

1 Single Scale Adaptive Wavelet Tight Frames

As a first step of the plan, we plan to construct wavelet tight frames in a manner that is adapted to the given data. The resulting model will be a building block of the multi-layer structure developed in later sections. We begin with a brief review of existing construction of wavelet tight frames.

1.1 Wavelet Tight Frames

In this subsection, we give a very brief introduction to wavelet tight frames, for a detailed introduction to this subject, the readers may refer to [1].

Let \mathcal{H} be a Hilbert space, we are mainly concerned with the case when $\mathcal{H} = L_p(\mathbb{R}^d)$. a system $X \subset \mathcal{H}$ is called a tight frames is

$$\|f\|_2^2 = \sum_{x \in X} |\langle f, x \rangle|^2, \quad \text{for any } f \in \mathcal{H}$$

There are two operators that are associated with a tight frame, one is the analysis operator defined as

$$W : f \in \mathcal{H} \rightarrow \{\langle x, f \rangle\}_{x \in X} \in l_2(\mathbb{N})$$

and the other is the synthesis operator W^T , which is the adjoint operator of the analysis operator defined as

$$W^T : \{a_n\} \in l_2(\mathbb{N}) \rightarrow \sum_{x \in X} a_n x \in \mathcal{H}.$$

The system X is called a tight frame if and only if $W^T W = I$, where $I : \mathcal{H} \rightarrow \mathcal{H}$ is the identity operator. In other words, given a tight frame X , we have the following canonical expansion:

$$f = \sum_{x \in X} \langle x, f \rangle x, \quad \text{for any } f \in \mathcal{H}$$

The sequence $Wf := \{\langle f, x \rangle\}_{x \in X}$ are called the canonical tight frame coefficients. Thus, tight frames are often viewed as generalizations of orthonormal basis. In fact, a tight frame X is an orthonormal basis for \mathcal{H} if and only if $\|x\| = 1, \forall x \in X$.

In signal processing applications, one widely used class of tight frames is the wavelet tight frames. The construction starts with a finite set of generators $\Psi := \{\psi^1, \dots, \psi^m\}$. Then consider the affine system defined by the shifts and dilations of the generators:

$$X(\Psi) = \{M^{j/2} \psi^l(M^j \cdot -k), 1 \leq l \leq m, j, k \in \mathbb{Z}, M \in \mathbb{Z}^+\}.$$

The affine system $X(\Psi)$ is called a wavelet tight frame if it is a tight frame satisfying

$$f = \sum_{x \in X(\Psi)} \langle f, x \rangle x, \forall f \in \mathcal{H}.$$

Wavelet tight frames used in practice are usually constructed from multi-resolution analysis(MRA). This is because a MRA structure is crucial for fast decomposition and reconstruction algorithms. The MRA construction usually starts with a compactly supported scaling function ϕ with a refinement mask a_0 satisfying

$$\hat{\phi}(M \cdot) = \hat{a}_0 \hat{\phi}.$$

where $\hat{\phi}$ is the Fourier transform of ϕ , and \hat{a}_0 is the discrete Fourier series defined as $\hat{a}_0(\omega) := \sum_{k \in \mathbb{Z}} a_0(k) e^{-ik\omega}$ and $\hat{a}_0(0) = 1$. After obtaining the scaling function, the next step is to find an appropriate set of filters $\{a_1, \dots, a_m\}$ and define the set of functions called framelets $\Psi = \{\psi_1, \dots, \psi_m\}$ by

$$\hat{\psi}_i(M \cdot) = \hat{a}_i \hat{\phi}, i = 1, \dots, m$$

such that the affine system $X(\Psi)$ forms a wavelet tight frame. It is natural to ask when does such a system form a wavelet tight frame. Sufficient and necessary conditions are given by the so called Unitary Extension Principle(UEP). There are different versions of UEP principles, we are only concerned with

one version that is associated with discrete wavelet tight frames, which we will state after a description of the decomposition and reconstruction operations. For a survey of UEP, the reader may refer to [1].

Given a filter $a \in l_2(\mathbb{Z})$, the discrete decomposition and reconstruction transform are defined in the following way. Define the one dimensional down-sampling and up-sampling operator:

$$\begin{aligned} (n) &:= v(Mn), \quad n \in \mathbb{Z} \\ [v \uparrow](Mn) &:= v(n), \quad n \in \mathbb{Z} \end{aligned}$$

where M a positive integer denoting the down-sampling or up-sampling factor. Down-sampling and up-sampling operators in higher dimensions are carried out by performing one dimensional operators along each dimension.

