ORTHOGONAL WAVELET FRAMES AND VECTOR-VALUED WAVELET TRANSFORMS

GHANSHYAM BHATT, BRODY DYLAN JOHNSON, AND ERIC WEBER

Abstract. Motivated by the notion of orthogonal frames, we describe sufficient conditions for the construction of orthogonal MRA wavelet frames in $L^2(\mathbb{R})$ from a suitable scaling function. These constructions naturally lead to filter banks in $\ell^2(\mathbb{Z})$ with similar orthogonality relations and, through these filter banks, the orthogonal wavelet frames give rise to a vector-valued discrete wavelet transform (VDWT). The novelty of these constructions lies in their potential for use with vector-valued data, where the VDWT seeks to exploit correlation between channels. Extensions to higher dimensions are natural and the constructions corresponding to the bidimensional case are presented along with preliminary results of numerical experiments in which the VDWT is applied to color image data.

1. Introduction

Wavelets are an important tool in a variety of data processing applications including compression, smoothing, and feature detection. Frequently, the data to be processed is multicomponent in nature in the sense that several data values may correspond to a common location in space or time. It is natural to view such multicomponent data as being *vector-valued*. Sources of such vector-valued data include digitized stereo audio and color images, and wireless multiuser communication. Moreover, sampling of vector and tensor fields in medical imaging or geophysics give rise to true vector-valued data.

It is reasonable to expect that an advantage could be achieved by utilizing a wavelet transform which is inherently vector-valued, in correspondence with the nature of the data. A number of approaches to vector-valued wavelet transforms exist in the literature [20, 7, 11]. The defining feature of these wavelet transforms is that they require matrix coefficients in the filterbanks. Wavelet transforms using multiwavelets is a specific example of such an approach. Multiwavelets can be constructed with more flexibility than traditional scalar wavelets, giving rise to wavelets with desirable properties that are not possible with scalar wavelets [12, 6, 16, 5]. However, the refinement equation for multiwavelets involves matrix coefficients, because the scaling functions interact with each other. Because of the matricial nature of the filterbanks, it has been found [19] that prefiltering is required to fully utilize the previously mentioned advantages of multiwavelets. Moreover, the matrix equations involved can be difficult to handle.

We present in this paper a vector-valued wavelet transform built on filterbanks that have scalar coefficients, and does not have a prefiltering stage. The construction presented here leads to a vector-valued discrete wavelet transform (VDWT), which arises from a number of

Date: November 1, 2006.

2000 *Mathematics Subject Classification.* Primary: 42C40.

Key words and phrases. wavelet, frame, orthogonal frame, discrete wavelet transform, vector valued data.

wavelet frames satisfying certain orthogonality relationships. A more technical description of the orthogonality relationships will be given in the subsection entitled "Motivation" below.

There are two defining features of the VDWT presented here. The first is that the filterbanks can be easily constructed from ordinary scalar wavelets or filter banks and thus is very flexible. The second is that the refinement equation associated to the VDWT can be regarded as having matrix coefficients which are *diagonal*. Therefore, there is no need to solve a complicated matrix valued refinement equation.

The outline of the paper is as follows. The remainder of this section is devoted to background material regarding frames and wavelets: definitions and results from the literature; a detailed description of the motivation; and a short description of the main results. The second section develops in detail the VDWT in one spatial dimension by presenting results and algorithms for constructing wavelet frames and the corresponding filter banks with the necessary orthogonality properties. The third section then mimics the second section in the setting of two spatial dimensions. The third section also includes a discussion of preliminary numerical results obtained by applying the VDWT constructed in this paper to color image data.

1.1. **Definitions.** Frames for (separable) Hilbert spaces were introduced by Duffin and Schaeffer [10] in their work on non-harmonic Fourier series. Later, Daubechies, Grossmann, and Meyer revived the study of frames in [8], and since then, frames have become the focus of active research in both theory and applications.

Let H be a separable Hilbert space and $\mathbb{J}$ a countable index set. A sequence $\mathbb{X} := \{x_j\}_{j\in\mathbb{J}} \subset H$ is a *frame* if there exist positive real numbers C_1, C_2 such that for all $v \in H$,

$$C_1\|v\|^2 \leq \sum_{j\in\mathbb{J}} |\langle v, x_j\rangle|^2 \leq C_2\|v\|^2. \tag{1}$$

If $\mathbb{X}$ satisfies the second inequality, then $\mathbb{X}$ is called a *Bessel sequence*, or simply *Bessel.* Given $\mathbb{X}$ which is Bessel, define the analysis operator

$$\Theta_{\mathbb{X}} : H \to l^2(\mathbb{J}) : v \mapsto (\langle v, x_j\rangle)_j;$$

and the synthesis operator

$$\Theta^*_{\mathbb{X}} : l^2(\mathbb{J}) \to H : (c_j)_j \mapsto \sum_{j\in\mathbb{J}} c_j x_j.$$

The analysis operator is well-defined and bounded by the frame inequality (1). Additionally, the sum $\sum_j c_j x_j$ converges (see [10]), and so the synthesis operator is also well-defined and bounded, and a simple computation shows that it is in fact the Hilbert space adjoint operator of the analysis operator.

If the Bessel sequence $\mathbb{X}$ satisfies the condition that $\Theta^*_{\mathbb{X}}\Theta_{\mathbb{X}} = I$, we say that $\mathbb{X}$ is a *Parseval frame*. This condition also goes by the names of *normalized tight* or *1-tight* frames.

Given two Bessel sequences $\mathbb{X}$ and $\mathbb{Y} := \{y_j\}_{j\in\mathbb{J}}$, define the operator

$$\Theta^*_{\mathbb{Y}}\Theta_{\mathbb{X}} : H \to H : v \mapsto \sum_{j\in\mathbb{J}} \langle v, x_j\rangle y_j;$$

this operator is sometimes called a "Mixed Dual Gramian". Note that again by the Bessel condition of both sequences, it is a well-defined and bounded operator. Typically in frame theory, one wants the above operator to be the identity; if this is the case, then the Bessel sequences $\mathbb{X}$ and $\mathbb{Y}$ are actually frames and are called *dual frames*. Our motivation here is for this operator to be the 0 operator.

Definition 1.1.1. Suppose $\mathbb{X}$ and $\mathbb{Y}$ are Bessel sequences in H and K, respectively, and both are indexed by $\mathbb{J}$. If

$$\Theta_{\mathbb{Y}}^*\Theta_{\mathbb{X}} := \sum_{j\in\mathbb{J}} \langle\cdot, x_j\rangle y_j = 0,$$

the Bessel sequences are said to be *orthogonal*.

This idea has been studied by Han and Larson [13], where the Bessel sequences were assumed to be frames and were called strongly disjoint, and also by Balan, et. al. in [1] and [2] for the Gabor (Weyl-Heisenberg) frame case.

Definition 1.1.2. For the purposes of this paper, we will define the Fourier transform for $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ to be

$$\hat{f}(\xi) = \int f(x)e^{-2\pi i x\cdot\xi}dx.$$

We shall consider the affine system in d spatial dimensions using dilation by 2 times the identity matrix:

$$D : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) : f(\cdot) \mapsto \sqrt{2}^d f(2\cdot).$$

For $\alpha \in \mathbb{R}^d$, let T_α denote the unitary translation operator

$$T_\alpha : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) : f(\cdot) \mapsto f(\cdot - \alpha).$$

Definition 1.1.3. If $\{\psi_1, \dots, \psi_r\} \subset L^2(\mathbb{R}^d)$, the affine system generated by $\{\psi_1, \dots, \psi_r\}$ is the collection $\{D^nT_l\psi_k : n \in \mathbb{Z};\ l \in \mathbb{Z}^d;\ k = 1, \dots, r\}$. We shall say that $\{\psi_1, \dots, \psi_r\}$ generates a wavelet frame if the affine system generated by it is a frame for $L^2(\mathbb{R}^d)$. We say $\{\psi_1, \dots, \psi_r\}$ generates an affine Bessel system if the affine system generated by it is a Bessel sequence in $L^2(\mathbb{R}^d)$.

Definition 1.1.4. By a filter we mean an element of $L^\infty([0, 1)^d)$; i.e. m is a filter if $m \in L^\infty([0, 1)^d)$. We shall call m a low-pass filter if $m(0) = 1$, and we shall call m a high-pass filter if $m(0) = 0$. Though not necessary, we will assume that every filter is continuous on a neighborhood of 0, so there will be no ambiguity in these definitions.

