STATISTICAL LEARNING OF RATIONAL WAVELET TRANSFORM FOR NATURAL IMAGES

Naushad Ansari* and Anubha Gupta

SBILab, Deptt. of ECE, IIIT-Delhi, India

ABSTRACT

Motivated with the concept of transform learning and the utility of rational wavelet transform in audio and speech processing, this paper proposes <u>Rational Wavelet</u> Transform <u>Learning</u> in the <u>Statistical sense</u> (RWLS) for natural images. The proposed RWLS design is carried out via lifting framework and is shown to have a closed form solution. The efficacy of the learned transform is demonstrated in the application of compressed sensing (CS) based reconstruction. The learned RWLS is observed to perform better than the existing standard dyadic wavelet transforms.

Index Terms— Rational Wavelet, Statistically Matched Wavelet, Natural Images, Lifting Framework

1. INTRODUCTION

Transform learning (TL) is an active research area where the sparsifying transform along with the transform domain signal is learned using some constraints for a class of signals. Currently, TL is being used in several applications including image/video denoising and MRI reconstruction [1–3]. While TL is being used actively, non-convexity of the problem having no closed form solution makes it difficult to solve. Hence, greedy algorithms are used to solve TL problem.

Among existing transforms, discrete wavelet transform (DWT) is widely used in applications because of its ability of efficient signal representation [4]. Non-uniqueness of wavelet basis motivates one to learn wavelet transform in applications. Wavelet transform learning can be viewed as a specific case of transform learning. Since the integer translates of the associated wavelet filters form the basis in l_2 -space, wavelet transform learning of wavelet transform learning of wavelet filters.

Generally, the dyadic wavelet transform is used that decomposes input signal spectrum into two uniform frequency bands via two-channel filterbank. On the other hand, rational wavelet transform (RWT) provides non-uniform frequency band representation of signal spectrum that is seen to be useful in some applications [5, 6]. RWT has also been used in pattern recognition [7] and feature extraction [8]. Although methods for RWT designs have been presented in the literature [9–12], designed wavelets are independent of the signal of interest. Recently, a method has been proposed in [13] to learn rational wavelet deterministically from a given signal. Since [13] requires full signal, it cannot be used in inverse problems such as CS-based reconstruction where one does not have access to the full original signal.

This paper proposes rational wavelet learning for natural images. It has been shown that natural images can be modeled as fractional Brownian motion (fBm) processes in [14]. fBm processes are Gaussian non-stationary random processes with stationary increments that form a class of statistically self-similar processes [15] and have been used widely in image processing [16, 17].

The above discussion of transform learning, flexibility of rational wavelet transform, and modeling of natural images via fBm processes motivates us to learn rational wavelet transform for natural images in the statistical sense. Specifically, statistics of a set of natural images are used to propose a method for learning separable rational wavelet transform for this class of images. Lifting framework for rational wavelet introduced in [13] is utilized in the proposed work and is called the RWLS method. The proposed formulation leads to a convex problem that can be solved by least squares making the RWLS method computationally efficient.

Following are the salient contributions of this work:

- 1. Statistical learning of rational separable wavelet transform for natural images is proposed.
- 2. Lifting framework, that is Digital Signal Processing (DSP) hardware friendly, is used in the proposed method making the learned transform easily implementable on hardware.
- 3. The proposed formulation leads to convex problem unlike conventional TL and can be solved easily.
- 4. The proposed RWLS is applied in compressed sensing based reconstruction and is observed to perform better than the existing dyadic wavelet transforms.

