Statistical Learning of Rational Wavelet Transform for Natural Images

Naushad Ansari and Anubha Gupta
{naushada, anubha}@iiitd.ac.in
Signal Processing and Bio-Medical Imaging Lab (SBILab)
Deptt. of ECE, IIIT-Delhi, India

Objective

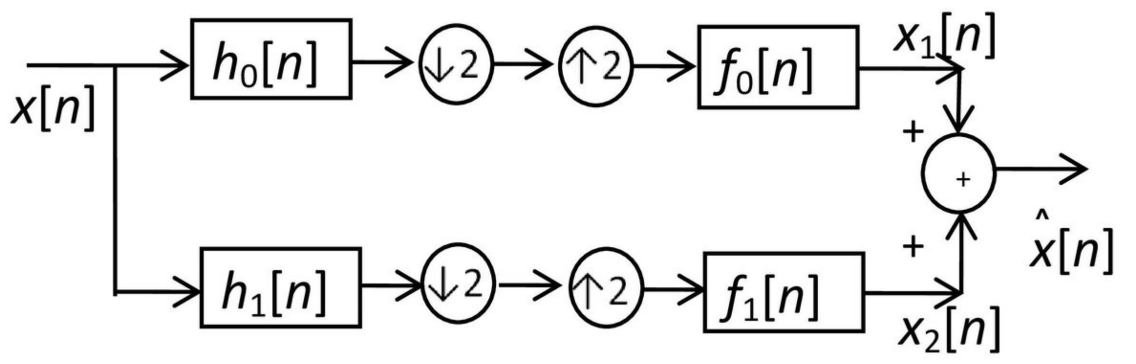
Motivated with the concept of transform learning and the utility of rational wavelet transform in audio and speech processing, the objective of this work is to propose Rational Wavelet Transform Learning in the Statistical sense (RWLS) for natural images. The learned rational wavelet is used as the sparsifying transform for CS based reconstruction of natural images.

Motivation

- Wavelets are used extensively as sparsifying basis for images.
- Rational wavelets provide non-uniform frequency band representation.
- Learned rational wavelet may provide better reconstruction results for images.
- Lifting framework, extended to rational wavelets in [1] can be used to learn signal-matched rational wavelets.

Lifting[2]

General dyadic wavelet



Lifting framework consists of three steps:

- Split: input into even and odd indexed samples.
- Predict: Predict one subband samples from the other

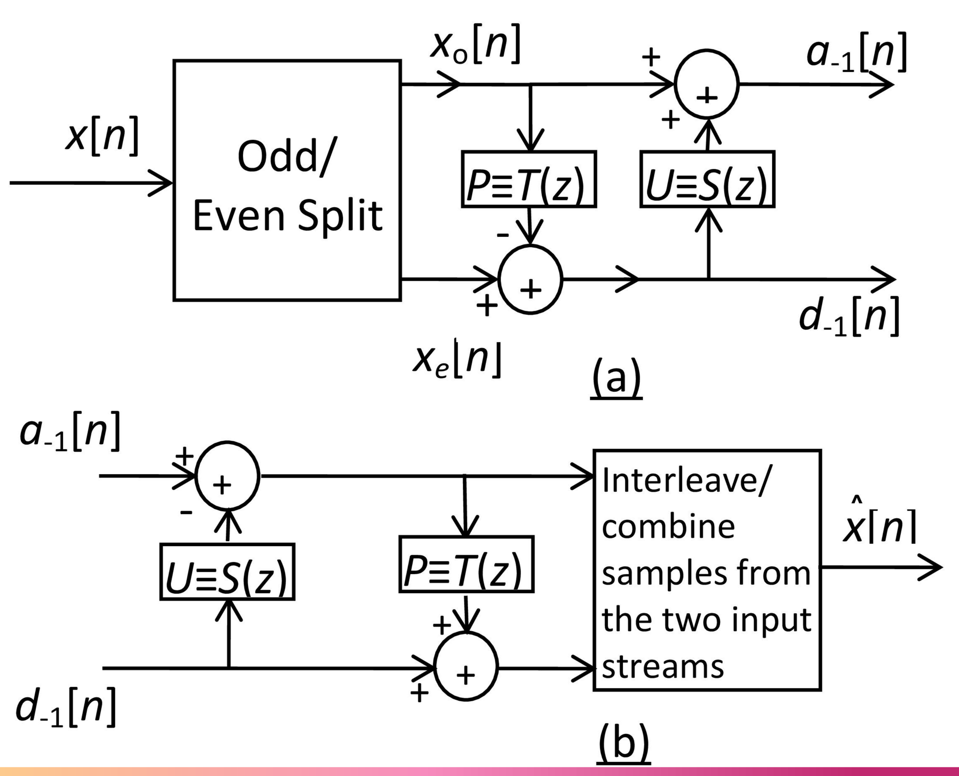
  Correspondingly, update analysis highpass and synthesis lowpass filters using:

$$H_1^{new}(z) = H_1(z) - H_0(z)T(z^2)$$
$$F_0^{new}(z) = F_0(z) + F_1(z)T(z^2) \qquad (1)$$

- Update: Update the other subband samples with the predicted subband samples

  Correspondingly, update analysis lowpass and synthesis highpass filters using:

$$H_0^{new}(z) = H_0(z) + H_1(z)S(z^2)$$
$$F_1^{new}(z) = F_1(z) - F_0(z)S(z^2) \qquad (2)$$



Fractional Brownian Motion

Fractional Brownian motion $B_H(t)$ is a Gaussian, zero mean, self similar, non-stationary random process with stationary increments[3]. The auto-covariance of the corresponding discrete time process $B_H[n]$ is given by:

$$r_B^H[n_1,n_2] = \frac{\sigma_H^2}{2}(|n_1|^{2H} - |n_1-n_2|^{2H} + |n_2|^{2H}) \qquad (3)$$

where H is the self-similarity index and $\sigma_H^2 = var(B_H[1]) = \frac{1}{\Gamma(2H+1)|sin(\pi H)|}$.

Compressive Sensing[3]

Mathematically: $\mathbf{y}_{m\times 1} = \mathbf{A}_{m\times n}\mathbf{x}_{n\times 1} + \boldsymbol{\eta}_{n\times 1}, \quad m < n$

Can be solved using following optimization framework[4]:

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2 \quad \text{subject to } \|\mathbf{W}\mathbf{x}\|_1 \le \tau.$$

REFERENCES

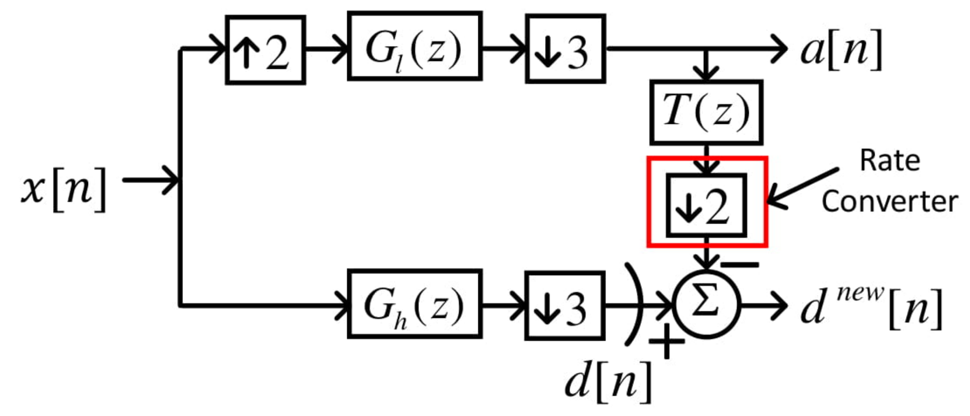
1. Naushad Ansari and Anubha Gupta, "M-RWTL: Learning Signal-Matched Rational Wavelet Transform in Lifting Framework," in *IEEE Access,* no.99, pp. 1-14, 2018.
2. W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," Applied and computational harmonic analysis, vol. 3, no. 2, pp. 186–200, 1996.
3. D. L. Donoho, "Compressed sensing," Information Theory, IEEE Transactions on, vol. 52, no. 4, pp. 1289–1306, 2006.
4. S. Chen and D. Donoho, "Basis pursuit," in Signals, Systems and Computers, 1994. 1994 Conference Record of the Twenty-Eighth Asilomar Conference on, vol. 1. IEEE, 1994, pp. 41–44.
5. Naushad Ansari and Anubha Gupta, "Image Reconstruction using Matched Wavelet Estimated from Data Sensed Compressively using Partial Canonical Identity Matrix," *IEEE Transactions on Image Processing (IEEE-TIP)*, vol. 26, no. 8, pp. 3680-3695, 2017.