1.2. **Motivation.** The main motivation of the present paper is to construct a wavelet transform for vector-valued data via orthogonal wavelet frames. Suppose we have a function (signal) on $\mathbb{R}$ taking values in a finite dimensional space–for convenience, say $\mathbb{C}^N$. The function f can be identified with its scalar valued components: $f = (f_1, \dots, f_N)$. Supposing that the coordinate functions are measurable and square integrable, then $f \in L^2(\mathbb{R}) \oplus \cdots \oplus L^2(\mathbb{R})$ (with N summands). The standard wavelet transform algorithm for f is to take a wavelet basis $\{D^nT_l\psi : n, l \in \mathbb{Z}\}$ for $L^2(\mathbb{R})$ and perform the wavelet decomposition on each component. Thus, the wavelet transform is

$$\Theta_\psi : \bigoplus_{k=1}^{N} L^2(\mathbb{R}) \to \bigoplus_{k=1}^{N} l^2(\mathbb{Z}^2) : f \mapsto (\langle f_1, D^nT_l\psi\rangle, \dots, \langle f_N, D^nT_l\psi\rangle)_{n,l}.$$

Instead, we propose to do the following: construct N frame wavelets $\psi_1, \dots, \psi_N$ such that each wavelet frame is Parseval and any two wavelet frames are orthogonal. Then, denoting $\Psi = (\psi_1, \dots, \psi_N)$, the wavelet transform becomes

$$\Theta_\Psi : \bigoplus_{k=1}^{N} L^2(\mathbb{R}) \to l^2(\mathbb{Z}^2) : f \mapsto (\sum_{k=1}^{N} \langle f_k, D^n T_l \psi_k \rangle)_{n,l}. \tag{2}$$

So, we are doing N wavelet transforms, each one distinct on the components of f, and then summing the outputs of those transforms based on scale and translation. The rationale for doing this is the following theorem:

Theorem 1.2.1. *Let H_k be a Hilbert space for $k = 1, \dots, N$; and suppose $\{x_j^k\}_{j\in\mathbb{J}} \subset H_k$ is a Bessel sequence for each k. Then the sequence $\{x_j^1 \oplus \cdots \oplus x_j^N\}_{j\in\mathbb{J}}$ is a Parseval frame for $H_1 \oplus \cdots \oplus H_N$ if and only if the following two conditions hold:*

1. *for each k, $\{x_j^k\}_{j\in\mathbb{J}}$ is a Parseval frame for H_k;*
2. *for $k, l = 1, \dots, N$ with $k \neq l$, $\{x_j^k\}_{j\in\mathbb{J}}$ and $\{x_j^l\}_{j\in\mathbb{J}}$ are orthogonal.*

Proof. This is a known result, but we include the proof for completeness, since it is fundamental to the construction of the VDWT. For each coordinate k of $H_1 \oplus \cdots \oplus H_N$, let P_k denote the orthogonal projection onto $0 \oplus \cdots \oplus H_k \oplus \cdots \oplus 0$.

($\Rightarrow$) If $\{x_j^1 \oplus \cdots \oplus x_j^N\}_{j\in\mathbb{J}}$ is a Parseval frame for $H_1 \oplus \cdots \oplus H_N$, then for each k, $\{x_j^k\}_{j\in\mathbb{J}}$ is a Parseval frame for H_k, since $\{x_j^k\}_{j\in\mathbb{J}}$ is the image of a Parseval frame under the projection P_k. Moreover, if $l \neq k$, then for every $v \in H_1 \oplus \cdots \oplus H_N$, we have

$$\begin{aligned} 0 &= P_l P_k v \\ &= P_l \left(\sum_{j\in\mathbb{Z}} \langle P_k v, (x_j^1 \oplus \cdots \oplus x_j^N) \rangle (x_j^1 \oplus \cdots \oplus x_j^N) \right) \\ &= \sum_{j\in\mathbb{Z}} \langle P_k v, (0 \oplus \cdots \oplus x_j^k \oplus \cdots \oplus 0) \rangle (0 \oplus \cdots \oplus x_j^l \oplus \cdots \oplus 0) \\ &= 0 \oplus \cdots \oplus \sum_{j\in\mathbb{Z}} \langle P_k v, x_j^k \rangle x_j^l \oplus \cdots \oplus 0. \end{aligned}$$

Whence, $\{x_j^k\}$ and $\{x_j^l\}$ are orthogonal frames.

($\Leftarrow$) Conversely, if the conditions of both 1. and 2. are satisfied, then for every $v \in H_1 \oplus \cdots \oplus H_N$,

$$\begin{aligned} \sum_{j\in\mathbb{Z}} \langle v, (x_j^1 \oplus \cdots \oplus x_j^N) \rangle (x_j^1 \oplus \cdots \oplus x_j^N) &= \left(\sum_{j\in\mathbb{Z}} \sum_{k=1}^{N} \langle P_k v, x_j^k \rangle x_j^1 \oplus \cdots \oplus \sum_{j\in\mathbb{Z}} \sum_{k=1}^{N} \langle P_k v, x_j^k \rangle x_j^N \right) \\ &= \left(\sum_{j\in\mathbb{J}} \langle P_1 v, x_j^1 \rangle x_j^1 \oplus \cdots \oplus \sum_{j\in\mathbb{J}} \langle P_N v, x_j^N \rangle x_j^N \right) \\ &= v. \end{aligned}$$

□

We have therefore that, under the orthogonality assumptions on $\psi_1, \dots, \psi_N$, the collection

$$\{\sqrt{2}^n (\psi_1(2^n \cdot -l), \dots, \psi_N(2^n \cdot -l)) : n, l \in \mathbb{Z}\}$$

forms a Parseval frame for $L^2(\mathbb{R}) \oplus \cdots \oplus L^2(\mathbb{R})$. Hence, we can naturally think of Ψ as being a vector-valued wavelet and Θ_Ψ as being a vector-valued wavelet transform, or more precisely, a wavelet transform for vector-valued data.

For an alternate construction of wavelets in $L^2(\mathbb{R}) \oplus \cdots \oplus L^2(\mathbb{R})$, see [3].

1.3. **Main Results.** The two main results of the paper are an algorithm for the construction of orthogonal wavelet frames and a construction of the VDWT. The construction of orthogonal wavelet frames in one spatial dimension is given in Theorems 2.1.1 and 2.1.2, where the basic ingredients consists of a fixed wavelet basis and a paraunitary matrix of an appropriate size. The number of orthogonal wavelet frames that can be constructed is arbitrary, and is determined by the size of the paraunitary matrix. Analogous constructions for two spatial dimensions are given in Theorems 3.1.1 and 3.1.2.

Theorems 2.2.4 and 2.2.6 give parallel constructions for orthogonality of filterbanks. There is a technical restriction, described in Theorem 2.2.3, about the low pass filter of a filterbank. Despite this restriction, we construct the VDWT in Definition 2.2.5 using filterbanks which satisfy the orthogonality condition of Definition 2.2.1. We note here that in Definition 2.2.5, several of the filter outputs are summed together, which is not done in normal DWT's. This summation corresponds to the direct sum nature of vector-valued data (Equation 2 Theorem 1.2.1).

1.4. **Background Results.** We present in this subsection some previously published results which we shall need on duality and orthogonality of wavelet frames.

Theorem 1.4.1. *Suppose $\{\psi_1, \ldots, \psi_r\}$ and $\{\eta_1, \ldots, \eta_r\}$ generate wavelet frames in $L^2(\mathbb{R}^d)$. The frames are dual if and only if*

1. $$\sum_{k=1}^{r} \sum_{j \in \mathbb{Z}} \hat{\psi}_k(2^j \xi) \overline{\hat{\eta}_k(2^j \xi)} = 1 \text{ a.e. } \xi;$$

2. *for every $q \in \mathbb{Z}^d \setminus 2\mathbb{Z}^d$,*
$$\sum_{k=1}^{r} \sum_{j=0}^{\infty} \hat{\psi}_k(2^j \xi) \overline{\hat{\eta}_k(2^j (\xi + q))} = 0 \text{ a.e. } \xi.$$

In particular, $\{\psi_1, \ldots, \psi_r\}$ generates a Parseval wavelet frame if the two equations hold for $\eta_k = \psi_k$.

Proof. See [15, 4]. □

Theorem 1.4.2. *Suppose $\{\psi_1, \ldots, \psi_r\}$ and $\{\eta_1, \ldots, \eta_r\}$ generate affine Bessel sequences in $L^2(\mathbb{R}^d)$; they are orthogonal if and only if*

1. $$\sum_{k=1}^{r} \sum_{j \in \mathbb{Z}} \hat{\psi}_k(2^j \xi) \overline{\hat{\eta}_k(2^j \xi)} = 0 \text{ a.e. } \xi;$$

2. *for every $q \in \mathbb{Z}^d \setminus 2\mathbb{Z}^d$,*
$$\sum_{k=1}^{r} \sum_{j=0}^{\infty} \hat{\psi}_k(2^j \xi) \overline{\hat{\eta}_k(2^j (\xi + q))} = 0 \text{ a.e. } \xi.$$

Proof. See [18]. □

The following theorem is stated for one spatial dimension only.

Theorem 1.4.3 (Unitary Extension Principle [9])**.** *Suppose $\phi \in L^2(\mathbb{R})$ is a refinable function, with low pass filter $m(\xi)$, which satisfies the following two conditions:*

1. $\lim_{\xi \to 0} \hat{\phi}(\xi) = 1$*;*
2. $\sum_{l \in \mathbb{Z}} |\hat{\phi}(\xi + l)|^2 \in L^\infty(\mathbb{R})$.

Let $m_1(\xi), \dots, m_r(\xi) \in L^\infty([0,1))$ such that the matrix

$$M(\xi) = \begin{pmatrix} m(\xi) & m(\xi + 1/2) \\ m_1(\xi) & m_1(\xi + 1/2) \\ \vdots & \vdots \\ m_r(\xi) & m_r(\xi + 1/2) \end{pmatrix}$$

satisfies the matrix equation

$$M^*(\xi) M(\xi) = I_2$$

for almost every ξ. Then, the affine system generated by $\{\psi_1, \dots, \psi_r\}$, where

$$\hat{\psi}_k(2\xi) = m_k(\xi)\hat{\phi}(\xi), \quad k = 1, \dots, r,$$

is a Parseval wavelet frame.