2. BRIEF BACKGROUND

2.1. Lifting in Dyadic Wavelet

Lifting methodology supports customized wavelet design [18], [19]. This design is modular, guarantees perfect reconstruction, and allows non-linear filters to be part of the

^{*}Thanks to CSIR (Council of Scientific & Industrial Research), Govt. of India for funding.

wavelet structures. A general lifting scheme consists of three steps: Split, Predict, and Update (Refer to Fig. 1). In the split step, given input signal is divided into even $x_e[n]$ and odd $x_o[n]$ indexed samples. The corresponding filterbank structure is called as the Lazy wavelet system [19] and is converted to the conventional wavelet system using successive predict and update stage filters as shown in Fig. 2 with analysis filters labeled as $G_l(z) = Z\{g_l[n]\}, G_h(z) = Z\{g_h[n]\}$ and the synthesis filters as $F_l(z) = Z\{f_l[n]\}, F_h(z) = Z\{f_h[n]\}.$

In the Predict Lifting step, odd samples are predicted from the neighboring even samples using the predictor $P \equiv T(z)$ or vice-versa. This step modifies the analysis highpass filter and the synthesis lowpass filter as:

$$G_h^{new}(z) = G_h(z) - G_l(z)T(z^2).$$
 (1)

$$F_l^{new}(z) = F_l(z) + F_h(z)T(z^2).$$
 (2)

The Update Lifting step modifies the analysis lowpass filter



Fig. 1: Steps of Lifting: Split, Predict and Update



Fig. 2: Two Channel Wavelet System

and the synthesis highpass filter. The update step filter is denoted with the symbol $U \equiv S(z)$ and the related equations are given as:

$$G_l^{new}(z) = G_l(z) + G_h(z)S(z^2).$$
 (3)

$$F_h^{new}(z) = F_h(z) - F_l(z)S(z^2).$$
 (4)

2.2. Rational Wavelet

Let us consider Fig. 3(a) with 2-channel $(\frac{2}{3}, \frac{1}{3})$ rational wavelet filterbank that can be converted into an equivalent uniformly decimated *M*-band structure (Fig. 3(b)). Filters $G_0(z)$ and $G_1(z)$ of Fig. 3(b) can be written as an equivalent filter $G_l(z)$ of Fig. 3(a) using the following equation:

$$G_l(z) = G_0(z^2) + z^3 G_1(z^2).$$
 (5)

Similarly, synthesis filters $F_0(z)$ and $F_1(z)$ of Fig. 3(b) can be written as an equivalent filter $F_l(z)$ of Fig. 3(a) using the following equation:

$$F_l(z) = F_0(z^2) + z^{-3}F_1(z^2),$$
(6)

while other filters remain same, i.e., $G_h(z) = G_2(z)$ and $F_h(z) = F_2(z).$

$$x[n] \rightarrow \underbrace{F_{l}(z)}_{G_{l}(z)} \rightarrow \underbrace{F_{l}(z)}_{F_{l}(z)} \rightarrow \underbrace{F_{l}(z)}_{F_{l}(z)} \rightarrow \widehat{x}[n]$$
(a) Two-channel $(\frac{2}{3}, \frac{1}{3})$ rational filterbank
$$x[n] \rightarrow \underbrace{G_{0}(z)}_{G_{1}(z)} \rightarrow \underbrace{F_{3}}_{F_{3}} \rightarrow \underbrace{F_{0}(z)}_{F_{0}(z)} \rightarrow \widehat{x}[n]$$

$$x[n] \rightarrow \underbrace{G_{1}(z)}_{G_{2}(z)} \rightarrow \underbrace{F_{3}}_{F_{3}} \rightarrow \underbrace{F_{2}(z)}_{F_{2}(z)} \rightarrow \underbrace{f_{1}(z)}_{F_{2}(z)} \rightarrow \underbrace{f_{1}($$

(b) Equivalent *M*-band (M = 3) filterbank

Fig. 3: Rational wavelet system and its equivalent M-band structure 2.3. Fractional Brownian Motion

Fractional Brownian motion $B_H(t)$ is a Gaussian, zero mean, self similar, non-stationary random process with stationary increments [20]. The auto-covariance of the corresponding discrete time process $B_H[n]$ is given by:

$$r_B^H[n_1, n_2] = \frac{\sigma_H^2}{2} (|n_1|^{2H} - |n_1 - n_2|^{2H} + |n_2|^{2H}), \quad (7)$$

where $\sigma_H^2 = var(B_H[1]) = \frac{1}{\Gamma(2H+1)|sin(\pi H)|}$, and H is the self-similarity index, also called as Hurst exponent. The statistical properties of fBm processes are completely characterized by the single parameter H that can be estimated using the maximum likelihood estimation method presented in [21].

