

# A NEW INTERPRETATION OF TRANSLATION INVARIANT DENOISING

Gang Hua, Michael T. Orchard

Department of Electrical and Computer Engineering  
Rice University  
Houston, TX 77005 USA

## ABSTRACT

Translation Invariant (TI) image denoising outperforms orthogonal wavelet thresholding by averaging a collection of denoised estimates from different orthogonal bases. This paper proposes a new perspective of TI processing as an average of a collection of *cyclic-basis* frame reconstructions, each a stationary signal estimate, contrasting with the nonstationary estimates of orthogonal wavelet thresholding. This viewpoint clarifies that certain characteristics of TI (i.e. reduced edge contour artifacts) are inherited from each cyclic-basis reconstruction, rather than from the process of averaging. We relate performance advantages of TI in smooth areas of images to statistical relationships of the cyclic-basis reconstructions. In edge regions, the quality of cyclic-basis reconstructions vary significantly with pixel position relative to the edge contour. These differences couple with convexity arguments to explain large performance gains of TI in edge regions. They also suggest an improved approach to frame reconstruction, based on estimating relative location information, and identifying the best cyclic-basis reconstruction for the estimated pixel location.

## 1. INTRODUCTION

Wavelet thresholding [1] for image denoising has been researched extensively because of its effectiveness and simplicity. This approach expands the image in a wavelet basis, keeps approximation coefficients unchanged, and compares each detail coefficient to a threshold. If it is below the threshold, the coefficient is set to zero; if it is above the threshold, it is kept unchanged or modified by some amount. Thresholding in an orthogonal wavelet basis has been observed to produce Gibbs-like visual artifacts around edges [2]. Translation invariant (TI) denoising has been proposed to mitigate these artifacts. TI can be considered as the average of the results of thresholding in different orthogonal bases (differing only in translation). The denoising performances in all these orthogonal bases are approximately the same. By averaging, TI gets a significant performance gain over each of these bases (Fig. 1(a)).

In [3, 4], the relationship between wavelet thresholding in an orthogonal basis and TI is explained in terms of the effective regularity of the denoising kernels, suggesting that TI can be viewed as contributing some extra smoothing to the single-basis denoising algorithm. While this is indeed correct, we observe that more fundamental differences between TI and basic wavelet-based denoising algorithms can be identified, relating to how they respect the stationarity of the desired signals. The stationarity property of the assumed signal model implies that the statistical relationship between each pixel and its neighborhood of pixels is invariant to translation. It can be easily recognized, however, that the same

translation invariance is not found in the way each pixel is reconstructed by the single-basis wavelet-denoising algorithm. For basic wavelet denoising, the treatment of each pixel differs depending on where the pixel is located relative to the arbitrary downsampling grid of the wavelet decomposition. For smooth, stationary Gaussian processes, we describe how this lack of translation invariance creates relationships between the approximation error and residual noise variance of the collection of single-basis wavelet denoised images that improve the quality of the averaged TI denoised image.

Building from these insights, we define a new collection of translation-invariant frame-based *cyclic-basis* reconstructions, and show that TI denoising can also be viewed as the average of cyclic-basis reconstructions, each reflecting the stationarity of the underlying signal model. Unlike the collection of single-basis reconstructions, which all have the same average noise performance, the quality of the cyclic-basis reconstructions differ significantly, and we investigate the cause and consequences of these differences for smooth Gaussian processes.

Although the analysis of TI denoising and cyclic-basis reconstruction is most tractable for smooth, stationary Gaussian processes, the most significant gains of TI denoising, and the most significant differences among cyclic-basis reconstructions are due to reconstruction of edge regions in images. Around edges, pixel statistics are highly dependent on the pixel location relative to the edge contour. (I.e. given edge contour information, the signal is non-stationary.) Thus, we show that the stationary cyclic-basis reconstructions vary significantly depending on pixel position. We show that these large variations in errors of cyclic-basis reconstructions, coupled with the convexity of the mean-squared-error performance measure, explain the substantial processing gains realized by TI. We also recognize that certain characteristics of TI (i.e. removal of contour artifacts) that have been viewed as results of averaging actually reflect characteristics of each and every cyclic-basis reconstruction, before any averaging. Finally, our understanding of TI performance also points to an improved approach to wavelet denoising based on estimating relative location information in the neighborhood of edges. When location information can be reliably estimated, the best cyclic-basis reconstruction given pixel location offers significant gain over TI denoising.