2. Orthogonal Wavelet Frames in One Spatial Dimension

In this section we shall construct a vector-valued wavelet transform in one spatial dimension. We present characterization and construction results for orthogonal wavelet frames in $L^2(\mathbb{R})$. We then discuss the analogous results for filter banks, and describe the VDWT.

We shall restrict our attention to dilation by 2. We assume that all refinable functions $\phi \in L^2(\mathbb{R})$ satisfy the conditions of the Unitary Extension Principle, and that all high pass filters m_k $(k \neq 0)$ are such that the affine system generated by ψ, defined by

$$\hat{\psi}_k(2\xi) = m_k(\xi)\hat{\phi}(\xi),$$

is a Bessel system. Given a collection of filters $\mathcal{M} = \{m_0, m_1, \dots, m_r\} \subset L^\infty([0,1))$ let $M(\xi)$ and $\widetilde{M}(\xi)$ be the matrices

$$M(\xi) = \begin{pmatrix} m_0(\xi) & m_0(\xi + 1/2) \\ m_1(\xi) & m_1(\xi + 1/2) \\ \vdots & \vdots \\ m_r(\xi) & m_r(\xi + 1/2) \end{pmatrix} \quad \text{and} \quad \widetilde{M}(\xi) = \begin{pmatrix} m_1(\xi) & m_1(\xi + 1/2) \\ m_2(\xi) & m_2(\xi + 1/2) \\ \vdots & \vdots \\ m_r(\xi) & m_r(\xi + 1/2) \end{pmatrix}. \tag{3}$$

In the remainder of the paper, the filter banks will be composed of a single low-pass filter (with index 0) and a number of high-pass filters.

2.1. **Construction of Orthogonal Wavelet Frames.** We present an algorithm for the construction of arbitrarily many orthogonal wavelet frames. The wavelet frames are MRA based (sometimes called framelets [9]), and the construction utilizes the Unitary Extension Principle.

Theorem 2.1.1. *Suppose $\phi \in L^2(\mathbb{R})$ be a refinable function which satisfies the conditions of the unitary extension principle, and let $m(\xi)$ be the associated low pass filter. Let $\mathcal{M} = \{m_0(\xi), m_1(\xi), \dots, m_r(\xi)\}$ and $\mathcal{N} = \{n_0(\xi), n_1(\xi), \dots, n_r(\xi)\}$ be filter banks with $m_0 = n_0 = m$. Suppose that the following matrix equations hold:*

1. *$M^*(\xi)M(\xi) = I_2$ for almost every ξ;*
2. *$N^*(\xi)N(\xi) = I_2$ for almost every ξ;*
3. *$\widetilde{M}^*(\xi)\widetilde{N}(\xi) = 0$ for almost every ξ.*

Let $\hat{\psi}_k(2\xi) = m_k(\xi)\hat{\phi}(\xi)$ and $\hat{\eta}_k(2\xi) = n_k(\xi)\hat{\phi}(\xi)$, $1 \le k \le r$. Then $\{\psi_1, \dots, \psi_r\}$ and $\{\eta_1, \dots, \eta_r\}$ generate orthogonal Parseval wavelet frames.

Proof. That $\{\psi_1, \dots, \psi_r\}$ and $\{\eta_1, \dots, \eta_r\}$ generate Parseval wavelet frames follows from the Unitary Extension Principle (Theorem 1.4.3). We use the characterization equations of Theorem 1.4.2 to prove orthogonality. Consider

$$\begin{aligned}\sum_{k=1}^{r}\sum_{j\in\mathbb{Z}} \hat{\psi}_k(2^j\xi)\overline{\hat{\eta}_k(2^j\xi)} &= \sum_{k=1}^{r}\sum_{j\in\mathbb{Z}} m_k(2^j\xi)\hat{\phi}(2^j\xi)\overline{n_k(2^j\xi)\hat{\phi}(2^j\xi)} \\ &= \sum_{j\in\mathbb{Z}} |\hat{\phi}(2^j\xi)|^2 \sum_{k=1}^{r} m_k(2^j\xi)\overline{n_k(2^j\xi)} \\ &= 0\end{aligned}$$

for almost every ξ by item 3 above. Note that the order of summation can be reversed since the sum is absolutely summable: for each k, by Hölder's inequality and by virtue of the fact that ψ_k and η_k generate Bessel sequences,

$$\sum_{j\in\mathbb{Z}} |\hat{\psi}_k(2^j\xi)\overline{\hat{\eta}_k(2^j\xi)}| \le \sum_{j\in\mathbb{Z}} |\hat{\psi}_k(2^j\xi)|^2 \sum_{j\in\mathbb{Z}} |\hat{\eta}_k(2^j\xi)|^2 < \infty.$$

See [14, Theorem 8.3.2].

Likewise, for q odd,

$$\begin{aligned}\sum_{k=1}^{r}\sum_{j=0}^{\infty} \hat{\psi}_k(2^j\xi)\overline{\hat{\eta}_k(2^j(\xi+q))} &= \sum_{k=1}^{r}\sum_{j=0}^{\infty} m_k(2^{j-1}\xi)\hat{\phi}(2^{j-1}\xi)\overline{n_k(2^{j-1}(\xi+q))\hat{\phi}(2^{j-1}(\xi+q))} \\ &= \sum_{j=0}^{\infty} \hat{\phi}(2^j\omega)\overline{\hat{\phi}(2^j(\omega+q/2))} \sum_{k=1}^{r} m_k(2^j\omega)\overline{n_k(2^j\omega+2^{j-1}q)} \\ &= 0\end{aligned}$$

again by item 3, where $\omega = \xi/2$. □

The above proof shows that each of the terms indexed over j in the sums of Theorem 2.1.1 is 0, i.e. for each j,

$$\sum_{k=1}^{r} \hat{\psi}_k(2^j\xi)\overline{\hat{\eta}_k(2^j\xi)} = 0 \text{ and } \sum_{k=1}^{r} \hat{\psi}_k(2^j\xi)\overline{\hat{\eta}_k(2^j(\omega+q))} = 0.$$

We call this a "local" orthogonality condition. In subsection 2.3 we shall discuss non-local orthogonality.

The following theorem describes a general construction algorithm for locally orthogonal wavelet frames.

Theorem 2.1.2. *Suppose $K(\xi)$ is a $r \times r$ paraunitary matrix with 1/2-periodic entries $a_{i,j}(\xi)$; let $K_j(\xi)$ denote the j-th column. Suppose m_0 and m_1 are low and high pass filters, respectively, for an orthonormal wavelet basis with scaling function ϕ. For $j = 1, \dots, r$, define new filters via*

$$\begin{pmatrix} n_1^j(\xi) \\ \vdots \\ n_r^j(\xi) \end{pmatrix} = K_j(\xi) m_1(\xi).$$

Then, for $j = 1, \dots, r$, the affine systems generated by $\{\psi_l^j : l = 1, \dots, r\}$ obtained via

$$\hat{\psi}_l^j(2\xi) = n_l^j(\xi)\hat{\phi}(\xi) \tag{4}$$

are Parseval frames and are pairwise orthogonal.

Proof. We verify that the construction satisfies the conditions of Theorem 2.1.1. We first verify the unitary extension principle. Letting $\mathcal{M}_j = \{m_0, n_1^j, \dots, n_r^j\}$, we must show that

$$M_j^*(\xi)M_j(\xi) = I_2, \quad 1 \le j \le r,$$

where M_j is defined according to (3). We examine the entries of $M_j^*(\xi)M_j(\xi)$ individually. Since the columns of $K(\xi)$ have length 1, it follows that

$$\begin{aligned}
\left[M_j^*(\xi)M_j(\xi)\right]_{1,1} &= |m_0(\xi)|^2 + \sum_{k=1}^{r} |a_{k,j}(\xi)m_1(\xi)|^2 \\
&= |m_0(\xi)|^2 + \sum_{k=1}^{r} |a_{k,j}(\xi)|^2 |m_1(\xi)|^2 \\
&= |m_0(\xi)|^2 + |m_1(\xi)|^2 \\
&= 1.
\end{aligned}$$

Likewise,

$$\left[M_j^*(\xi)M_j(\xi)\right]_{2,2} = |m_0(\xi + 1/2)|^2 + \sum_{k=1}^{r} |a_{k,j}(\xi + 1/2)m_1(\xi + 1/2)|^2 = 1.$$

Now, since the entries of $K(\xi)$ are 1/2-periodic,

$$\begin{aligned}
\left[M_j^*(\xi)M_j(\xi)\right]_{1,2} &= \overline{m_0(\xi)}m_0(\xi + 1/2) + \sum_{k=1}^{r} \overline{a_{k,j}(\xi)m_1(\xi)}a_{k,j}(\xi + 1/2)m_1(\xi + 1/2) \\
&= \overline{m_0(\xi)}m_0(\xi + 1/2) + \sum_{k=1}^{r} |a_{k,j}(\xi)|^2 \overline{m_1(\xi)}m_1(\xi + 1/2) \\
&= \overline{m_0(\xi)}m_0(\xi + 1/2) + \overline{m_1(\xi)}m_1(\xi + 1/2) \\
&= 0.
\end{aligned}$$

Finally, the $(2,1)$-entry must be zero by conjugate symmetry of $M_j^*(\xi)M_j(\xi)$.