3. PROPOSED RWLS LEARNING METHOD

This section presents the proposed RWLS learning method on $(\frac{2}{3}, \frac{1}{3})$ rational wavelet statistically matched to natural images. Learning of separable two-dimensional (2D) rational wavelet is presented that requires learning 1-D RWLS separately matched to the row space and the column space of natural images. The proposed strategy is identical on either the row or the column space. For the sake of readers' ease, let us first consider design for the column space.

3.1. Proposed Learning for the Column Space

Consider the initial architecture of uniformly decimated 3band Lazy wavelet with filters (in Fig. 3(b)):

$$G_i(z) = z^i, \quad i = 0, 1, 2$$

 $F_i(z) = z^{-i}, \quad i = 0, 1, 2.$ (8)

This Lazy wavelet is subsequently transformed to equivalent $\left(\frac{2}{3},\frac{1}{3}\right)$ rational wavelet via (5) and (6). On feeding the vectorized column form of collection of natural images, labeled as x[n], through this rational Lazy wavelet filterbank, following approximate a[n] and detail d[n] subband coefficients are obtained:

$$a[n] = \begin{cases} x \left[\frac{3n}{2}\right] & \text{if } n \text{ is a even} \\ x \left[\frac{3n-1}{2}\right] & \text{if } n \text{ is a odd,} \end{cases}$$
(9)

$$d[n] = x[3n+2]$$
(10)

Next, the lowpass and highpass filters of the *Lazy* rational wavelet structure are lifted via predict and update stage polynomial learned as explained in the following subsections.

3.1.1. Predict stage

We require to predict one branch of samples with the help of the other branch in the predict stage. In rational wavelet structure, this requires the concept of rate converter as proposed in [13] because the output sample rate of two branches is unequal (refer to Fig. 4). Considering the predict polynomial filter T(z) as

 $T(z) = t_0 z + t_1 z^2,$

we obtain

$$d^{new}[n] = d[n] - t_0 x[3n+1] - t_1 x[3n+3],$$

= $x[3n+2] - t_0 x[3n+1] - t_1 x[3n+3].$ (12)

Thus, the choice of T(z) in (11) allows d[n] = x[3n + 2] to be exactly predicted from the neighboring samples. These updated detail coefficients $d^{new}[n]$ can also be viewed as the error in predicting the lower branch samples. Hence, (12) is re-written as

$$e[n] = d^{new}[n] = x[3n+2] - t_0 x[3n+1] - t_1 x[3n+3].$$
(13)

T(z) is learned by minimizing the mean squared prediction error (mse) given by:

$$\zeta[n] = E(e^2[n]) = E(\{x[3n+2] - t_0x[3n+1] - t_1x[3n+3]\}^2),$$
(14)

where E(.) denotes the expectation operator. To minimize





mse, mse vector ζ is we differentiated with respect to t and is equated to zero as:

$$\frac{\partial \boldsymbol{\zeta}}{\partial \mathbf{t}} = 0 - 2E[\mathbf{A}'\mathbf{b}] + 2E[\mathbf{A}'\mathbf{A}]\mathbf{t} = 0$$
$$\implies E[\mathbf{A}'\mathbf{A}]\mathbf{t} = E[\mathbf{A}'\mathbf{b}]. \tag{15}$$

Assuming that the input signal x[n], corresponding to the column space of natural images, belongs to an fBm process, $E[\mathbf{A}'\mathbf{A}]$ and $E[\mathbf{A}'\mathbf{b}]$ are computed using (7) and (15) is solved for t. On simplifying the structure of Fig. 4, the updated equivalent analysis highpass filter, using the learned predict filter T(z), can be written as:

$$G_h^{new}(z) = G_h(z) - \sum_{k=0}^1 G_l(z^{\frac{1}{2}} W_2^{2k}) T(z^{\frac{3}{2} W_2^{3k}}), \quad (16)$$

where $W_k = e^{-j\frac{2\pi}{k}}$. For the update of the corresponding synthesis lowpass filter, the rational wavelet structure is converted to the equivalent 3-band structure and the polyphase matrix $\mathbf{E}(z)$ of the analysis side is computed using $G_0(z)$, $G_1(z)$ and $G_2^{new}(z)$. On applying the condition of perfect reconstruction [22] in (17), the polyphase matrix $\mathbf{R}(z)$ of the synthesis side is computed.