## 2. CYCLIC-BASIS RECONSTRUCTION

### 2.1. Orthogonal and translation invariant denoising

In 1D case, assume  $f$  is a discrete signal with  $N = 2^m$  samples. Let  $W_k$  and  $W_k^{-1}$  denote the orthogonal wavelet transform with translation  $k$  ( $k = 0, 1, \dots, M - 1$ ), corresponding to periodi-

cally translating  $f$  by  $k$  before applying the ordinary DWT, and translating by  $-k$  after applying the ordinary inverse DWT. For an  $L$ -level decomposition,  $M = 2^L$ . For each  $k$ ,  $W_k$  defines an orthogonal basis. In the 2D case,  $f$  is a  $N$  by  $N$  signal and there are  $M^2$  orthogonal bases, with  $k = m * M + n$ . We will sometimes use  $(m, n)$  to denote basis  $k$  in 2D case. Then the estimation in the  $k^{th}$  basis is

$$\hat{f}_k = W_k^{-1}(D_T(W_k(g))), \quad (g = f + n),$$

where  $n$  is i.i.d. Gaussian noise, and  $D_T(x)$  is soft thresholding. Soft thresholding keeps all approximation coefficients unchanged, and replaces detail coefficient  $y$  with  $sgn(y)(|y| - T)_+$ . Given  $M$  (or  $M^2$  in 2d) such estimates, the TI estimate is formed:

$$\hat{f}_{TI} = \frac{1}{M} \sum_{k=0}^{M-1} \hat{f}_k$$

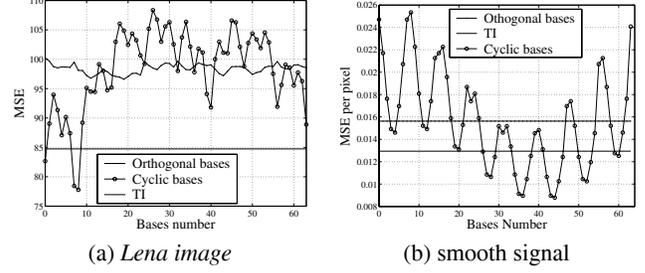
For stationary signals, all the estimates  $\hat{f}_k$  should have equal quality. For natural images, it is observed [2, 5] that the mean square errors of  $\hat{f}_k$  ( $E\|\hat{f}_k - f\|^2$ ) are approximately equal, and MSE of  $\hat{f}_{TI}$  ( $E\|\hat{f}_{TI} - f\|^2$ ) is about 20 - 25% lower than that of any  $\hat{f}_k$ .

## 2.2. Cyclic-basis Reconstruction

Since natural images are stationary, an ideal signal estimate should preserve this property. Because orthogonal wavelet transforms are not translation invariant, orthogonal wavelet shrinkage produces non-stationary estimates. If the tight frame formed by taking together basis vectors from all the  $M$  ( $M^2$  in 2d) bases is considered, then all  $\hat{f}_k$  and  $\hat{f}_{TI}$  can be considered different frame reconstructions from the same frame. Assuming that the denoising operation on the coefficients is translation invariant (true for [2, 5]), the whole denoising process preserves stationarity if and only if the reconstruction is also translation invariant.