For orthogonality, using the notation of Equation 3, we have for $j = 1, \dots, r$,

$$\widetilde{M}_j(\xi) = \begin{pmatrix} n_1^j(\xi) & n_1^j(\xi+1/2) \\ \vdots & \vdots \\ n_r^j(\xi) & n_r^j(\xi+1/2) \end{pmatrix} = K_j(\xi) \begin{pmatrix} m_1(\xi) & m_1(\xi+1/2) \end{pmatrix}.$$

Thus, we have for $j \neq j'$

$$\widetilde{M}_j^*(\xi)\widetilde{M}_{j'}(\xi) = \begin{pmatrix} \overline{m_1(\xi)} \\ \overline{m_1(\xi+1/2)} \end{pmatrix} K_j^*(\xi) K_{j'}(\xi) \begin{pmatrix} m_1(\xi) & m_1(\xi+1/2) \end{pmatrix} = 0,$$

since the product of the middle two matrices is 0 by the orthogonality of the columns of $K(\xi)$. □

The following proposition is directly related to the construction algorithm in Theorem 2.1.2. The multiplication of the high pass filter by the entries of the paraunitary matrix will increase the length of the filter and, consequently, the support of the wavelet frames, which is undesirable. Theorem 2.1.2 assumes the entries of the paraunitary matrix are 1/2 periodic, increasing the length of the filter by at least 2. Except for constant entries, this 1/2 periodicity is necessary.

Proposition 2.1.3. *If ϕ is compactly supported, the paraunitary matrix K in Theorem 2.1.2 must have entries which are 1/2-periodic.*

Proof. The proof will follow the notation of Theorem 2.1.2. Since we require that the matrix

$$M(\xi) = \begin{pmatrix} m_0(\xi) & m_0(\xi+1/2) \\ a_{1,1}(\xi)m_1(\xi) & a_{1,1}(\xi+1/2)m_1(\xi+1/2) \\ \vdots & \vdots \\ a_{1,N}(\xi)m_1(\xi) & a_{1,N}(\xi+1/2)m_1(\xi+1/2) \end{pmatrix}$$

satisfy the equation

$$M^*(\xi)M(\xi) = I_2 \text{ a.e. } \xi,$$

we must have that for almost every ξ,

$$0 = \overline{m_0(\xi)}m_0(\xi+1/2) + \overline{m_1(\xi)}m_1(\xi+1/2)\sum_{j=1}^{N}\overline{a_{1,j}(\xi)}a_{1,j}(\xi+1/2).$$

Since we also have that

$$0 = \overline{m_0(\xi)}m_0(\xi+1/2) + \overline{m_1(\xi)}m_1(\xi+1/2),$$

we must have that either $\overline{m_1(\xi)}m_1(\xi+1/2) = 0$ or $\sum_{j=1}^{N}\overline{a_{1,j}(\xi)}a_{1,j}(\xi+1/2) = 1$. If ϕ is compactly supported, then the first possibility is eliminated except possibly on a set of measure 0, whence the second must hold almost everywhere. Now, the sum is precisely the inner product of the two vectors $(a_{1,j}(\xi))$ and $(a_{1,j}(\xi+1/2))$, each of which has length 1. Applying Cauchy-Schwarz yields that the two vectors must be identical for almost every ξ. □

2.2. **Vector-valued Discrete Wavelet Transform.** This section is concerned with the discrete implementation of the orthogonal wavelet frames discussed above, which will ultimately lead to the definition of a vector-valued discrete wavelet transform for multichannel data. Naturally, filter banks will play an important role in the developments of this section, so we shall begin with essential notation before moving on to describe a notion of orthogonality for filter banks and, shortly thereafter, filter bank counterparts to the results of Theorems 2.1.1 and 2.1.2.

Figure 1 depicts the block diagram for a filtering scheme with analysis filterbank $\mathcal{M} = \{m_0, m_1, \dots, m_r\}$ and synthesis filterbank $\mathcal{N} = \{n_0, n_1, \dots, n_r\}$. The notation $\downarrow_2$ represents downsampling by 2 in $\ell^2(\mathbb{Z})$. Similarly, $\uparrow_2$ will represent upsampling by 2 in $l^2(\mathbb{Z})$. We will succumb to a slight abuse of notation in that no distinction shall be made between the sequence corresponding to a filter and its continuous domain counterpart.

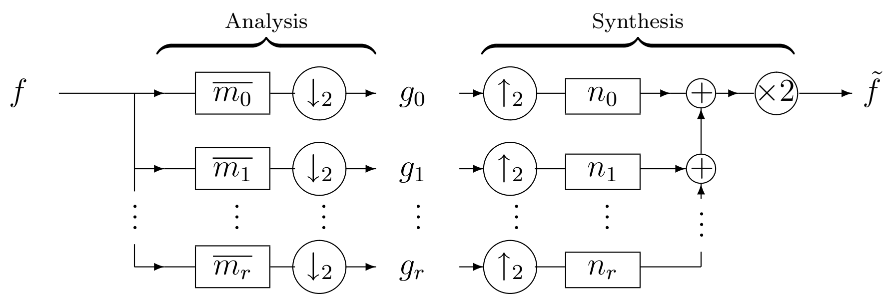

FIGURE 1. Block-Diagram of a filter bank.

For $f \in \ell^2(\mathbb{Z})$, it is well known that the outputs from the analysis stage may be expressed as

$$\hat{g}_\ell(\xi) = 1/2 \left[\hat{f}(\xi/2)\overline{m_\ell(\xi/2)} + \hat{f}(\xi/2 + 1/2)\overline{m_\ell(\xi/2 + 1/2)} \right], \quad 0 \leq \ell \leq r,$$

while the output of the filter bank after synthesis is given by

$$\hat{\tilde{f}}(\xi) = 2 \sum_{\ell=0}^{r} \hat{g}_\ell(2\xi) n_\ell(\xi).$$

Expanding this last equality in terms of $\hat{f}$ and the filter bank $\mathcal{M}$, we arrive at

$$\hat{\tilde{f}}(\xi) = \sum_{\ell=0}^{r} \left[\overline{m_\ell(\xi)} n_\ell(\xi) \hat{f}(\xi) + \overline{m_\ell(\xi + 1/2)} n_\ell(\xi) \hat{f}(\xi + 1/2) \right]. \tag{5}$$

Definition 2.2.1. We say that the filter banks $\mathcal{M}$ and $\mathcal{N}$ are *orthogonal* if, for any input vector, the composition of the analysis stage of $\mathcal{M}$ with the synthesis stage of $\mathcal{N}$ yields 0, i.e. for any input $f \in \ell^2(\mathbb{Z})$, $\tilde{f} = 0$.

Given a filter bank $\mathcal{M} = \{m_0, m_1, \dots, m_r\}$, the matrices $M(\xi)$ and $\widetilde{M}(\xi)$ will again be defined according to (3). The following characterization of orthogonality for the filterbanks $\mathcal{M}$ and $\mathcal{N}$ follows immediately from (5) and the 1-periodicity of the filters.

Theorem 2.2.2. *The filter banks $\mathcal{M} = \{m_0, \dots, m_r\}$ and $\mathcal{N} = \{n_0, \dots, n_r\}$ are orthogonal if and only if $M^*(\xi)N(\xi) = 0$ holds a.e. ξ.*

The following result shows that we cannot hope to achieve complete orthogonality for filter banks that include a single low pass filter, as is the case of our construction of orthogonal wavelet frames.

Theorem 2.2.3. *Suppose $\mathcal{M} = \{m_0, \ldots, m_r\}, \mathcal{N} = \{n_0, \ldots, n_r\}$ are filter banks in which m_0 and n_0 are low pass filters and m_ℓ and n_ℓ are high pass filters, $1 \leq \ell \leq r$. If each filter is continuous on a neighborhood of $\xi = 0$, then the filter banks $\mathcal{M}$ and $\mathcal{N}$ cannot be orthogonal.*

Proof. It follows from Theorem 2.2.2 that a necessary condition for orthogonality is

$$\sum_{\ell=0}^{r} \overline{m_\ell(\xi)} n_\ell(\xi) = 0 \ a.e.\ \xi.$$

The definitions of low and high pass filters imply that the above sum is 1 for $\xi = 0$. Moreover, the fact that each filter is continuous on a neighborhood of $\xi = 0$ guarantees that the sum is non-zero on some set of positive measure. ☐

It is natural to consider the counterparts to Theorems 2.2.2 and 2.2.3 for undecimated filter banks. In the undecimated case, the condition for orthogonality in Theorem 2.2.2 is replaced by

$$\sum_{\ell=0}^{r} \overline{m_\ell(\xi)} n_\ell(\xi) = 0 \ a.e.\ \xi,$$

from which a counterpart to Theorem 2.2.3 follows immediately. We leave the details to the reader.

In the absence of complete orthogonality we can still desire that the high-pass portions of two filter banks be orthogonal and, in fact, we can attempt to simultaneously achieve perfect reconstruction following the model provided by the orthogonal wavelet frames of Theorem 2.1.1. It is a familiar fact that a filter bank $\mathcal{M} = \{m_0, \ldots, m_r\}$ has the perfect reconstruction property (in the sense that $\tilde{f} = f$ in Figure 1 with $n_\ell = m_\ell,\ 0 \leq \ell \leq r$) provided that $M^*(\xi)M(\xi) = I_2$ for almost every ξ. Combining this fact with Theorem 2.2.2 we have the foundation for a discrete wavelet transform for multichannel data.

