LOCAL ADAPTIVE TRANSFORM BASED IMAGE DE-NOISING WITH VARYING WINDOW SIZE

Hakan Öktem (1), Vladimir Katkovnik (2), Karen Egiazarian (1) and Jaakko Astola (1)

1) Signal Processing Laboratory, Tampere Universty of Technology,
PO Box 553, Tampere, Finland
e-mails: hakan.oktem@tut.fi, karen@cs.tut.fi, jta@cs.tut.fi
2) Department of Mechatronics, Kwangju Institute of Science and Technology
Oryong-dong, Puk-gu, Kwangju 500-712 Republic of Korea, e-mail: katkov@kjist.ac.kr

ABSTRACT

Local adaptive image de-noising in transform domain is a powerfull tool for adapting to unknown smoothness of the images. In this work we propose to perform local adaptive denoising with adaptively varying local transform support size rather than using a transform with fixed size. We use a special rule (Intersection of Confidence Intervals - ICI) to select the optimum window sizes locally. The algorithm provides significant improvements in the de-noising performance.

1. INTRODUCTION

Signal and 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 the spatial domain into the transform domain is especially useful if it is applied locally rather than globally.

Local adaptive filters [2] 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 at the central pixel. Nonlinear filtering in the wavelet transform domain were introduced in terms of wavelet denoising by Donoho and Johnstone [3, 4] and has been extended by several authors. In [1, 5] translation invariant wavelet denoising algorithms were introduced and tested on different one dimensional signals and SAR images, respectively. In [6] wavelet transform domain denoising was combined with the empirical Wiener filtering for a better performance.

In [7] the local average transform domain denoising was presented. The difference between this filter and the one in [2] 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 the multiple estimates are obtained for each pixel. All of the above estimates are averaged in order to obtain the final estimate for these pixels. In this work we will use this filter as a prototype filter and add some additional features to implement within varying size locally optimum windows.

In [8] another local adaptive filter fitting the samples within a local window to a local polynomial was introduced.

When a locally applied transform is used the performance highly depends on key parameters, especially the window (transform support) size. When a local signal model is considered the neighborhood within which the model is valid has a crucial importance. In this paper we will investigate the performance of the filters by selecting the locally optimal window size for each location rather than a global optimum.

2. FILTERING IN TRANSFORM DOMAIN

A noise model for images is considered as:

$$y(i,j) = x(i,j) + e(i,j), \qquad (1)$$

where $\mathsf{x} = (x(i,j))$ is a matrix of a noise-free image and $\mathsf{e} = (e(i,j))$ is a Gaussian noise, $e(i,j) \sim N(0, \sigma_n^2)$ with zero mean and variance σ_n^2.

In an obvious vector-matrix notation both formulas have a form

$$\mathsf{y} = \mathsf{x} + \mathsf{e}, \qquad (2)$$

where the corresponding y, x, e are matrices.

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

Let the orthogonal transform be defined by a $N \times N$ matrix H. The observation model (2) in a transform domain

has a form

$$Y = HyH^T, X = HxH^T, E = HeH^T$$

where Y, X are observed and noise-free images and E is noise in the transform domain.

The optimum Wiener filter coefficient η can be found by minimizing the expected value of the quadratic error between a spectrum $\tilde{X} = Y \odot \eta$, where $\odot$ stays for a element-by-element product, of estimated output and the spectrum X of the noise-free image:

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

From this one can get an estimate of the Wiener filter coefficients:

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

Using (2) we obtain [9]

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

Thus, the filtering procedure can be done by the so-called rejective filter [9], defined by

$$\hat{\eta}(i,j) = \begin{cases} 1, & \text{if } |Y(i,j)| \geq \sigma_n \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

It was shown in [9] that this filter can be adapted to different noise distributions (e.g. to multiplicative noise) by adjusting the rejection rule (threshold). For a Gaussian noise the rule (5) is quite similar to the standard wavelet shrinkage.

2.1. Local Versus Global Filtering in Transform Domain

Before doing a filtering by (5) we need to compute the spectrum Y of the observed signal y. How to do it - locally or globally? One thing is clear that the basis functions of the transform should be compactly supported in order to achieve a reasonable decorrelation of the observed image. There are two alternative solutions reported in literature: apply a wavelet transform or a local sliding window transform.