This paper proposes a new translation-invariant frame reconstruction, called *cyclic-basis* reconstruction. In the 2D case, define cyclic-basis reconstruction  $\hat{f}_{(m,n)}^c$ , ( $0 \leq m, n \leq M - 1$ ) to be the reconstruction that assigns at pixel  $(i, j)$  the estimate from orthogonal basis  $([i + m]_M, [j + m]_M)$ , where  $[x]_M$  is the modulo  $M$  operation. I.e.  $\hat{f}_{(m,n)}^c(i, j) = \hat{f}_{([i+m]_M, [j+m]_M)}(i, j)$ . The definition in 1D is similar. This reconstruction has the form of processing the frame coefficients on each scale with a linear filter associated with that scale and then adding the results together. Thus, cyclic-basis reconstruction is translation invariant and all  $\hat{f}_{(m,n)}^c$  are stationary. It is worth noting that, just as each  $\hat{f}_k^c$  is a periodic assignment of pixels from the collection of  $\hat{f}_k$ ,  $0 \leq k \leq M - 1$ , we can recognize each  $\hat{f}_k$  to be a periodic assignment of pixels from the collection of  $\hat{f}_k^c$ ,  $0 \leq k \leq M - 1$ . Thus, the nonstationarity of each estimate  $\hat{f}_k$  can be recognized as reflecting the statistical variation among the collection of stationary estimates  $\hat{f}_k^c$ ,  $0 \leq k \leq M - 1$ . Moreover, since TI can be considered as the average of cyclic-basis reconstructions, it inherits many of the properties of these reconstructions. Studying thresholding in cyclic-basis reconstructions sheds new light on the performance of TI, as well as suggests directions to improve performance.

For natural images, the performances of cyclic bases have very large difference, while those of orthogonal bases are about the same (Fig. 1(a)).



**Fig. 1.** Denoising performance of TI, orthogonal and cyclic bases for signals with additive white Gaussian noise  $N(0, 400)$ , threshold  $T=60$  and 3-level symlet-8 wavelet decomposition.

## 3. ANALYSIS OF SMOOTH AREA

### 3.1. Smooth signal model

In smooth areas of an image, detail wavelet coefficients are small compared to the noise, and do not survive the thresholding operation (in [5] very large thresholds are used for smooth areas). Then  $\hat{f}_k$  is the linear projection of  $g$  onto the wavelet approximation subspace  $V_L^k$  ( $\hat{f}_k = P_k(g)$ ) of orthogonal basis  $k$ . And the MSE in basis  $k$  is

$$\begin{aligned} E\|\hat{f}_k - f\|^2 &= E\|P_k(f + n) - f\|^2 \\ &= E\|P_k(f) - f\|^2 + E\|P_k(n)\|^2 \end{aligned}$$

The two terms on the last line, approximation error and residual noise, have the same value for all orthogonal bases. The approximation error  $P_{TI}(f)$  and residual noise  $P_{TI}(n)$  of TI are the average of those of all the orthogonal bases. Since the denoising operation on frame coefficients is linear (in smooth areas), and both cyclic-basis reconstruction and TI are time-invariant, both reconstructions can be viewed representing the output of equivalent denoising filter, where the TI filter is recognized as the average of the collection of equivalent cyclic-basis reconstruction filters.

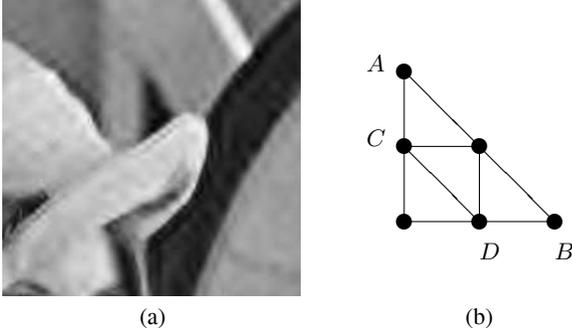
We model the signal in smooth areas as the sum of two components: a polynomial function with order less than  $D$ , and a stationary random Gaussian process. For natural images the polynomial component often dominates, and the Gaussian component may come from processing noise. Assuming  $D$  is less than or equal to the number of the vanishing moments of the wavelet used, then approximation error for the polynomial component in any orthogonal wavelet bases is zero. Correspondingly, it contributes zero error to the approximation error of TI. Thus only the approximation error of the second component need to be considered.