$$\mathbf{R}(z)\mathbf{E}(z) = cz^{-n_0}\mathbf{I},\tag{17}$$

where $c \in \mathbb{R}$, $n_0 \in \mathbb{Z}$, and I is 3×3 identity matrix. From $\mathbf{R}(z)$ and (6), updated filter $F_1^{new}(z)$ of the rational wavelet is computed. This completes the predict stage.

3.1.2. Update Stage

(11)

Next, the update stage filter S(z) shown in Fig.5 is learned. Again, rate converter, shown in Fig.5 as proposed in [13], is required. The reconstructed signal at the upper branch is shown as $x_u[n]$. Since the natural images are generally rich in low frequency content, $x_u[n]$ should be as close as possible to the input signal x[n]. This allows us to learn the update stage filter by minimizing the energy difference of the two signals as below:

$$\tilde{\mathbf{s}} = \min ||\mathbf{x} - \mathbf{x}_u||^2, \tag{18}$$

where $\mathbf{s} \equiv S(z) = s_0 + s_1 z^{-2}$. Signal \mathbf{x}_u can be written in terms of update stage filter \mathbf{s} that allows us to solve (18). Once S(z) is learned, analysis lowpass filter is updated as:

$$G_l^{new}(z) = G_l(z) + G_h(z^2)S(z^3)$$
(19)

Synthesis highpass filter is updated using the method similar to the one used to update the synthesis lowpass filter. This completes the proposed learning. Since the lifting framework is modular, more predict and update stages can be appended to get longer length filterbanks. This is to note that for learning the RWT for the column space of natural images, we vectorized an ensemble of natural images column-wise and stacked them below each other to build a 1-D signal. Next, we estimate the Hurst exponent H of this column vector and learn the RWT as presented above.

3.2. Proposed Design for the Row Space

Corresponding to the row-space design, we vectorize all images row-wise and stack them to build a 1-D signal. Next, we estimate the Hurst exponent H of this row vector and learn the RWT using the method presented in the previous sub-section.



Fig. 6: Images used in the experiment. Img1 to Img10 were used to design the matched rational wavelet system and the performance in CS reconstruction is tested on images Img1, Img4 and Img11.



Fig. 7: Frequency response of analysis side filters matched to the column space of natural images.

4. APPLICATION

The proposed RWLS method is applied on natural images as separable wavelets. The performance of the learned RWLS is compared with standard bi-orthogonal 5/3 and 9/7 wavelets in the application of compressive sensing based reconstruction of natural images of dimension 512×512 . An ensemble of ten natural images shown in Fig. 6 is considered for learning the statistically matched rational wavelet structure for the row space and the column space of images. The value of Hurst exponent is observed to be between 0.5 to 1.0 for all the ten images considered. Fig. 7 shows the frequency response of the analysis side lowpass and highpass filters matched to the column space of natural images.

Bernoulli measurement matrix with entries taken as ± 1 is considered in CS. Since it is computationally expensive to apply CS on big images, we use the concept of block CS [23], where block-size of 32×32 is considered. Recently, multilevel wavelet decomposition has been proposed over L-shaped pyramid (L-Pyramid) (Fig. 8b) in [24] and is observed to perform better in CS application compared to the existing multilevel regular pyramid (R-Pyramid) wavelet decomposition (Fig. 8a). We decompose our input images to 3-level using this new L-Pyramid wavelet decomposition in our experiments. Table-I presents reconstruction results in terms of PSNR (peak signal to noise ratio) for sampling ratios varying from 90% to 30%, where sampling ratio is the percentage of total samples measured.