Theorem 2.2.4. *Let $\mathcal{M} = \{m_0, \ldots, m_r\}$ and $\mathcal{N} = \{n_0, \ldots, n_r\}$ be filterbanks in which m_0 and n_0 are low pass filters and m_ℓ and n_ℓ are high pass filters for $\ell = 1, \ldots, r$. Suppose that $\mathcal{M}$ and $\mathcal{N}$ satisfy the following matrix equalities:*

1. *$M^*(\xi)M(\xi) = I_2$ for almost every ξ;*
2. *$N^*(\xi)N(\xi) = I_2$ for almost every ξ;*
3. *$\widetilde{M}^*(\xi)\widetilde{N}(\xi) = 0$ for almost every ξ.*

Then the filter banks $\mathcal{M}$ and $\mathcal{N}$ each have the perfect reconstruction property and the high-pass filters $\{m_1, \ldots, m_r\}$ and $\{n_1, \ldots, n_r\}$ form orthogonal filter banks.

Comparison of the hypotheses of Theorem 2.2.4 with those of Theorem 2.1.1 reveals the fact that we are indeed describing the discrete counterpart to the orthogonal wavelet frames of Theorem 2.1.1. In fact, from Theorem 2.2.4 comes the notion of a Vector-valued Discrete Wavelet Transform (VDWT).

Definition 2.2.5. Let $\mathcal{M}_\ell = \{m_0^\ell, \ldots, m_s^\ell\}$ be filterbanks for $1 \leq \ell \leq r$. Let $f_0 = \oplus_{\ell=1}^{r} f_{0,\ell} \in \oplus_{\ell=1}^{r} \ell^2(\mathbb{Z})$. The *Vector-valued Discrete Wavelet Transform (VDWT)* of f_0 to scale $J > 0$

consists of:

$$f_J = \oplus_{\ell=1}^r f_{J,\ell} \quad \text{and} \quad g_{j,k}, \quad 1 \le j \le J,\ 1 \le k \le s,$$

where

$$f_{j+1,\ell} = \downarrow_2 (f_{j,\ell} * \widetilde{m}_0^\ell) \quad \text{and} \quad g_{j+1,k} = \sum_{\ell=1}^r \downarrow_2 (f_{j,\ell} * \widetilde{m}_k^\ell). \tag{6}$$

In Definition 2.2.5, $\widetilde{m}_k^\ell$ is the involution of the filter m_k^ℓ in the $\ell^2(\mathbb{Z})$ sense, which corresponds to conjugation under the Fourier transform. Note that the involution of a sequence $h = \{h_k\}_{k\in\mathbb{Z}}$ is $\tilde{h} = \{\overline{h_{-k}}\}_{k\in\mathbb{Z}}$.

Now suppose that each filterbank, $\mathcal{M}_\ell$, in Definition 2.2.5 has the perfect reconstruction property and, moreover, that the high-pass portions of the filter banks are pairwise orthogonal, i.e., $\widetilde{M_\ell^*}(\xi)\widetilde{M_{\ell'}}(\xi) = 0$ for almost every ξ whenever $\ell \ne \ell'$. In this case, the VDWT may be inverted using the identity:

$$f_{j-1,\ell} = 2\left[\uparrow_2 f_{j,\ell} * m_0^\ell + \sum_{k=1}^s \uparrow_2 g_{j,k} * m_k^\ell\right], \quad 1 \le j \le J,\ 1 \le \ell \le r. \tag{7}$$

Theorem 2.2.6. *Suppose $\mathcal{M}_\ell = \{m_0^\ell, \dots, m_s^\ell\}$, $1 \le \ell \le r$ are filterbanks which pairwise satisfy the hypotheses of Theorem 2.2.4. For any $f_0 = \oplus_{\ell=1}^r f_{0,\ell} \in \oplus_{\ell=1}^r \ell^2(\mathbb{Z})$,*

$$f_{0,\ell} = 2\left[\uparrow_2 f_{1,\ell} * m_0^\ell + \sum_{k=1}^s \uparrow_2 g_{1,k} * m_k^\ell\right], \quad 1 \le \ell \le r$$

where $f_{1,\ell}$ and $g_{1,k}$ are given in Equation 6.

Proof. Fix ℓ_0 such that $1 \le \ell_0 \le r$. We reconstruct f_{0,ℓ_0} via a filter scheme similar to that of Figure 1 in which $g_0 = f_{1,\ell_0}$ and $g_k = g_{1,k}$, $1 \le k \le s$, where the synthesis filters n_k are replaced by $m_k^{\ell_0}$ of the filter bank $\mathcal{M}_{\ell_0}$. The filter output in this case is then described analogously to that of (5),

$$\begin{aligned}
\hat{\tilde{f}}_{0,\ell_0}(\xi) &= 2\left[\hat{f}_{1,\ell_0}(\xi) m_0^{\ell_0}(\xi) + \sum_{k=1}^s \hat{g}_{1,k}(\xi) m_k^{\ell_0}(\xi)\right] \\
&= \overline{m_0^{\ell_0}(\xi)} m_0^{\ell_0}(\xi)\hat{f}_{0,\ell_0}(\xi) + \overline{m_0^{\ell_0}(\xi+1/2)} m_0^{\ell_0}(\xi)\hat{f}_{0,\ell_0}(\xi+1/2) \\
&\quad + \sum_{k=1}^s \sum_{\ell=1}^r \left[\overline{m_k^\ell(\xi)} m_k^{\ell_0}(\xi)\hat{f}_{0,\ell}(\xi) + \overline{m_k^\ell(\xi+1/2)} m_k^{\ell_0}(\xi)\hat{f}_{0,\ell}(\xi+1/2)\right] \\
&= \sum_{k=0}^s \left[\overline{m_k^{\ell_0}(\xi)} m_k^{\ell_0}(\xi)\hat{f}_{0,\ell_0}(\xi) + \overline{m_k^{\ell_0}(\xi+1/2)} m_k^{\ell_0}(\xi)\hat{f}_{0,\ell_0}(\xi+1/2)\right] \\
&= \hat{f}_{0,\ell_0}(\xi).
\end{aligned}$$

In this calculation the second to last equality uses the pairwise orthogonality of the high-pass portions of the filterbanks and the last equality uses the perfect reconstruction property of $\mathcal{M}_{\ell_0}$. □

In Definition 2.2.5 we have allowed for the possibility that the number of high-pass filters is not equal to the number of channels of the data. This allows room for redundant filterbank representations through the VDWT. (Note the following theorem, which establishes a lower bound on the number of filters based on the number of channels.) It is also perfectly reasonable to consider different sets of filters for analysis and synthesis; however, we have omitted any discussion of such dual filterbanks for clarity of presentation. We note that Theorem 2.1.2 provides a simple and flexible method for constructing filterbanks for use in the VDWT from any standard orthonormal wavelet filterbank. This approach will be discussed in greater detail below.

Theorem 2.2.7. *If the VDWT corresponding to the filterbanks $\mathcal{M}_\ell = \{m_0^\ell, \ldots, m_s^\ell\}$, $1 \le \ell \le r$ on $\oplus_{k=1}^r \ell^2(\mathbb{Z})$ has perfect reconstruction via Equation 7, then $s \ge r$.*

Proof. The fact that the filter banks provide perfect reconstruction via (7) implies that, pairwise, the filter banks satisfy the filter conditions of Theorem 2.2.4. Namely, for almost every ξ, $M_\ell^*(\xi) M_\ell(\xi) = I_2$, $1 \le \ell \le r$, and $\widetilde{M}_\ell^*(\xi)\widetilde{M}_{\ell'}(\xi) = 0$ when $1 \le \ell \ne \ell' \le r$. For $\xi \in [0,1]$ let

$$v_\ell(\xi) = \begin{pmatrix} m_0^\ell(\xi) \\ \vdots \\ m_s^\ell(\xi) \end{pmatrix} \in \mathbb{C}^{s+1} \quad \text{and} \quad \tilde{v}_\ell(\xi) = \begin{pmatrix} m_1^\ell(\xi) \\ \vdots \\ m_s^\ell(\xi) \end{pmatrix} \in \mathbb{C}^s, \quad 1 \le \ell \le r.$$

The perfect reconstruction property guarantees for almost every ξ that $v_\ell(\xi)$ and $v_\ell(\xi + 1/2)$ form an orthonormal pair of vectors in $\mathbb{C}^{s+1}$. Moreover, the pairwise orthogonality condition on the filter banks $\mathcal{M}_\ell$, $1 \le \ell \le r$ implies that the collection $\{\tilde{v}_\ell(\xi)\}_{\ell=1}^r \subseteq \mathbb{C}^s$ is orthogonal for almost every ξ. Provided that each of the vectors $\tilde{v}_\ell(\xi)$ is nonzero almost everywhere, it follows from dimensional considerations that $s \ge r$. Suppose by way of contradiction it happens that some $\tilde{v}_\ell(\xi)$ is zero for almost every ξ. Then $m_0^\ell(\xi) = 1$ on a set of positive measure for some fixed ℓ, contradicting the fact that $v_\ell(\xi)$ and $v_\ell(\xi + 1/2)$ are orthonormal almost everywhere. Thus, the collection $\{\tilde{v}_\ell(\xi)\}_{\ell=1}^r$ consists of r nonzero, orthogonal vectors for almost every ξ, implying $s \ge r$. □

