NOISE REDUCTION OF IMAGES WITH MULTIPLE SUBBAND TRANSFORMS

Toshihisa Tanaka

Lab. for Advanced Brain Signal Processing
RIKEN Brain Science Institute
Wako-shi, Saitama, 351–0198, Japan
t.tanaka@riken.jp

Laurent Duval

Technology Department
Institut Français du Pétrole
92852 Rueil-Malmaison Cedex, France
laurent.duval@ifp.fr

ABSTRACT

It is reported that the use of multiple number of subband transforms for thresholding-based denoising gains performance in the sense of the mean square error. In traditional thresholding-based methods, a noisy image is decomposed by linear transformation such as wavelets, FFT, and so on, and the transformed coefficients are hard-/soft-thresholded. In particular, it is well-known that wavelets work well for denoising. From the viewpoint that wavelets are in a class of subband transforms, we propose a strategy in which multiple number of subband transforms are switched region by region, i.e. block by block. For reconstruction, the projection-based iterative method is used. Experimental results are pretty good and promising.

1. INTRODUCTION

Restoration of an image or signal contaminated by noise is a fundamental problem in the field of image and signal processing. We deal in this paper with the additive noise model. This is very simple but still difficult, and can be applied to a lot of real practical problems. One of well-known solutions for this problem is a denoising method with linear transformation and thresholding [1]. In particular, wavelet transforms are very successful for pre-processing before thresholding [1]. Moreover, there have been several reports regarding how to choose a threshold value. Conventional thresholding (hard-thresholding) as well as shrinkage (soft-thresholding) are well analyzed.

It has been also reported [2] that uniform subband transforms or filter banks work quite well in denoising. They have been developed originally for image coding [3], and provide performance similar or superior to wavelets in image coding applications [4, 5]. This is due to the fact that those transforms provide good energy compaction property, which can be exploited by denoising. It has been moreover suggested in image coding that the so-called time-varying subband transform results in better performance than conventional subband transforms [6]. In this transform, multiple number of subband transforms are switched region by region, i.e., block by block depending on local images. In this paper, we would like to apply this successful strategy to image denoising. Our main motivations is to use the idea that there should exist more appropriate basis functions depending on local statistics. In the rest of this paper, we describe the denoising framework which we deal with and the difficulty of reconstruction in the case of the use of multiple subband transforms. We then show the projection-based reconstruction method used in the proposed procedure and illustrate how to choose a transform at each region (block). Finally, we show that the use of multiple number of transforms gains denoising performance in PSNR over conventional single transform methods throughout a simulation study.

This work was partially supported by Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Scientific Research (C) (2), 15560344, 2003.

2. MULTIPLE TRANSFORMS AND THRESHOLDING FOR DENOISING

2.1. Thresholding-Based Denoising Framework

Denoising scheme based on the combination of transformation and thresholding is generally described as follows. Let f be an original image. This image is transformed by a properly chosen linear transformation (e.g., wavelets, FFT, and lapped transforms) denoted by T. The transformed coefficients are hard-/soft-thresholded, and then the corresponding inverse transformation T^{-1} is applied to the thresholded coefficients. This non-linear operation can be written as

$$\hat{f} = T^{-1}\mathrm{Thr}[Tf], \tag{1}$$

where $\mathrm{Thr}[\cdot]$ denotes the thresholding and $\hat{f}$ is the reconstructed image. Various sophisticated thresholding operators have been investigated (see [1], for instance).

Usually, the transformation operator T (FFT, wavelets, subband transform, and so on) can be implemented by a block-diagonal matrix, or a periodically time-invariant system. This implies that the identical filters are applied to all regions. It is, however, not guaranteed that for all regions in an image, the same filters are suitable for denoising. Therefore, we introduce a notion that we switch the transformation region by region. Roughly speaking, when we switch two transforms, this idea can be formulated as

$$\hat{f} = T^{-1}\mathrm{Thr}[\tilde{P}_1 T_1 f + \tilde{P}_2 T_2 f], \tag{2}$$

where $\tilde{P}_1$ and $\tilde{P}_2$ are diagonal matrices with entries 0 or 1 such that $\tilde{P}_1 + \tilde{P}_2 = I$, and T^{-1} means here the inverse of the operator $\tilde{P}_1 T_1 + \tilde{P}_2 T_2$. An overview of this concept in denoising is depicted in Fig. 1. Even if we know T_1^{-1} and T_2^{-1}, the derivation of T^{-1} is not a trivial task. In the following subsections, we will have a more precise discussion of the use of multiple subband transforms.