### 3.2. Residual noise and approximation error

The residual noise  $P_k(n)$  is a cyclo-stationary Gaussian process. The cycle is  $2^L$  for 1D signal, and  $(2^L, 2^L)$  for 2D signal. And the residual noises in cyclic-basis reconstruction and TI are stationary Gaussian processes. It can be shown that, for any  $L$

$$\begin{aligned} E\|P_k(n)\|^2 &= N^K \sigma_n^2 / K^L \\ E\|P_{TI}(n)\|^2 &= E\|P_k(n)\|^2 \int \Phi^2(x) dx \end{aligned}$$

where  $\sigma_n^2$  is the noise variance,  $K$  is 2 for 1D signal and 4 for 2D signal, and  $\Phi(x)$  is the autocorrelation function of the wavelet



**Fig. 2.** Artifacts in orthogonal basis (a) and Edge Geometry (b)

scaling function. The energy of  $\Phi(x)$  is smaller than or equal to one. For 1D Haar wavelet it is  $2/3$ , and for 1D and 2D Daubechies' symlet wavelet with eight vanishing moments (used in [5] for image denoising) they are 0.91 and 0.83 respectively. For each cyclic basis, the residual noise is stationary. The residual noise variance (Fig. 1(b)) is determined by its frame reconstruction filter.

For the signal component comes from a stationary random Gaussian process model, approximation error can be analyzed in the same way as the residual noise. It can be shown

$$E\|f - P_k(f)\|^2 = \int 2^{-L} P_s(2^{-L}f)(1 - P_\Phi(f)) df$$

$$E\|f - P_{TI}(f)\|^2 = \int 2^{-L} P_s(2^{-L}f)(1 - P_\Phi(f))^2 df$$

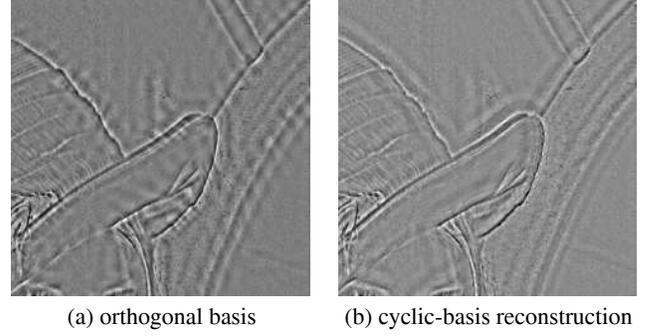
where  $P_s(f)$  and  $P_\Phi(f)$  are the power spectrum of the signal and wavelet scaling function respectively.  $P_\Phi(f)$  is less than or equal to 1, and it is 1 at  $f = 0$ . So TI has smaller approximation error than orthogonal bases (TI recovers some of the signal lost in thresholding). If the wavelet has  $D$  vanishing moments, then the  $1^{st}$  to  $(2D-1)^{th}$  derivatives of  $P_\Phi(f)$  at  $f = 0$  vanish. This result is consistent with [3, 4]. For each cyclic basis, the error spectrum can be characterized by its frame reconstruction filter.

The projection of stationary signal and noise onto subspaces  $\{V_L^k, 0 \leq k < M\}$  produces a collection of nonstationary processes. The "angle" between these subspaces can be measured by the correlation matrix of the collection of projected estimates at any pixel. Though space does not permit a full treatment, larger angles between subspaces increase the independence of the residual noise seen by each estimate, and thus increases the effectiveness of TI averaging in reducing noise variance. However, larger angles between subspaces also increase the nonstationarity of the estimate, thus distancing the estimate from the true stationary signal, and increasing approximation error, both for each  $\hat{f}_k$  and  $f_{TI}$ . The "angle" can be controlled by design of the wavelet, thus providing a trade-off between noise reduction and approximation error.

## 4. ANALYSIS OF EDGE AREA

### 4.1. Artifacts in orthogonal wavelet thresholding

Orthogonal wavelet thresholding produces pseudo-Gibbs phenomena around edges. In images, we recognize two kinds of artifacts: (i) overshoots and undershoots in the direction vertical to the edge contour (section 4.3), and (ii) discontinuities along the direction of