From Table-1, we note that the performance of the proposed RWLS is superior (comparable at 90% for Img1 and Img4) to standard wavelets on natural images. Although image 'Img11' was not used in the ensemble of images used to

Table 1: CS based reconstruction results of natural images						
		PSNR (in dB) over different				
Image	Wavelet	sampling ratios				
		90%	70%	50%	30%	
Img1	5/3	34.83	31.22	26.94	22.13	
	9/7	35.30	32.86	29.54	25.44	
	RWLS	34.86	33.37	30.82	26.55	
Img4	5/3	38.83	35.38	31.46	27.18	
	9/7	39.36	36.78	33.35	29.53	
	RWLS	39.07	37.11	34.58	30.63	
Img11	5/3	39.51	37.41	33.57	27.48	
	9/7	39.74	38.59	35.66	30.97	
	RWLS	40.24	39.50	37.44	32.63	

LL ₃ LH ₃ HL ₃ HH ₃	LH ₂	14
HL₂	HH₂	
ŀ	IL₁	ΗH1

$\begin{array}{c c} L_{3}L_{3} & L_{3}H_{3} & L_{3}H_{2} & L_{3}H_{1} \\ \hline H_{3}L_{3} & H_{3}H_{3} & H_{3}H_{2} & H_{3}H_{1} \\ \hline H_{2}L_{3} & H_{2}H_{3} & H_{2}H_{2} & H_{2}H_{1} \\ \hline H_{1}L_{3} & H_{1}H_{3} & H_{1}H_{2} & H_{1}H_{1} \\ \hline \end{array}$				
$\begin{array}{cccccccccccccccccccccccccccccccccccc$	L ₃ L ₃	L₃H₃	L_3H_2	L_3H_1
$H_{2}L_{3}H_{2}H_{3}$ $H_{2}H_{2}$ $H_{2}H_{1}$ $H_{1}L_{3}H_{1}H_{3}$ $H_{1}H_{2}$ $H_{1}H_{1}$	H ₃ L ₃	H₃H₃	H ₃ H ₂	H ₃ H ₁
$H_1L_3H_1H_3$ H_1H_2 H_1H_1	H ₂ L ₃	H ₂ H ₃	H ₂ H ₂	H_2H_1
	H ₁ L ₃	H ₁ H ₃	H ₁ H ₂	H ₁ H ₁

(a) 3-level R-Pyramid Wavelet Decomposition

elet (b) 3-level L-Pyramid Wavelet Decomposition

Fig. 8: 3-level Wavelet Decomposition Notations for (b): First Letter- wavelet band of columns (Lowpass or <u>Highpass</u>); Second Letter-wavelet band of rows (Lowpass or <u>Highpass</u>); Subscripts 1, 2, and 3 denote the number of times wavelet has been applied on that row/column

learn the RWT, the performance of the learned RWLS over this image is also superior indicating that the proposed learning indeed provides the statistically-matched rational system for the class of natural images.

5. CONCLUSION

Statistical learning for rational wavelet transform (RWLS) method for natural images is presented in this work. The natural images are modeled as fBm processes and their statistical properties are used to learn separable rational wavelet transform. Lifting framework for the rational wavelet is used in the proposed work that provides closed-form solution for learning making the method computationally efficient. The learned rational wavelet transform is tested in the application of CS based reconstruction of natural images and is observed to perform better compared to the existing standard bi-orthogonal wavelet transforms.