2.2. Transform Settings

We formulate in this subsection the proposed analysis transformation system which use multiple uniform subband transforms (filter

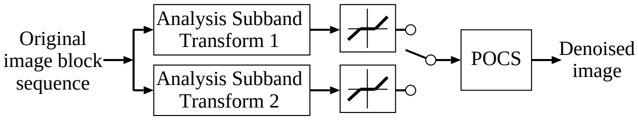


Fig. 1. Proposed denoising structure.

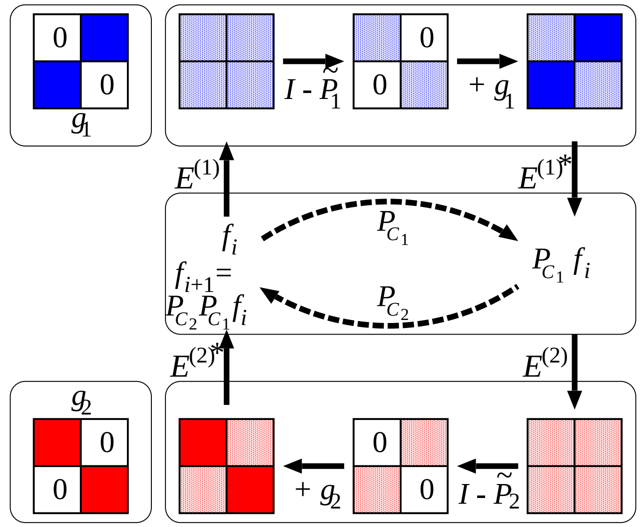


Fig. 2. POCS part in Fig. 1: Illustrative iteration procedure for reconstruction.

banks). As we mentioned earlier, the reason why we introduce the use of multiple subband transforms is that a natural image is generally a non-stationary process. This can be intuitively understood from the fact that a natural image can be roughly classified into smooth, texture, and edge regions, whose correlation functions might be different from each other. Therefore, there is a doubt that a single transform is applied to a whole image. From this point of view, the use of various subband transforms for decomposing an image will improve performance in denoising.

Suppose that there are L subband transforms implemented as an N-channel maximally decimated uniform filter bank, which consists of a set of N filters followed by downsampling by a factor of N. We also assume here that they have the same length KN, where K is an positive integer. Define the analysis transform matrix of lth subband transform as for $l = 0, \ldots, L-1$,

$$\boldsymbol{E}^{(l)} = [\boldsymbol{E}_{K-1}^{(l)}, \ldots, \boldsymbol{E}_0^{(l)}] \tag{3}$$

where $\boldsymbol{E}_k$, $k = 0, \ldots, K-1$, which leads to a polyphase matrix of the analysis part described by $\boldsymbol{E}^{(l)}(z) = \sum_{k=0}^{K-1} \boldsymbol{E}_k^{(l)} z^{-k}$. The rows of $\boldsymbol{E}^{(l)}$ correspond to the impulse responses of the analysis filters. Let $\{\boldsymbol{f}(t)\}_{t\in\mathcal{T}}$ be a set of tth blocked signals (or vectors at time instance t) consisting of N consecutive samples, where $\mathcal{T}$ describes a set of time indexes. The analysis subband transform $\boldsymbol{E}^{(l)}$ generates a sequence of the transformed vectors $\{\boldsymbol{g}(t)\}_{t\in\mathcal{T}}$ as

$$\boldsymbol{g}(t) = \boldsymbol{E}(t)\bar{\boldsymbol{f}}(t) \tag{4}$$

where $\bar{\boldsymbol{f}}(t) = [\boldsymbol{f}(t-(K-1))^T, \ldots, \boldsymbol{f}(t)^T]^T$, and $\boldsymbol{E}(t) = \boldsymbol{E}^{(l(t))}$, that is, $\boldsymbol{E}(t)$ is equal to some element in the subband transform set $\mathcal{E} = \{\boldsymbol{E}^{(l)}\}_{l=0}^{L-1}$. If we use a single subband transform in the same manner as conventional methods, $\boldsymbol{E}(t)$ is identical, say, independent of time instance t.

In most of applications, it is desired that the original sequence $\boldsymbol{f}(t)$ is perfectly reconstructed from the transformed sequence $\boldsymbol{g}(t)$. As we stated in 2.1, if we use a single perfect reconstruction subband transform, the reconstruction is performed just by applying the corresponding synthesis transform to the thresholded coefficients. However, in our case, i.e., when we switch several transforms region by region, even if we apply each synthesis transform to the corresponding coefficients, perfect reconstruction is generally impossible. As proved by Tanaka *et al* [6], however, if all subband transforms are unitary, the reconstruction can be done by applying the theory of projection onto convex sets (POCS), which is usually used in the area of image restoration [7].