Let us illustrate the differences for the VDWT versus the ordinary DWT for two-channel data on $\mathbb{Z}$, as one might find in stereo audio applications. Let $f_1 \oplus f_2 \in \ell^2(\mathbb{Z}) \oplus \ell^2(\mathbb{Z})$. To apply an ordinary Discrete Wavelet Transform (DWT), induced by filters $\mathcal{M} = \{m_0, m_1\}$, to $f_1 \oplus f_2$, one applies the analysis stage of the filtering scheme of Figure 1 to both f_1 and f_2. The result after the first stage of analysis consists of four $\ell^2(\mathbb{Z})$ sequences:

$$\downarrow_2 (f_1 * \widetilde{m}_0), \quad \downarrow_2 (f_2 * \widetilde{m}_0), \quad \downarrow_2 (f_1 * \widetilde{m}_1), \quad \downarrow_2 (f_2 * \widetilde{m}_1).$$

On the other hand, utilizing orthogonal filter banks, suppose $\{m_0, m_1, m_2\}$ and $\{n_0, n_1, n_2\}$ satisfy the conditions of Theorem 2.2.4. The output of the analysis part of the first stage filter bank here also consists of four $\ell^2(\mathbb{Z})$ sequences:

$$(8) \quad \downarrow_2 (f_1 * \widetilde{m}_0), \quad \downarrow_2 (f_2 * \widetilde{n}_0), \quad \downarrow_2 (f_1 * \widetilde{m}_1) + \downarrow_2 (f_2 * \widetilde{n}_1), \quad \downarrow_2 (f_1 * \widetilde{m}_2) + \downarrow_2 (f_2 * \widetilde{n}_2).$$

The block-diagram for the analysis and synthesis stages of this two-channel VDWT are presented in Figure 2. In order to relate (8) to the block-diagram, observe that $f_{j+1}^1 = \downarrow_2 (f_j^1 * \widetilde{m}_0)$, $f_{j+1}^2 = \downarrow_2 (f_j^2 * \widetilde{n}_0)$, $g_{j+1,1} = \downarrow_2 (f_j^1 * \widetilde{m}_1) + \downarrow_2 (f_j^2 * \widetilde{n}_1)$, and $g_{j+1,2} = \downarrow_2 (f_j^1 * \widetilde{m}_2) + \downarrow_2 (f_j^2 * \widetilde{n}_2)$.

Upon first glance it is a point of curiosity that each of these transforms uses the same amount of information, especially when we recall that we began with redundant filter banks for each channel. The key observation here comes from the fact that by summing the outputs of the appropriate pairs of high pass filters, the redundancy is eliminated. Whence, the output described in (8) corresponds to an orthonormal basis. Given sufficient correlation between

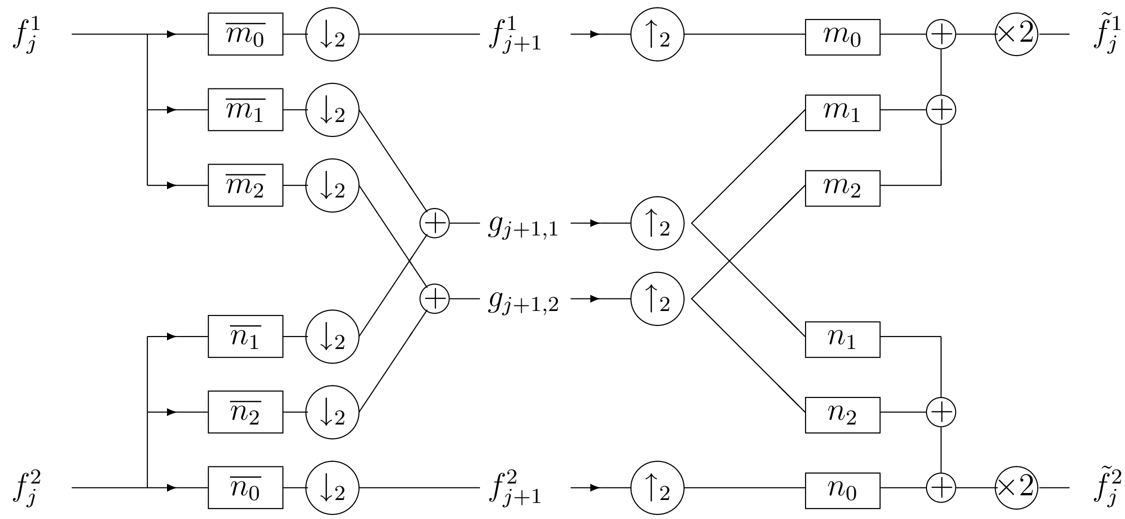

FIGURE 2. Interscale analysis and synthesis with the two-channel VDWT.

the two channels it is reasonable to expect that appropriately chosen filters may lead to a more efficient representation through the VDWT than is possible by analyzing each channel independently.

The elimination of redundancy possible with the VDWT goes beyond the above example. In particular, whenever the number of high-pass filters introduced is equal to the number of channels the resulting VDWT will not be redundant. The following theorem makes this claim rigorous by showing that the synthesis stage of the VDWT is injective provided that $s = r$ in the statement of Theorem 2.2.7.

Theorem 2.2.8. *Let $\mathcal{M}_k = \{m_0^k, \dots m_r^k\}$, $1 \le k \le r$ be filterbanks which satisfy the conditions of Theorem 2.2.4 pairwise. Let $\{w_0^k, w_k : 1 \le k \le r\} \subseteq \ell^2(\mathbb{Z})$ be inputs to the synthesis stage of the VDWT associated to the filterbanks $\mathcal{M}_k$, $1 \le k \le r$. If*

$$\left(\uparrow_2 w_0^k\right) * m_0^k + \sum_{j=1}^{r} \left(\uparrow_2 w_j\right) * m_j^k = 0, \quad 1 \le k \le r, \tag{9}$$

then $w_0^k = w_k = 0$ for $1 \le k \le r$.

Proof. Computing the Fourier transform of Equation (9) yields for almost every ξ:

$$\hat{w}_0^k(2\xi) m_0^k(\xi) + \sum_{j=1}^{r} \hat{w}_j(2\xi) m_j^k(\xi) = 0, \quad 1 \le k \le r.$$

By considering the above equation both as is and after substituting the variable ξ with $\xi + 1/2$, one obtains $2r$ independent equations. Writing these equations in matrix form and noting the 1/2-periodicity of the upsampled Fourier transforms results in

$$\begin{pmatrix} m_0^1(\xi) & & & m_1^1(\xi) & \cdots & m_r^1(\xi) \\ & \ddots & & \vdots & \ddots & \vdots \\ & & m_0^r(\xi) & m_1^r(\xi) & \cdots & m_r^r(\xi) \\ m_0^1(\xi+\frac{1}{2}) & & & m_1^1(\xi+\frac{1}{2}) & \cdots & m_r^1(\xi+\frac{1}{2}) \\ & \ddots & & \vdots & \ddots & \vdots \\ & & m_0^r(\xi+\frac{1}{2}) & m_1^r(\xi+\frac{1}{2}) & \cdots & m_r^r(\xi+\frac{1}{2}) \end{pmatrix} \begin{pmatrix} \hat{w}_0^1(2\xi) \\ \vdots \\ \hat{w}_0^r(2\xi) \\ \hat{w}_1(2\xi) \\ \vdots \\ \hat{w}_r(2\xi) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

where the omitted entries are zero. The matrix conditions of Theorem 2.2.4 imply that the rows of the above $2r \times 2r$ matrix are orthogonal, whence we must have $w_0^k = w_k = 0$, $1 \le k \le r$, as required. □

2.3. **Non-Local Orthogonality.** The local orthogonality of the wavelet frames constructed in Theorem 2.1.2 is a strong condition. Indeed, it says essentially (but not exactly) that the orthogonality of the wavelet frames is independent of the scale, that none of the cancellations occur across different scales. Strictly speaking, for orthogonal wavelet frames, this local orthogonality is not necessary, as Example 2.3.1 demonstrates. However, Theorem 2.3.3 shows for the orthogonality of the corresponding discrete wavelet transforms, or filter banks, the local orthogonality of the wavelet frames is necessary.

Example 2.3.1. Consider ϕ the scaling function for the Shannon wavelet: $\hat{\phi} = \chi_{(-1/2,1/2)}$. Let $m_1(\xi) = n_1(\xi) = \chi_{(-1/2,-1/4)\cup(1/4,1/2)}$, and let $m_2(\xi) = -n_2(\xi) = \chi_{(-1/4,-1/8)\cup(1/8,1/4)}$. Let $\hat{\psi}_k(2\xi) = m_k(\xi)\hat{\phi}(\xi)$, and $\eta_k(2\xi) = n_k(\xi)\hat{\phi}(\xi)$. One then verifies that

$$\sum_{j\in\mathbb{Z}} \sum_{k=1}^{2} \hat{\psi}_k(2^j\xi)\overline{\hat{\eta}_k(2^j\xi)} = 0$$

and

$$\sum_{j\in\mathbb{Z}} \sum_{k=1}^{2} \hat{\psi}_k(2^j\xi)\overline{\hat{\eta}_k(2^j(\xi+q))} = 0$$

for q odd, so that the wavelet frames are orthogonal. However, they are not strongly locally orthogonal since $m_1(\xi)\overline{n_1(\xi)} + m_2(\xi)\overline{n_2(\xi)}$ is not identically 0.