Define the linear convolution operator $S_a : l_2(\mathbb{Z}) \rightarrow l_2(\mathbb{Z})$ by

$$[S_a(v)](n) := [a * v](n) = \sum_{k \in \mathbb{Z}} (a(-\cdot) * v \downarrow)M, \forall v \in l_2(\mathbb{Z})$$

For a set of filters $\{a_i\}_{i=1}^m \subset l_2(\mathbb{Z})$, we define its analysis operator W by

$$W = [S_{a_1}(-\cdot), S_{a_2}(-\cdot), \dots, S_{a_m}(-\cdot)]^T.$$

Its synthesis operator is defined as the transpose of W :

$$W^T = [S_{a_1}(\cdot), S_{a_2}(\cdot), \dots, S_{a_m}(\cdot)].$$

Now we may state the UEP for this situation.

Proposition 1 (cite). *Let a_1, \dots, a_m be finitely supported filters, the following are equivalent:*

1. $W_a^T W_a = I$

2. for all $\omega \in [0, 1)^d \cup M^{-1}\mathbb{Z}^d$,

$$\sum_{i=1}^m \hat{a}_i \overline{\hat{a}_i(\xi + 2\pi\omega)} = \delta(\omega);$$

3. for all $k, \gamma \in \mathbb{Z}^d$,

$$\sum_{i=1}^m \sum_{n \in \mathbb{Z}^d} \overline{a_i(k + Mn + \gamma)} a_i(Mn + \gamma) = M^{-d} \delta(k).$$

In particular, if the data are real numbers and no down-sampling is performed, then $W^T W = I$ is equivalent to

$$\sum_{i=1}^m \sum_{n \in \mathbb{Z}^d} a_i(k + n) a_i(n) = \delta_k, \forall k \in \mathbb{Z}. \quad (1)$$

The linear B-spline wavelet tight frame used in many image restoration tasks is constructed via the UEP. Its associated tree filters are :

$$a_1 = \frac{1}{4}(1, 2, 1)^T; \quad a_2 = \frac{\sqrt{2}}{4}(1, 0, -1)^T; \quad a_3 = \frac{1}{4}(-1, 2, -1)^T.$$

Once the 1D filter $\{a_i\}_{i=1}^m$ for generating a tight frame for $l_2(\mathbb{Z})$ is constructed, the traditional way of generating higher dimensional tight frames is to use tensor products of 1D filters. But in this paper, we are going to construct 2D filters directly.

1.2 Adaptive Construction

Let W_a be the matrix form of the analysis operator generated by the filters $\{a_i\}_{i=1}^m$, and let W_a^T be the matrix form of the synthesis operator. Define $\mathcal{C} = \{\{a_i\}_{i=1}^m : \{a_i\}_{i=1}^m \text{ satisfies the full UEP condition}\}$, in other words, $\mathcal{C} = \{\{a_i\}_{i=1}^m : W_a^T W_a = I\}$. We also consider the filters that approximately satisfy the full UEP condition. $\mathcal{C}_\delta = \{\{a_i\}_{i=1}^m : \|W_a^T W_a - I\| \leq \delta\}$. For a particular class of images, there are many wavelet tight frames for us to choose from. Although they all provide perfect reconstruction of the images, some of them may yield sparser representations over the rest. Therefore, we propose the following model to find the wavelet tight frame that provides the sparsest representation:

$$\min_{a_1, \dots, a_m} \sum_j \|a_j(\cdot) * x\|_0 \quad \text{subject to } \{a_i\}_{i=1}^m \text{ satisfies UEP condition (1)} \quad (2)$$

For signals with noise, it is natural to consider the modification:

$$\min_{a_1, \dots, a_m, v^1, \dots, v^m} \sum_{j=1}^m \|v^j - a_j(\cdot) * x\|_2^2 + \lambda \sum_{j=1}^m \|v^j\|_1 \quad \text{s.t. } \{a_i\}_{i=1}^m \text{ satisfies UEP condition (1)} \quad (3)$$

Another modification is also reasonable, in fact, may yield better results sometimes.

$$\min_{a_1, \dots, a_m} \sum_j \|a_j(\cdot) * x\|_1 \quad \text{s.t. } \{a_i\}_{i=1}^m \text{ satisfies UEP condition (1)} \quad (4)$$

In the following, we mainly focus on (4) and its modifications. A special case of (3) is also considered. Before we proceed to the algorithms, we make a few comments about this model.

As both models involve optimization over both a and v and is non-convex, it is typically solved using alternating directions. As a result, generally, global minimum may not be obtained. Surprisingly, the local minimum obtained by alternating directions is often good enough in sparse coding and dictionary learning models.