2.3. Reconstruction With POCS

On the basis of the theory described in [6], we show in this section that by introducing appropriate closed convex sets, we can exactly reconstruct the original image from the transformed vector.

We obtain the transformed vector by the transformation described in (4). Recall that each $\boldsymbol{E}(t)$ corresponds one of elements in the analysis transform set $\mathcal{E}$, i.e., $\boldsymbol{E}(t) \in \mathcal{E}$. Now, we can rewrite (4) as

$$\boldsymbol{g}(t) = \sum_{l=0}^{L-1} \boldsymbol{\Delta}_l(t)\boldsymbol{E}^{(l)}\bar{\boldsymbol{f}}(t), \tag{5}$$

where

$$\boldsymbol{\Delta}_l(t) = \begin{cases} \boldsymbol{I}_M, & \text{if } \boldsymbol{E}(t) = \boldsymbol{E}^{(l)} \\ \boldsymbol{0}_M, & \text{otherwise,} \end{cases} \tag{6}$$

where $\boldsymbol{I}_M$ and $\boldsymbol{0}_M$ denote the identity and the null matrices of both size M. It is clear that $\sum_{l=0}^{L-1} \boldsymbol{\Delta}_l(t) = \boldsymbol{I}_M$. Let

$$\tilde{P}_l = \operatorname{diag}[\ldots, \boldsymbol{\Delta}_l(t), \boldsymbol{\Delta}_l(t+1), \ldots], \tag{7}$$

and

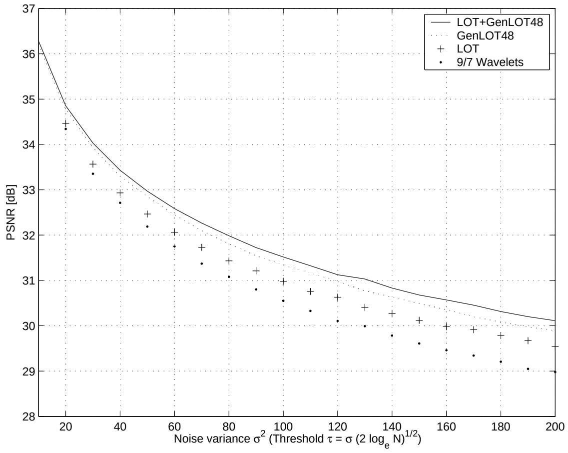
$$E^{(l)} = \begin{bmatrix} \ddots & & & & & 0 \\ & \boldsymbol{E}_{K-1}^{(l)} & \cdots & \boldsymbol{E}_0^{(l)} & & \\ & & \boldsymbol{E}_{K-1}^{(l)} & \cdots & \boldsymbol{E}_0^{(l)} & \\ 0 & & & & & \ddots \end{bmatrix}. \tag{8}$$

Then, $\tilde{P}_l$ represents an orthogonal projector with diagonal entries 0 or 1, and $E^{(l)}$ describes a subband transform operator which is not time-varying. Notice that $\sum_{l=0}^{L-1} \tilde{P}_l = I$. We can further modify (4) as $g = \sum_{l=0}^{L-1} \tilde{g}_l$, where $g_l = \tilde{P}_l E^{(l)} f$ and $f = [\ldots, \boldsymbol{f}(t-1)^T, \boldsymbol{f}(t)^T, \boldsymbol{f}(t+1)^T, \ldots]^T$, $g = [\ldots, \boldsymbol{g}(t-1)^T, \boldsymbol{g}(t)^T, \boldsymbol{g}(t+1)^T, \ldots]^T \in H$ (e.g. $H = \ell^2(Z)$). If the subband transform operator $E^{(l)}$ is unitary, i.e., $E^{(l)*}E^{(l)} = I$, where $\cdot^*$ denotes the adjoint, the operator P_l is an orthogonal projector, since it can be easily checked that $P_l^2 = P_l$ and $P_l^* = P_l$.