Definition 2.3.2. Suppose the affine systems generated by $\{\psi_1, \ldots, \psi_r\}$ and $\{\eta_1, \ldots, \eta_r\}$ are both frames for $L^2(\mathbb{R})$, which are MRA based, and suppose $\{m_0, m_1, \ldots, m_r\}$ and $\{n_0, n_1, \ldots, n_r\}$ are the corresponding filter banks. We say that the discrete wavelet transforms of $\{\psi_1, \ldots, \psi_r\}$ and $\{\eta_1, \ldots, \eta_r\}$ are orthogonal if the filter banks $\{m_1, \ldots, m_r\}$ and $\{n_1, \ldots, n_r\}$ are orthogonal as in Definition 2.2.1.

Our motivation is to apply the VDWT to discrete vector-valued data implemented by filter banks. Thus, for our purposes, we require that wavelet frames $\{\psi_1, \ldots, \psi_r\}$ and $\{\eta_1, \ldots, \eta_r\}$ not only satisfy the orthogonality condition as frames, but also possess orthogonal discrete wavelet transforms. Therefore, the corresponding high pass filters must also satisfy the orthogonality condition of Theorem 2.2.4, which in turn implies that the wavelet frames must in fact be locally orthogonal. Thus, for discrete wavelet transforms, non-local orthogonality is not possible, as we have just demonstrated and which proves the following statement.

Theorem 2.3.3. *Suppose the affine systems generated by $\{\psi_1, \ldots, \psi_r\}$ and $\{\eta_1, \ldots, \eta_r\}$ are orthogonal frames in $L^2(\mathbb{R})$ and both are MRA based. If the corresponding discrete wavelet transforms are orthogonal, then the wavelet frames are locally orthogonal.*

3. The VDWT in Multiple Spatial Dimensions

The main purpose of this section is to describe the construction of orthogonal wavelet frames and filterbanks in higher dimensions, analogous to those used in Theorems 2.1.2 and 2.2.4. There is no obstruction in generalizing the results to dimension 2 and higher, yet, for convenience, our discussion here will be limited to the bidimensional case. Color image data provides a natural testing ground for the VDWT in the context of two spatial dimensions and three independent channels and, for this reason, the last section will be used to present preliminary results with the VDWT motivated by the problem of color image compression.

Before delving into the details of the construction of orthogonal wavelet frames in higher dimensions we pause to consider two fundamentally different approaches to this problem. For the purposes of this discussion, let the term "orthogonalization" refer to the process of multiplying the high pass filters by the columns of a paraunitary matrix, as in Theorem 2.1.2. Ultimately, our constructions will make use of tensor products of one-dimensional filters and it is natural to consider the order of the orthogonalization and tensor product operations.

First, we could construct orthogonal wavelet frames in one-dimension and then form the tensor products in two-dimensions. If we begin with filters m_0 and m_1 and extend for r channels using m_0 and high-pass filters $n_1, n_2, \ldots, n_r$ then form tensor products, we will wind up with $(r+1)^2$ filters for each of the r channels of the VDWT. Second, we could begin with m_0 and m_1, form the four filters for two-dimensions via the tensor product and then orthogonalize, leading to a total of $3r+1$ filters in each of the r channels of the VDWT. It is rather curious that the first approach leads to a redundant representation. The explanation behind this curiosity is that one could actually construct orthogonal wavelet frames over r^2 channels using the approach of orthogonalization for r channels followed by the tensor product. Since we seek a non-redundant representation over three channels in two-dimensions the first approach is undesirable. Hence, we shall adopt the latter approach below, in which filters are first constructed in two dimensions via the tensor product and then orthogonalized over three channels.

If $\mathcal{M} = \{m_0, m_1, \ldots, m_r\} \subset L^\infty([0,1) \times [0,1))$, we construct the following matrices:

$$M(\xi,\omega) = \begin{pmatrix} m_0(\xi,\omega) & m_0(\xi+1/2,\omega) & m_0(\xi,\omega+1/2) & m_0(\xi+1/2,\omega+1/2) \\ \vdots & \vdots & \vdots & \vdots \\ m_r(\xi,\omega) & m_r(\xi+1/2,\omega) & m_r(\xi,\omega+1/2) & m_r(\xi+1/2,\omega+1/2) \end{pmatrix}$$

and

$$\widetilde{M}(\xi,\omega) = \begin{pmatrix} m_1(\xi,\omega) & m_1(\xi+1/2,\omega) & m_1(\xi,\omega+1/2) & m_1(\xi+1/2,\omega+1/2) \\ \vdots & \vdots & \vdots & \vdots \\ m_r(\xi,\omega) & m_r(\xi+1/2,\omega) & m_r(\xi,\omega+1/2) & m_r(\xi+1/2,\omega+1/2) \end{pmatrix}.$$

These filter matrices will play the same role as in one dimension.

3.1. **Construction.**

In two spatial dimensions, the orthogonality relations of Theorem 2.1.1 are unchanged, with the above matrices replacing those in the statement of Theorem 2.1.1.

We state the result explicitly for the direct sum of three copies of $L^2(\mathbb{R}^2)$, the (continuous) model for color image data.

Theorem 3.1.1. *Suppose* $\mathcal{M}_1 = \{m_0, m_1^1, \ldots, m_r^1\}$, $\mathcal{M}_2 = \{m_0, m_1^2, \ldots, m_r^2\}$, *and* $\mathcal{M}_3 = \{m_0, m_1^3, \ldots, m_r^3\}$ *are collections of filters that satisfy for almost every* (ξ, ω) *the following matrix equations:*

1. $M_j^*(\xi,\omega)M_j(\xi,\omega) = I_4,\ j = 1, 2, 3;$
2. $\widetilde{M}_j^*(\xi,\omega)\widetilde{M}_{j'}(\xi,\omega) = 0,\ 1 \le j, j' \le 3,\ j \ne j'.$

Suppose $\phi \in L^2(\mathbb{R}^2)$ *is a refinable function which satisfies the Unitary Extension Principle and has low pass filter* m_0. *Define for* $j = 1, 2, 3$ *and* $k = 1, \ldots, r$

$$\hat{\psi}_k^j(2\xi, 2\omega) = m_k^j(\xi,\omega)\hat{\phi}(\xi,\omega).$$

Then the affine system

$$\{2^n(\psi_k^1(2^n\cdot -l), \psi_k^2(2^n\cdot -l), \psi_k^3(2^n\cdot -l)) : n \in \mathbb{Z};\ l \in \mathbb{Z}^2;\ k = 1, \ldots, r\}$$

is a Parseval wavelet frame for $L^2(\mathbb{R}^2) \oplus L^2(\mathbb{R}^2) \oplus L^2(\mathbb{R}^2)$.

Theorem 3.1.2. *Let* $\phi \in L^2(\mathbb{R}^2)$ *be an orthonormal scaling function with low pass filter* m_0, *and suppose* $\{m_1, m_2, m_3\}$ *are corresponding high pass filters. Suppose* $P(\xi,\omega)$ *is a* 3×3 *paraunitary matrix, with entries* $a_{i,j}(\xi,\omega)$ *which are* $1/2$*-periodic in both variables. For* $i, j, k = 1, 2, 3$, *define the filters*

$$m_{i+3k-3}^j = a_{i,j}m_k \tag{10}$$

and for $l = 1, \ldots, 9$, *the wavelets*

$$\hat{\psi}_l^j(2\xi, 2\omega) = m_l^j(\xi,\omega)\hat{\phi}(\xi,\omega).$$

Then, for $j = 1, 2, 3$, *the affine systems generated by* $\{\psi_1^j, \ldots, \psi_9^j\}$ *are Parseval wavelet frames and are pairwise orthogonal. Therefore,*

$$\{2^n(\psi_k^1(2^n\cdot -l), \psi_k^2(2^n\cdot -l), \psi_k^3(2^n\cdot -l)) : n \in \mathbb{Z};\ l \in \mathbb{Z}^2;\ k = 1, \ldots, 9\}$$

is a Parseval wavelet frame for $L^2(\mathbb{R}^2) \oplus L^2(\mathbb{R}^2) \oplus L^2(\mathbb{R}^2)$.