6. REFERENCES

- S. Ravishankar, B. Wen, and Y. Bresler, "Online sparsifying transform learningpart i: Algorithms," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 625–636, 2015.
- [2] B. Wen, S. Ravishankar, and Y. Bresler, "Video denoising by online 3d sparsifying transform learning," in *Image Processing (ICIP), 2015 IEEE International Conference on.* IEEE, 2015, pp. 118–122.
- [3] S. Ravishankar and Y. Bresler, "Efficient blind compressed sensing using sparsifying transforms with convergence guarantees and application to magnetic resonance imaging," *SIAM Journal on Imaging Sciences*, vol. 8, no. 4, pp. 2519–2557, 2015.
- [4] S. Mallat, A wavelet tour of signal processing. Academic press, 1999.
- [5] T. Blu, "An iterated rational filter bank for audio coding," in *Time-Frequency and Time-Scale Analysis*, 1996., Proceedings of the IEEE-SP International Symposium on. IEEE, 1996, pp. 81–84.
- [6] —, "Iterated filter banks with rational rate changes connection with discrete wavelet transforms," *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3232–3244, 1993.
- [7] O. Chertov, V. Malchykov, and D. Pavlov, "Non-dyadic wavelets for detection of some click-fraud attacks," in *Signals and Electronic Systems (ICSES), 2010 International Conference on.* IEEE, 2010, pp. 401–404.
- [8] T.-T. Le, M. Ziebarth, T. Greiner, and M. Heizmann, "Optimized size-adaptive feature extraction based on content-matched rational wavelet filters," in *Signal Processing Conference (EUSIPCO)*, 2014 Proceedings of the 22nd European. IEEE, 2014, pp. 1672–1676.
- [9] I. Bayram and I. W. Selesnick, "Frequency-domain design of overcomplete rational-dilation wavelet transforms," *Signal Processing, IEEE Transactions on*, vol. 57, no. 8, pp. 2957–2972, 2009.
- [10] —, "Overcomplete discrete wavelet transforms with rational dilation factors," *Signal Processing, IEEE Transactions on*, vol. 57, no. 1, pp. 131–145, 2009.
- [11] A. Baussard, F. Nicolier, and F. Truchetet, "Rational multiresolution analysis and fast wavelet transform: application to wavelet shrinkage denoising," *Signal Processing*, vol. 84, no. 10, pp. 1735–1747, 2004.
- [12] S. T. N. Nguyen and B. W.-H. Ng, "Bi-orthogonal rational discrete wavelet transform with multiple regularity

orders and application experiments," *Signal Processing*, vol. 93, no. 11, pp. 3014–3026, 2013.

- [13] N. Ansari and A. Gupta, "Lifting-based rational wavelet design from a given signal," in *Digital Signal Processing* (*DSP*), 2015 IEEE International Conference on. IEEE, 2015, pp. 853–857.
- [14] A. P. Pentland, "Fractal-based description of natural scenes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6, pp. 661–674, 1984.
- [15] B. B. Mandelbrot and J. W. Van Ness, "Fractional brownian motions, fractional noises and applications," *SIAM review*, vol. 10, no. 4, pp. 422–437, 1968.
- [16] C. V. Stewart, B. Moghaddam, K. J. Hintz, and L. M. Novak, "Fractional brownian motion models for synthetic aperture radar imagery scene segmentation," *Proceedings of the IEEE*, vol. 81, no. 10, pp. 1511–1522, 1993.
- [17] C.-C. Chen, J. S. Daponte, and M. D. Fox, "Fractal feature analysis and classification in medical imaging," *Medical Imaging, IEEE Transactions on*, vol. 8, no. 2, pp. 133–142, 1989.
- [18] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *Journal of Fourier analysis and applications*, vol. 4, no. 3, pp. 247–269, 1998.
- [19] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Applied and computational harmonic analysis*, vol. 3, no. 2, pp. 186– 200, 1996.
- [20] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing.* Artech House Norwood, 2005, vol. 46.
- [21] T. Lundahl, W. J. Ohley, S. M. Kay, and R. Siffert, "Fractional brownian motion: A maximum likelihood estimator and its application to image texture," *Medical Imaging, IEEE Transactions on*, vol. 5, no. 3, pp. 152– 161, 1986.
- [22] P. P. Vaidyanathan, *Multirate systems and filter banks*. Pearson Education India, 1993.
- [23] L. Gan, "Block compressed sensing of natural images," in Digital Signal Processing, 2007 15th International Conference on. IEEE, 2007, pp. 403–406.
- [24] N. Ansari and A. Gupta, "Image reconstruction using matched wavelet estimated from data sensed compressively using partial canonical identity matrix," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 3680–3695, Aug 2017.