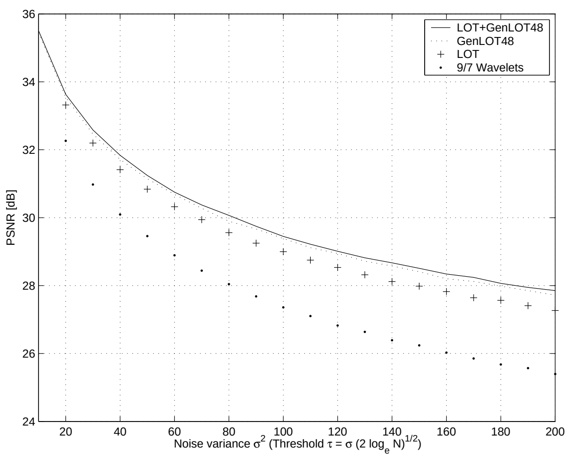
Define then the operator P_{C_l} as for $x \in H$,

$$P_{C_l} x = E^{(l)*}((I - \tilde{P}_l)x + g_l), \tag{9}$$

which gives the projector onto the convex set (linear manifold) defied as $C_l = E^{(l)} g_l + H_l^{\perp}$, where H_l is the subspace defined by P_l. Finally, the original signal is obtained as the limit of the recursion given as $f_{i+1} = P_{C_{L-1}} \cdots P_{C_1} P_{C_0} f_i$.

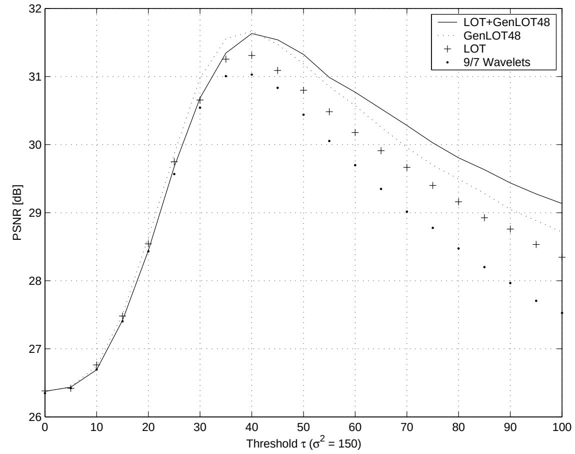


(a) Lena

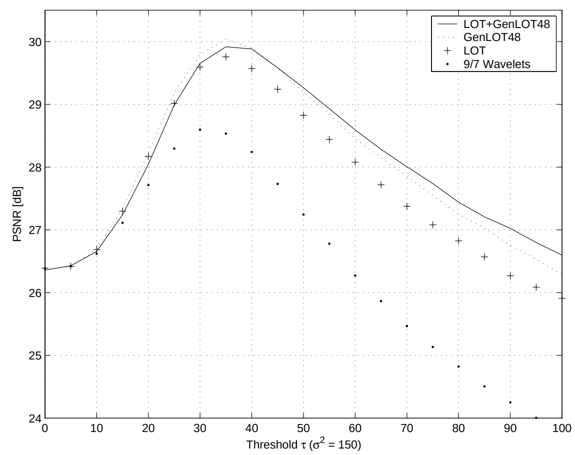


(b) Barbara

Fig. 3. Noise variance σ^2 v.s. peak signal-to-noise ratio (PSNR)



(a) Lena



(b) Barbara

Fig. 4. Threshold v.s. PSNR at the noise variance $\sigma^2 = 150$

2.4. Transform Selection

We have to define the criteria to choose the most appropriate transform at the encoder. At each time instance (block) t, we can obtain L different transformed vectors denoted by $\boldsymbol{g}^{(l)}(t)$, $l = 0, \ldots, L-1$, as $\boldsymbol{g}^{(l)}(t) = \boldsymbol{E}^{(l)}\bar{\boldsymbol{f}}(t)$. Then, we choose one vector from the L vectors. Indeed, how to choose the best basis in this method is a very difficult problem. Theoretical consideration would be needed; in this paper, however, the main purpose is to show that the use of multiple subband transforms improve performance in denoising. Therefore, the following simple criterion is adopted. Let $\hat{\boldsymbol{g}}^{(l)}(t)$ be a coefficients vector at block t after hard-/soft-thresholding. Firstly, we define the following cost function:

$$J[\boldsymbol{g}^{(l)}(t)] = \|\hat{\boldsymbol{g}}^{(l)}(t)\| / \|\boldsymbol{g}^{(l)}(t)\|. \tag{10}$$

We choose the filter bank which gives the maximum value of $J[\boldsymbol{g}^{(l)}]$. If no noise is added, this criteria selects the filter bank that mostly concentrates the energy after thresholding.

3. EXPERIMENTAL RESULTS

To illustrate the advantage of the use of multiple filter banks, we show some examples of image denoising. Recall that the additive noise model is supposed.

