 image) corrupted by different types of noises. The results are tabulated in Table 1. Figure 4 shows the noise-free, noisy, filtered by the Kuan and by the new adaptive filters.

CONCLUSION

New local adaptive transform based de-noising technique for removing additive and multiplicative noises has been proposed and studied in this paper. Filtering is based a number of transforms used instead of the single one. These transforms are equipped with a varying adaptive window size for which we use the ICI rule. Finally, we combine all the estimates for a pixel from neighboring windows by weighted averaging them in order to produce the final estimate. Comparison of the algorithm with known techniques for noise removal from images shows the advantages of the new approach, both quantitatively and visually.

Table 1: Performance of the Kuan filter and the proposed filter for different corrupted versions of image "House" (Gaussian noise with standard deviation $\sigma = 15$, "Film-grain" noise with $\sigma = 3, \gamma = 0.5$, and speckle noises with $\sigma = 0.4$ and $\sigma = 0.7$)

Image "House": / SNR	Gaussian	Film-grain	Speckle1	Speckle2
Noisy	24.6	23.44	25.03	20.22
Kuan (7×7)	35.27	34.66	35.42	33.16
New method	37.95	36.67	38.15	34.87

4 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] D.L. Donoho, I.M. Johnstone, "Ideal spatial adaptation by Wavelet Shrinkage", *Biometrica*, 81(3), pp. 425-455, 1994.

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

[8] S. P. Ghael, A. M. Sayeed and R. G. Baraniuk, "Improved Wavelet Denoising via Empirical Wiener Filtering", *SPIE Proc. Series,* vol. 3169, pp. *389-399, 1997.*

[9] 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.

[10] 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.

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

[12] R. Oktem, L. Yaroslavsky, K. Egiazarian, and J. Astola, "Transform Domain Approaches to Image Denoising", submitted to *Journal of Electronic Imaging*, 2000.

[13] 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.

[14] 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.

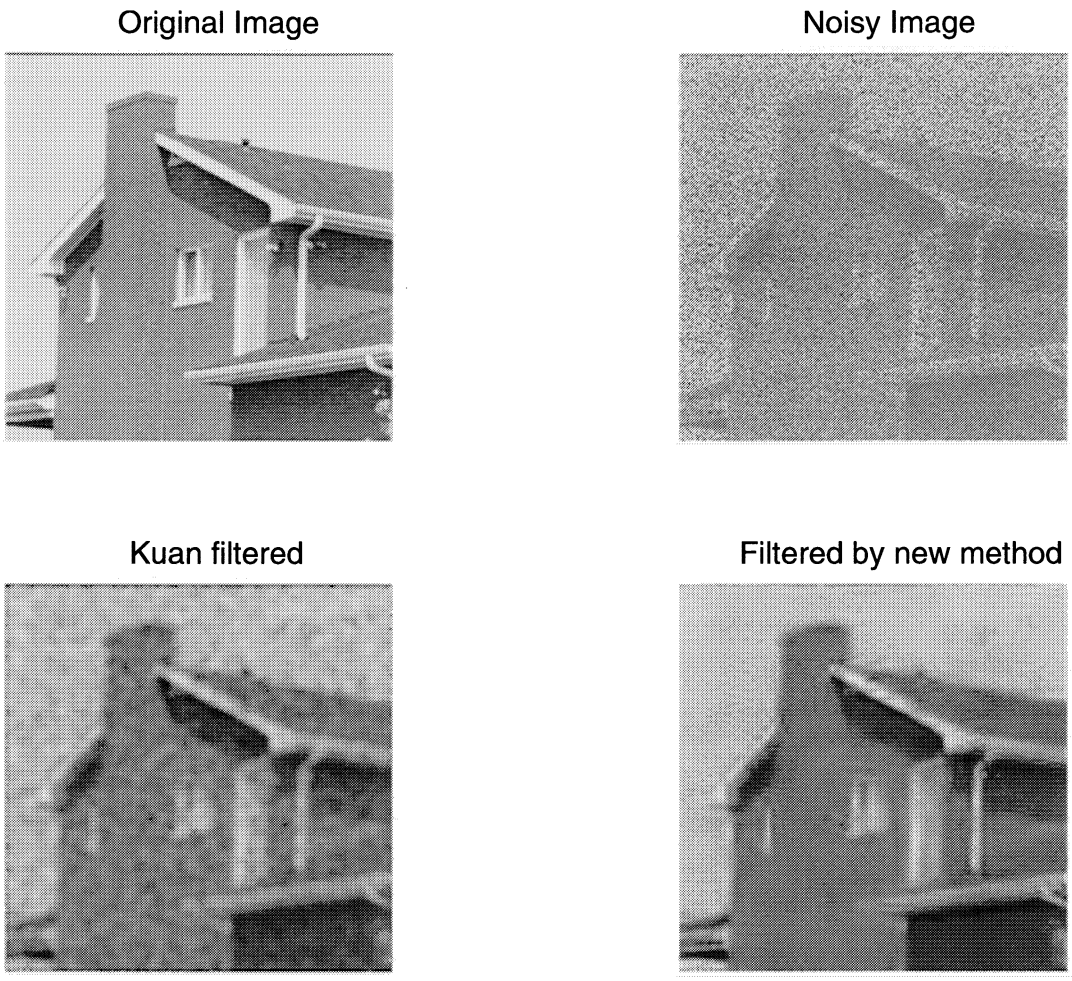

Figure 4: "House" images: original, noisy (speckle noise, standard deviation 0.7), and de-noised Kuan and new adaptive fiter

[15] Wickerhauser, M. V. *Adapted Wavelet Analysis from Theory to Software*, IEEE Press, A.K.Peters, Wellesley, Massachusetts, 1994.

[16] WaveLab web page (WaveLab .701) http://playfair.stanford.edu/ wavelab/.