**Fig. 3.** Estimation Bias for *Lena* image

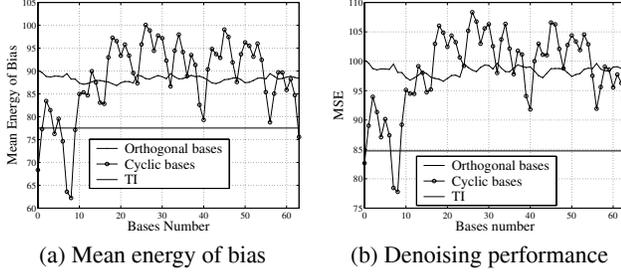
the edge contour (section 4.2), as shown in Fig. 2(a). Fig. 2(b) illustrates the edge geometry producing the second type of artifact. Suppose an edge runs from  $A$  to  $B$ , with pixel  $C$  and  $D$  having the same relative location relative to the edge (and the same gray level). In a given orthogonal basis  $(m, n)$ ,  $C$  and  $D$  are represented by a common collection of wavelet coefficients, but different weights are applied to those coefficients to reconstruct pixels  $C$  and  $D$ . Thresholding these coefficients thus has different effects on  $C$  and  $D$ . Treating pixels differently though they have the same location relative to the edge creates periodic processing structures, which we recognize as the second type of artifact discussed above.

To eliminate this type of artifact,  $C$  and  $D$  should be computed with the same reconstruction weights applied to the same coefficient values, and this should be true for any pixels  $C$  and  $D$  equidistant from any arbitrary edge orientation. This can only hold if  $C$  and  $D$  have access to a translation-invariant collection of coefficients, and are computed with a translation-invariant set of weights. Since these conditions are met for each cyclic-basis reconstruction  $\hat{f}_{(m,n)}^c$ , all of these reconstructions are images without these types of artifacts. Naturally, TI, the average of all these images, also lacks these artifacts.

Figure 3 shows how the cyclic-basis reconstruction reduces the periodic artifacts along edges in the *Lena* image compared with the orthogonal wavelet reconstruction. The pictures show the estimation bias of the two reconstructions, where bias is defined as  $E_n\{[f - \hat{f}]|f\}$ , for each of the estimators  $\hat{f}_k$  and  $\hat{f}_k^c$ . We note that the estimation bias is the dominant source of error energy around edges when the threshold is large compared with the noise standard deviation. We also note that stationarity of  $\hat{f}_k^c$  only ensures that the periodic artifacts along edges are reduced, not that  $\hat{f}_k^c$  is without error. In fact, since the collection of estimates  $\{\hat{f}_k^c\}$  is formed by a permutation of the pixels in  $\{\hat{f}_k\}$ , the total error in the two sets must be equal. The next section discusses how the cyclic-basis reconstructions provide a clearer organization of that error, and give us a slightly better handle on characterizing the error.

### 4.2. Convexity and the gain of TI

The processing that produces each  $\hat{f}_k$  estimate is a shifted version of the processing that results in  $\hat{f}_{k-1}$ . Assuming an underlying stationary signal and noise model, we can conclude that the denoising performance of all  $\hat{f}_k$  are statistically indistinguishable. In contrast, each  $\hat{f}_k^c$  represents quite different processing than any other one, and thus we have no reason to expect that their denoising per-



**Fig. 4.** Performance of soft thresholding (with threshold  $T=60$  and 3-level wavelet decomposition) in orthogonal and cyclic bases for *Lena image* with additive white Gaussian noise  $N(0, 400)$

formance would be the same. In fact, Figure 4 both confirms the similarity in quality of all  $\hat{f}_k$ , and shows significant differences in the error energy of the  $\hat{f}_k^c$ . Before considering the possibility of exploiting these differences by estimating the minimum-error  $\hat{f}_k^c(i, j)$ , we discuss here how these differences, coupled with the convexity of the MSE metric, explain the large gains achieved by TI in edge regions.

The mean square error ( $E\|x\|^2$ ) metric is a convex function, and according to the definition of  $\hat{f}_{TI}$

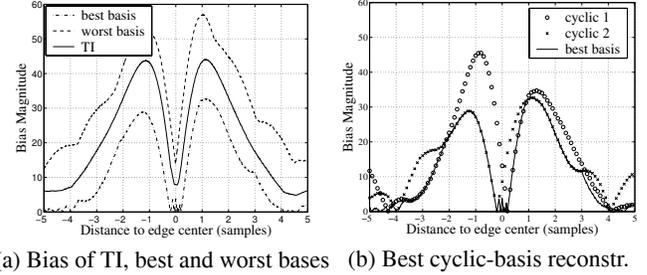
$$\begin{aligned} E\|\hat{f}_{TI} - f\|^2 &\leq \frac{1}{M} \sum_{k=0}^{M-1} E\|\hat{f}_k - f\|^2 \\ &= E\|\hat{f}_k - f\|^2 \end{aligned}$$