3.1. Choice of Threshold

We use here just the single fixed threshold suggested by Donoho *at al* [1] given by

$$\tau = \sigma\sqrt{2\log_e N}, \tag{11}$$

where σ is the standard deviation of the noise and N is the number of pixels of the image. By using this threshold, we apply "hard-thresholding" to the transform coefficients.

In denoising, how to choose the threshold is an important issue and several sophisticated thresholds have been investigated [1]; however, in this paper, in order to clearly emphasize and show the

(a) Noisy image (b) 9/7 biorthogonal wavelet (29.61 dB) (c) Proposed method (30.67 dB)

Fig. 5. Noisy and denoised images ‘Lena’ when $\sigma^2 = 150$

effectiveness of the use of multiple transforms and POCS reconstruction, we make a system simple, i.e., we use a fixed threshold for a whole image as done in [8]. For practical applications, of course, we should estimate the variance of noise. Furthermore, more sophisticated threshold might lead to better performance.

3.2. Results

In this test, we compare denoising performance of several transforms in the following: 1) 8-channel lapped orthogonal transform (LOT) [9], 2) 8-channel GenLOT of filter length 48 [9], 3) 9/7 biorthogonal wavelet [1], and 4) the proposed multiple-transform method in which the LOT and the GenLOT are selectively used. For reconstruction, the number of iteration is 20, which yields sufficient quality. In order to show the gain of performance at the same threshold, we depict PSNRs at various noise variances with respect to the well-known pictures Lena and Barbara of both size 512×512 in Fig. 3. As we can see in this figure, the use of multiple transforms improve PSNR consistently. In Fig. 4, next, we show the effect of our proposed procedure for various thresholds. At smaller thresholds, the GenLOT provides the best performance; however, after the peak, the proposed method gives the best quality in PSNR. Finally, we provide subjective comparison in Fig. 5. We show here the noisy image of the noise variance $\sigma^2 = 150$, the denoised images by 9/7 biorthogonal wavelet and the proposed multiple transforms. In the image generated by the wavelets, blurring around strong edges are very significant. It is observed that the proposed multiple-transform method suppresses blurring more than the wavelet. Moreover, PSNR of our proposed method outperforms that of wavelets.

4. CONCLUSIONS

We have shown that the use of multiple number of subband transforms in denoising gains performance in the sense of the mean square error compared to the case of single transform. We have proposed a concept that several subband transforms are switched region by region, and illustrated a reconstruction procedure which accomplishes perfect reconstruction on the basis of the POCS. Now, we have some open problems. We haven’t mentioned in this paper how to find the optimal threshold. This would be addressed in future.

5. REFERENCES

[1] S. Mallat, *A Wavelet Tour of Signal Processing*. New York, NY/London: Academic Press, 1998.

[2] L. Duval and T. Tanaka, “Denoising of seismic signals with oversampled filter banks,” in *Proc. ICASSP 2003*, vol. VI, (Hong Kong), pp. 189–192, 2003.

[3] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley MA: Wellesley-Cambridge Press, 1996.

[4] T. D. Tran, R. L. de Queiroz, and T. Q. Nguyen, “Linear-phase perfect reconstruction filter bank: Lattice structure, design, and application in image coding,” *IEEE Trans. Signal Processing*, vol. 48, pp. 133–147, Jan. 2000.

[5] T. Tanaka, Y. Hirasawa, and Y. Yamashita, “A novel class of variable-length lapped transform for image coding,” in *Proc. ICIP 2003*, vol. III, (Barcelona, Spain), pp. 201–204, 2003.

[6] T. Tanaka, T. Saito, and Y. Yamashita, “A time-varying subband transform with projection-based reconstruction,” *IEICE Trans. Fundamentals*, vol. E86-A, pp. 1935–1941, Aug. 2003.

[7] D. C. Youla and H. Webb, “Image restoration by the method of convex projections: Part 1 – Theory,” *IEEE Trans. Medical Imaging*, vol. MI-1, pp. 81–94, Oct. 1982.

[8] P. Ishwar, K. Ratakonda, P. Moulin, and N. Ahuja, “Image denoising using multiple compaction domains,” in *Proc. ICASSP ’98*, (Seattle, Washington, USA), pp. 1889–1892, 1998.

[9] R. L. de Queiroz, T. Q. Nguyen, and K. R. Rao, “The GenLOT: Generalized linear-phase lapped orthogonal transform,” *IEEE Trans. Signal Processing*, vol. 44, pp. 497–507, Mar. 1996.