In the first case we make a fit of the image by decomposing it into a linear combination of the scaled and shifted wavelet basis functions known as the "mother" wavelets. In second case, we also fit an image by locally supported and shifted basis functions which could have supports of different scales. Thus, the only difference between these two cases is in the used basis functions. It deserves to be mentioned that the local functions used in the latter case can be interpreted as the wavelet frames.

In the case of the wavelet transforms the overall filtering procedure [3, 5], consists of the following three steps:

1. Compute a transform of noisy observations: $Y = HyH^T$.
2. Apply a filtering procedure:

$$\tilde{X}(i,j) = \hat{\eta}(i,j)Y(i,j) \tag{6}$$

3. Transform back to the original domain:

$$b = H^T \tilde{X} H$$

For the local sliding window transform based filtering we perform the same 3 steps as described above but for every sliding window location k, l, obtaining the matrix of the local estimates: $b = [b_{k,l}]$. It is emphasized that the sizes of the matrix b are uniquely defined by the sizes of the windows used for the estimation. As the sliding windows are overlapped, we have multiple estimates for every k or k, l. The simplest way to deal with these multiple estimates and obtain the final result is to calculate the sample average over the all available overlapped estimates for every i, j.

3. THE *ICI* RULE FOR WINDOW SIZE SELECTION

The following is the *ICI* statistic, which is used in order to test the very existence of this common point and in order to obtain the adaptive window size value. Since the transform we are using is separable, from the simplicity point of view we just describe the ICI rule for 1-D case.

Consider the intervals $D(j)$, $j = 1, 2, ...,$, and let k^+ be the largest value of 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 estimate as follows

$$\hat{x}^+(i) = \hat{X}_{N^+}(i), \quad N^+ = N_{k^+}. \tag{7}$$

The following algorithm implements the procedure (7). Determine the sequence of the upper and lower bounds of the confidence intervals $D(j)$ as follows

$$D(j) = [L_j, U_j], \tag{8}$$
$$U_j = \hat{x}_{N_j}(i) + \Gamma \cdot std_i(N_j),$$
$$L_j = \hat{x}_{N_j}(i) - \Gamma \cdot std_i(N_j).$$

Let

$$\bar{L}_{j+1} = \max[\bar{L}_j, L_{j+1}], \quad \underline{U}_{j+1} = \min[\underline{U}_j, U_{j+1}], \tag{9}$$
$$j = 1, 2, ..., J, \quad \bar{L}_1 = L_1, \quad \underline{U}_1 = U_1$$

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

$$\bar{L}_j \leq \underline{U}_j \tag{10}$$

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 ICI rule is graphically illustrated into Figure 1, where the arrows show the successive intersections of the confidence intervals (1,2), (1,2,3), and (1,2,3,4). Assuming that the intersection with the forth confidence interval (corresponding $h = h_4$) is empty we obtain the adaptive window size $h^+ = h_3$.
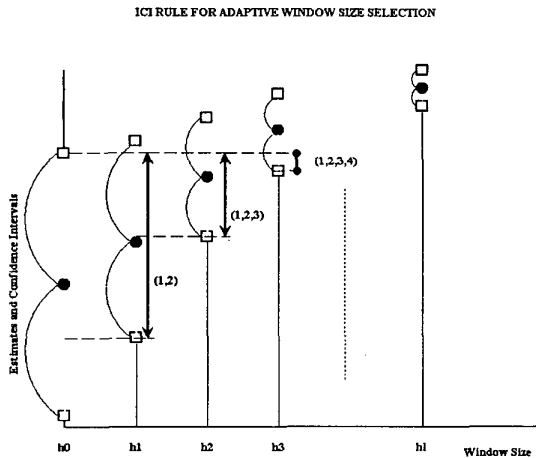
ICI RULE FOR ADAPTIVE WINDOW SIZE SELECTION

Figure 1: Graphical illustration of ICI rule.