For each pixel  $(i, j)$ , the difference between the two sides of the inequality is the sample variance of the collection of  $M$  estimates,  $\{\hat{f}_k^c(i, j)\}$ . Thus, the gain of TI in edge areas depends on the differences in error energies of the estimates  $\{\hat{f}_k^c\}$ , which is illustrated in figure 5(a). Figure 5(a) depicts the range of absolute estimation biases for a collection of 8 cyclic-basis reconstructions (1D example with  $L = 3$ ), and compares it to the absolute estimation bias of TI. The two curves defining the range of biases represent the pixel-wise minimum and maximum biases over all the cyclic-basis reconstructions, as functions of the location of the pixel relative to the center of the edge. Note that most absolute bias is contributed by pixels approximately 1-3 samples from edge centers (actual distances depends on sharpness of edges). Across all distances, we see significant range of variation between minimum and maximum biases, explaining the large gains of TI in edge regions.

### 4.3. Best cyclic basis selection

Both figures 4 and 5(a) point towards the possibility of more aggressively combining the set of estimates  $\{\hat{f}_k^c\}$  to compute a better estimate than  $\hat{f}_{TI}$ . Specifically, we see in figure 4(b) that two cyclic-basis reconstructions (8 and 9) appear to have significantly better performance than  $\hat{f}_{TI}$ . We see in figure 5(a) that the minimum-bias estimator for each pixel location has significantly lower bias than  $\hat{f}_{TI}$ .

As attractive as figure 5(a) appears, we recognize that the feasibility of every achieving the minimum-bias estimator is limited by the structure involved in selecting the best  $\hat{f}_k^c$  from the set  $\{\hat{f}_k^c\}$ .



**Fig. 5.** Best cyclic-basis reconstruction for a Gaussian filtered ideal edge with symlet-8 wavelet.

E.g. if the  $k$  defining the lower curve in figure 5 is randomly selected for each pixel position from the range  $0 \leq k < M$ , it would be unreasonable to expect any algorithm to find the best  $k$ . Figure 5(b) suggests that finding the best  $k$  may be quite feasible, at least for the most important pixel positions. The solid line in 5(b) corresponds to the minimum-bias curve of 5(a), and shows that this minimum-bias curve coincides with the bias curves of only two cyclic-basis reconstructions for most pixel positions with the highest biases. It is worth noting that the two "good" cyclic-basis reconstructions in this figure are the 1D corresponding reconstructions to the two cyclic-basis reconstructions (8 and 9) of figure 4(b).

The choice of the best basis for a pixel around an edge depends mostly on the location and slope of the edge. If this information is available, the best basis for each pixel with very small bias (cyclic bases 1 or 2 for symlet-8) can be found. The performance of this adaptive best basis promises to be much better than any of the cyclic-basis reconstructions shown in Fig.4(b). A practical algorithm for reliably estimating best cyclic-basis reconstruction remains a subject of ongoing research.

## 5. REFERENCES

- [1] D.L. Donoho, "De-noising by soft thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [2] R. R. Coifman and D. L. Donoho, "Translation-invariant denoising," in *Lecture Notes in Statistics: Wavelets and Statistics*, A. Antoniadis and G. Oppenheim, Eds. Springer-Verlag, Berlin, Germany, 1995.
- [3] Kathrin Berkner and Jr Raymond O. Wells, "Smoothness estimates for soft-threshold denoising via translation-invariant wavelet transforms," *Applied and Computational Harmonic Analysis*, vol. 12, no. 1, pp. 1–24, Jan 2002.
- [4] Juan Liu and P. Moulin, "Approximation-theoretic analysis of translation invariant wavelet expansions," *Proceedings of 2001 International Conference on Image Processing*, vol. 1, pp. 622–625, 2001.
- [5] S.G. Chang, Bin Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding based on context modeling for image denoising," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1522–1531, Sep 2000.