Proposed Method

Learn statistically matched rational wavelet transform for column space and row space of natural images as follows:

Rational Lazy Wavelet: Start with rational Lazy wavelet: $G_i(z) = z^i, F_i(z) = z^{-i};\ i = 0,1,2$

Predict stage: Use rate converter to obtain and $T(z) = t_0z + t_1z^2$ to obtain the prediction error $e[n] = d^{new}[n] = x[3n+2] - t_0x[3n+1] - t_1x[3n+3]$.



Learn $T(z)$ by minimizing least square prediction error:

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

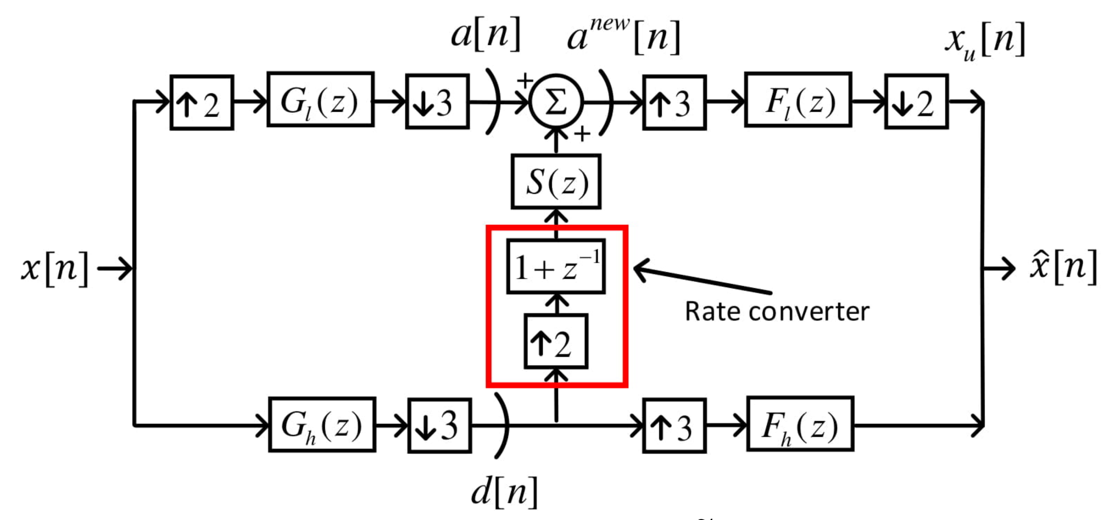
Differentiating and equating to zero: $\frac{\partial \zeta}{\partial \mathbf{t}} = 0 - 2E[\mathbf{A}'\mathbf{b}] + 2E[\mathbf{A}'\mathbf{A}]\mathbf{t} = 0$

$$\Rightarrow E[\mathbf{A}'\mathbf{A}]\mathbf{t} = E[\mathbf{A}'\mathbf{b}]$$

$E[\mathbf{A}'\mathbf{A}]$ and $E[\mathbf{A}'\mathbf{b}]$ are computed using (3). Update analysis highpass filter as follows:

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

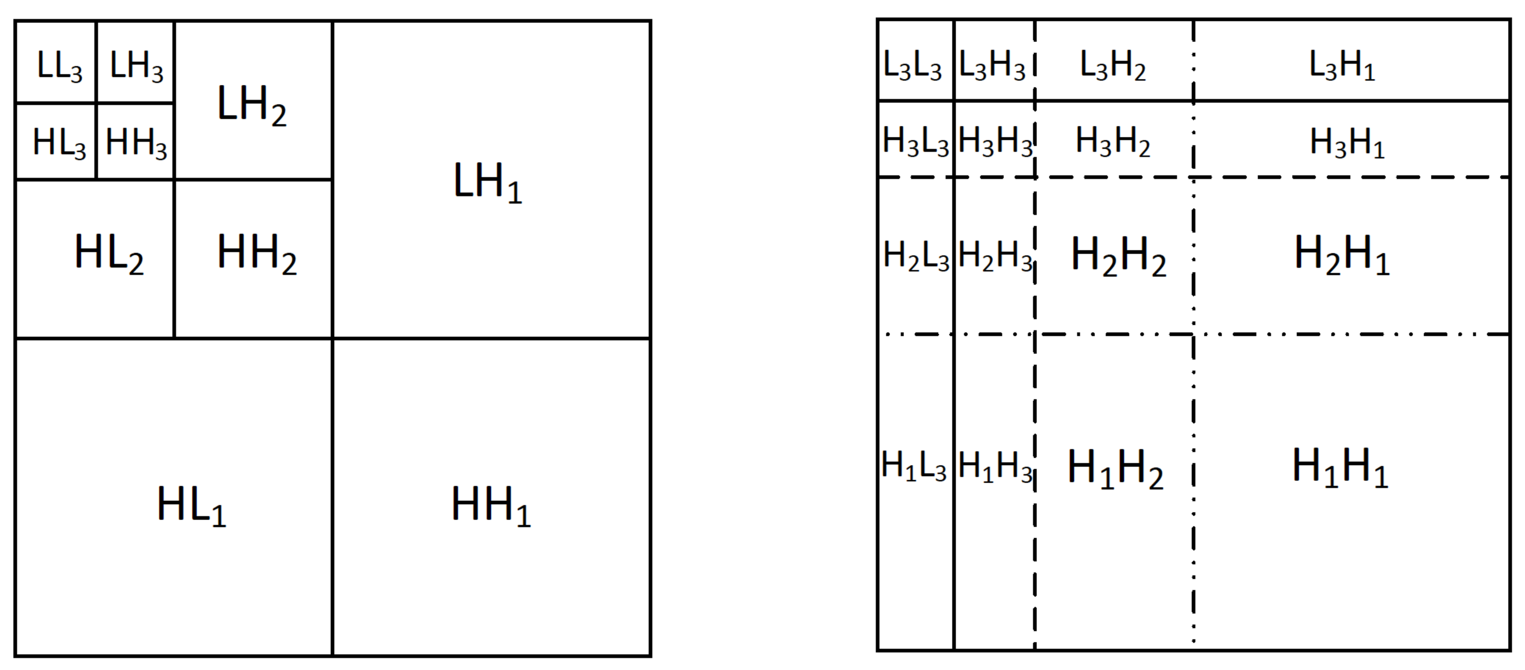
Update stage: Use rate converter and $S(z) = s_0 + s_1z^{-2}$.



As images are rich in low frequency. Learn S(z) by solving: $\tilde{\mathbf{s}} = \min_{\mathbf{s}} \|\mathbf{x} - \mathbf{x}_u\|^2$

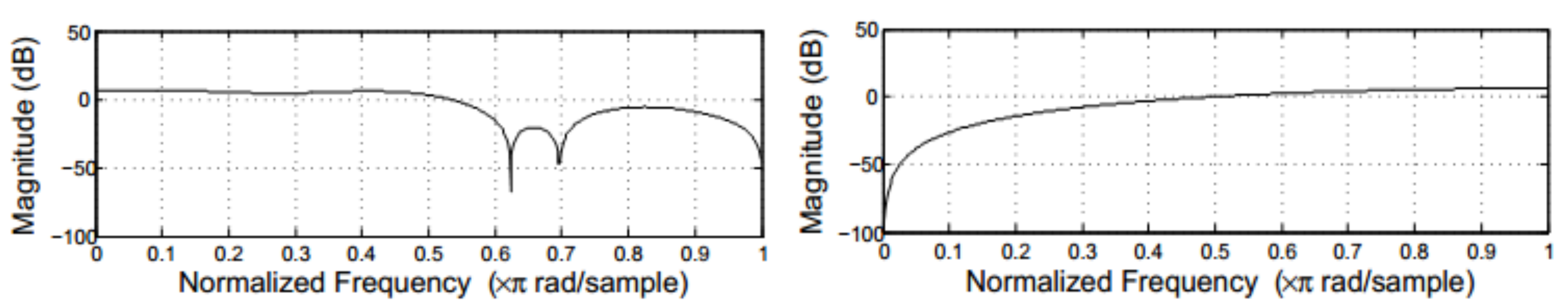
Update analysis lowpass filter as: $G_l^{new}(z) = G_l(z) + G_h(z^2)S(z^3)$

Results



(a) 3-level R-Pyramid Wavelet Decomposition (b) 3-level R-Pyramid Wavelet Decomposition[5]

Frequency response of learned filters:



a) Analysis lowpass filter b) Analysis highpass filter

Compressive sensing based reconstruction results of natural images:

Table 1: CS based reconstruction results of natural images

Image	Wavelet	PSNR (in dB) over different 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**

Some natural images used in experiments

The first author would like to thank Council of Scientific & Industrial Research(CSIR), Govt. of India, for financial support in research.