3.1. DCT Based Local Transform Domain De-noising with ICI Rule Adaptive Window

When discussing transform domain denoising we have mentioned that we were using the decorrelation property of some transforms. In particular, the Discrete Cosine Transform (DCT) is a good approximation of the ideal Karhunen Loeve Transform (KLT) for highly correlated data [10]. When we treat a signal as composed of more or less correlated portions whose sizes are not known in advance (as is usually the case) the selection of the transform support size becomes important. The de-noising performance increase with the increasing local transform size due to improved spectral resolution if the transform support is within the correlated region. At the same time, the compaction property of the local transform (which was the main consideration of the transform domain de-noising algorithm) will be lost if the transform support is out of the correlated range which will lead to detail loss or poor noise removal depending on the threshold. In any case, the estimate of a particular reference pixel will run out of the confidence interval when the transform size exceed the correlated region. On the other hand the introduced ICI rule helps to select the "optimum"

window size as a smoothing parameter when the signal is behaving as a local polynomial [8] (which is generally valid for simple signals but not necessary to be hold for relatively complex signals). Even after replacing the local polynomial with an adaptive model allowing the preservation of relatively smaller details, ICI rule remains to be a valid tool for verifying the validity of the local model with increasing size. When we use ICI rule for selecting the local transform size we simply equip local transform domain de-noising algorithm by a tool for selecting the locally optimum transform size for each window. Employing a local estimator adaptive to unknown smoothness instead of a local polynomial makes the adaptive bandwidth algorithm applicable for a larger range of images.

In Figure 2 the varying size local DCT filter is compared with the same filter with an optimum fixed size for the "montage" image. On the top left the original montage image is shown. This 256 x 256 8-bit image is a composition of 4 different types of images. On the top right of the Figure 2 the image corrupted by zero mean white Gaussian noise with a variance of 0.1 is shown. The root mean square error (RMSE) of the filtered image by DCT filter (bottom left) using adaptive window size was attenuated to 0.028 while the same filter with optimum fixed window of 5×5 attenuated the RMSE to 0.031 (bottom right).
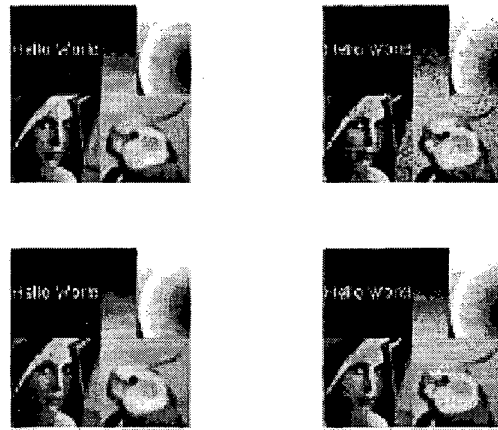

Figure 2: Comparison of adaptive and fixed window sizes in transform domain de-noising

In the following Table 1 some of the results for the same "montage" image by different filters were tabulated. As it can be seen from the performance figures, Local Adaptive

DCT filtering with adaptive size selected by ICI rule gives significantly better result than some other well known algorithms, namely, the classical Wiener filter, the wavelet shrinkage, wavelet packet shinkage and translation-invariant wavelet shrinkage using the Haar and Symmlet 8 filters [3, 4, 5], as well as the Local Adaptive DCT filtering using fixed window size [7].

Table1: Comparative results.

Used Filter	RMSE	MAE
LAD with adaptive transform size	0.0281	0.0191
LAD with fixed transform size	0.0314	0.0204
Wiener Filter (5x5)	0.0408	0.0274
Wavelet Packet Haar	0.0585	0.0444
Wavelet Packet Sym 8	0.0464	0.0310
Wavelet PO Haar (4 levels)	0.0567	0.0411
Wavelet PO Sym 8 (4 levels)	0.0693	0.0485
Wavelet TI Haar (5 levels)	0.0317	0.0205
Wavelet TI Sym 8 (5 levels)	0.0365	0.0261

MAE=Maximum Absolute Error
PO= Periodic, TI= Translation invariant

CONCLUSIONS

A new local adaptive transform based de-noising technique for removing additive Gaussian noise has been proposed and studied in this paper. The 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.

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] 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.

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

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

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

[6] 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.

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

[10] N. Ahmed, K. R. Rao "Orthogonal Transforms for Digital Signal Processing", Springer, New York, 1975