Proof. We simply need to verify that the filters defined in Equation 10 satisfy the matrix equations of Theorem 3.1.1. For $\mathcal{M}_j = \{m_0^j, m_1^j, \ldots, m_9^j\}$, we have that the $1,1$ entry of the matrix $M_j^*(\xi,\omega)M_j(\xi,\omega)$ is:

$$\begin{aligned}
[M_j^*(\xi,\omega)M_j(\xi,\omega)]_{1,1} &= |m_0^j(\xi,\omega)|^2 + \sum_{l=1}^{9} |m_l^j(\xi,\omega)|^2 \\
&= |m_0(\xi,\omega)|^2 + \sum_{i=1}^{3}\sum_{k=1}^{3} |a_{i,j}(\xi,\omega)m_k(\xi,\omega)|^2 \\
&= |m_0(\xi,\omega)|^2 + \sum_{k=1}^{3} |m_k(\xi,\omega)|^2 \\
&= 1,
\end{aligned}$$

since the columns of $P(\xi,\omega)$ have length 1. Likewise, the $1,2$ entry of the matrix is:

$$\begin{aligned}
[M_j^*(\xi,\omega)M_j(\xi,\omega)]_{1,2} &= \overline{m_0^j(\xi,\omega)}m_0^j(\xi+1/2,\omega) + \sum_{l=1}^{9}\overline{m_l^j(\xi,\omega)}m_l^j(\xi+1/2,\omega) \\
&= \overline{m_0(\xi,\omega)}m_0(\xi+1/2,\omega) \\
&\quad + \sum_{i=1}^{3}\sum_{k=1}^{3}\overline{a_{i,j}(\xi,\omega)m_k(\xi,\omega)}a_{i,j}(\xi+1/2,\omega)m_k(\xi+1/2,\omega) \\
&= \overline{m_0(\xi,\omega)}m_0(\xi+1/2,\omega) + \sum_{i=1}^{3}|a_{i,j}(\xi,\omega)|^2\sum_{k=1}^{3}\overline{m_k(\xi,\omega)}m_k(\xi+1/2,\omega) \\
&= \overline{m_0(\xi,\omega)}m_0(\xi+1/2,\omega) + \sum_{k=1}^{3}\overline{m_k(\xi,\omega)}m_k(\xi+1/2,\omega) \\
&= 0,
\end{aligned}$$

since the entries of $P(\xi,\omega)$ are 1/2-periodic. Similar computations show that the off-diagonal entries of $M_j^*(\xi,\omega)M_j(\xi,\omega)$ are 0 and the diagonal entries are 1. Moreover, the orthogonality conditions follow analogously. □

We see from these two theorems that the building blocks for the VDWT for three channel data in two dimensions consist of filters corresponding to an orthonormal wavelet basis (which we will take to be a tensor product of one dimensional filters) and a 3×3 paraunitary matrix P. The VDWT in two dimensions is analogous to the VDWT in one dimension given in Definition 2.2.5. Using the columns of P to orthogonalize (Equation 10) the three high pass filters of the wavelet basis, we have three low pass filters and 27 high pass filters. However, using the orthogonality of P, we sum the outputs of the high pass filters corresponding to each column of P, thus reducing the actual high pass outputs to 9. Hence, we end up with 12 outputs after 1 stage of the VDWT, so just as in Theorem 2.2.8, the VDWT here corresponds to a basis, i.e. there is no redundancy.

3.2. **Construction of the Paraunitary Matrix.** We saw in the previous section that Theorem 3.1.2 uses columns from a paraunitary matrix in two variables for the construction of the orthogonal filterbanks in two dimensions. One approach to the construction of such a matrix is to take the independent product of two building-block paraunitary matrices:

$$P(\xi,\omega) = (I - vv^T + vv^Te^{2\pi i2\xi})(I - ww^T + ww^Te^{2\pi i2\omega}),$$

where $v, w \in \mathbb{C}^r$ are column vectors of unit length [17]. The resulting $r\times r$ matrix allows orthogonalization of a two-dimensional filterbank over r channels.

Alternatively, the following proposition describes a direct approach to the construction of paraunitaries in two variables. Since we want to minimize the length of the filters after orthogonalization, we consider only trigonometric polynomials of degree 2.

Proposition 3.2.1. *Let A, B, C, D be matrices of size $N\times N$. The matrix polynomial*

$$P(\xi,\omega) = A + Be^{2\pi i2\xi} + Ce^{2\pi i2\omega} + De^{2\pi i2(\xi+\omega)}$$

is paraunitary if

$$A^*A + B^*B + C^*C + D^*D = I_N,$$

and for $a, b \in \{A, B, C, D\}$ with $a \neq b$,

$$a^* b = 0.$$

If $N \leq 4$, then this can be accomplished by choosing N elements of $\{A, B, C, D\}$ to be one dimensional projections onto an orthonormal basis of $\mathbb{C}^N$, and any remaining elements to be 0.

Proof. Consider the computation:

$$\begin{aligned} P(\xi,\omega)^* P(\xi, \omega) \\ &= (A^* + B^* e^{-2\pi i 2\xi} + C^* e^{-2\pi i 2\omega} + D^* e^{-2\pi i 2(\xi+\omega)})(A + Be^{2\pi i 2\xi} + Ce^{2\pi i 2\omega} + De^{2\pi i 2(\xi+\omega)}) \\ &= A^*A + B^*B + C^*C + D^*D + \Lambda, \end{aligned}$$

where Λ consists of non-constant terms whose coefficients are sums of cross products. Therefore, if all of the cross products are 0, then $P(\xi, \omega)^* P(\xi, \omega) = I_N$. □

We begin with any standard one-dimensional orthonormal wavelet filters, m_0 and m_1, and form the tensor product before orthogonalization. We seek to orthogonalize by means of Proposition 3.2.1 and, hence, we need a 3×3 paraunitary matrix in two variables $P(\xi, \omega)$ for use with Theorem 2.1.2. As in the proposition, we let u, v, w be any orthonormal basis of $\mathbb{R}^3$. The matrices obtained by uu^T, vv^T, and ww^T are all projections onto orthogonal one-dimensional subspaces, so the product of any two must be 0. Moreover, the sum of the three projections is the identity. By choosing any one of A, B, C, D to be zero, say B, we can then let $A = uu^T$, $C = vv^T$, and $D = ww^T$.

3.3. **Application of the VDWT to Color Image Data.** The goal of this section is to compare the representation of color image data (using the standard Red-Green-Blue colorspace) of the VDWT to the that of a standard DWT applied to each color channel. We pause to recall the basic structure of a compression scheme, which consists of three essential steps:

1. Transformation into basis/frame coefficients;
2. Quantization and/or threshholding of coefficients;
3. Encoding of coefficients.

In practice, each component is tailored to the specific application in order to achieve the most efficient encoding possible. Here, our goal is to specialize the transformation of Step 1, above, through the use of the VDWT in order to take advantage of correlation among the components of the data. Our comparison will be limited to the transformation into basis coefficients followed by the implementation of a threshhold and, therefore, will not include an examination of quantization or encoding issues.

Because of the difficulty in displaying color images in print, we will also examine a one-dimensional example in which the data comes from a single row of pixels in a color image. In each experiment we will briefly describe the methodology of the comparison and then present compression and signal-to-noise ratios to quantify the sparseness of each transform. Note that the VDWT used in these comparisons will not be redundant and, therefore, the space of coefficients for the VDWT will have the same cardinality as the total coefficient space of the DWT over three channels.

Throughout our examination the Daubechies D4 filters will be used as the base filters in the case of the VDWT or in each channel with the usual DWT. We will define the *compression ratio* by

$$\text{CR} \;= \frac{\#\text{ samples} \times 3}{\#\text{ coefficients kept}}.$$

where the number of samples refers to the number of data points in each of the three channels in the original data and the number of coefficients kept is summed over all three channels. The *signal-to-noise ratio* of the reconstruction will be computed as

$$\text{SNR} \;= 20\log_{10}\left(\frac{\|\text{Original}\|_2}{\|\text{Original} - \text{Reconstruction}\|_2}\right),$$

where it should be clear that a higher SNR corresponds to better approximation in the $\|\cdot\|_2$ norm.

We begin with the one-dimensional data which originates from row 270 of the standard color 512×512 Lena image. This row was chosen for its relative non-smoothness in comparison with other portions of the image. A hard threshold was used at the finest scale and reduced at each coarser scale by a factor of $\sqrt{2}$. Both the DWT and VDWT were computed at all 9 scales of the data. The results of the compression experiment are presented in Table 1 and depicted in Figure 3. In Figure 3 (e), the labels "one", "two", and "three" refer to the summed outputs of the three orthogonal high-pass filters over the three color channels. After this summation, the wavelet coefficients no longer correspond directly to the respective color channels, but rather an amalgam of all three. In this experiment, the VDWT provides both a higher compression ratio and a better SNR. We should note that the orthogonalization in this case was achieved as described in Theorem 2.1.2 using a scalar unitary matrix,

$$K(\xi) = \begin{pmatrix} 0.407996 & -0.671184 & 0.618911 \\ -0.162407 & 0.613723 & 0.772630 \\ -0.898423 & -0.415745 & 0.141395 \end{pmatrix}.$$

Hereafter we shall refer to orthogonalization with a scalar unitary matrix as *scalar orthogonalization*. The term *polynomial orthogonalization* will be used when a unitary matrix with polynomial entries is used for orthogonalization.

Remark 3.3.1. A brief remark about the display of the wavelet coefficients in Figure 3 is in order. First, the wavelet coefficients at the lowest scale (most coarse) are depicted left-most in the graph and the wavelet coefficients at the highest scale are depicted at the right. For example, the wavelet coefficients after the first filtering stage occupy the range 257 to 512 in the figure. Second, the coefficients are normalized at each scale so that all scales may be clearly depicted on one axis. The same normalization was used for the DWT and the VDWT.

Source	Method	Threshhold	Comp. Ratio	SNR
Lena: row 270	D4, none	30	4.70	25.76
Lena: row 270	D4, scalar	30	5.54	26.55

Table 1. One-dimensional compression results.

We now turn our attention to compression results with two-dimensional color image data. As in the one-dimensional case, a chosen hard threshold was implemented at the finest scale

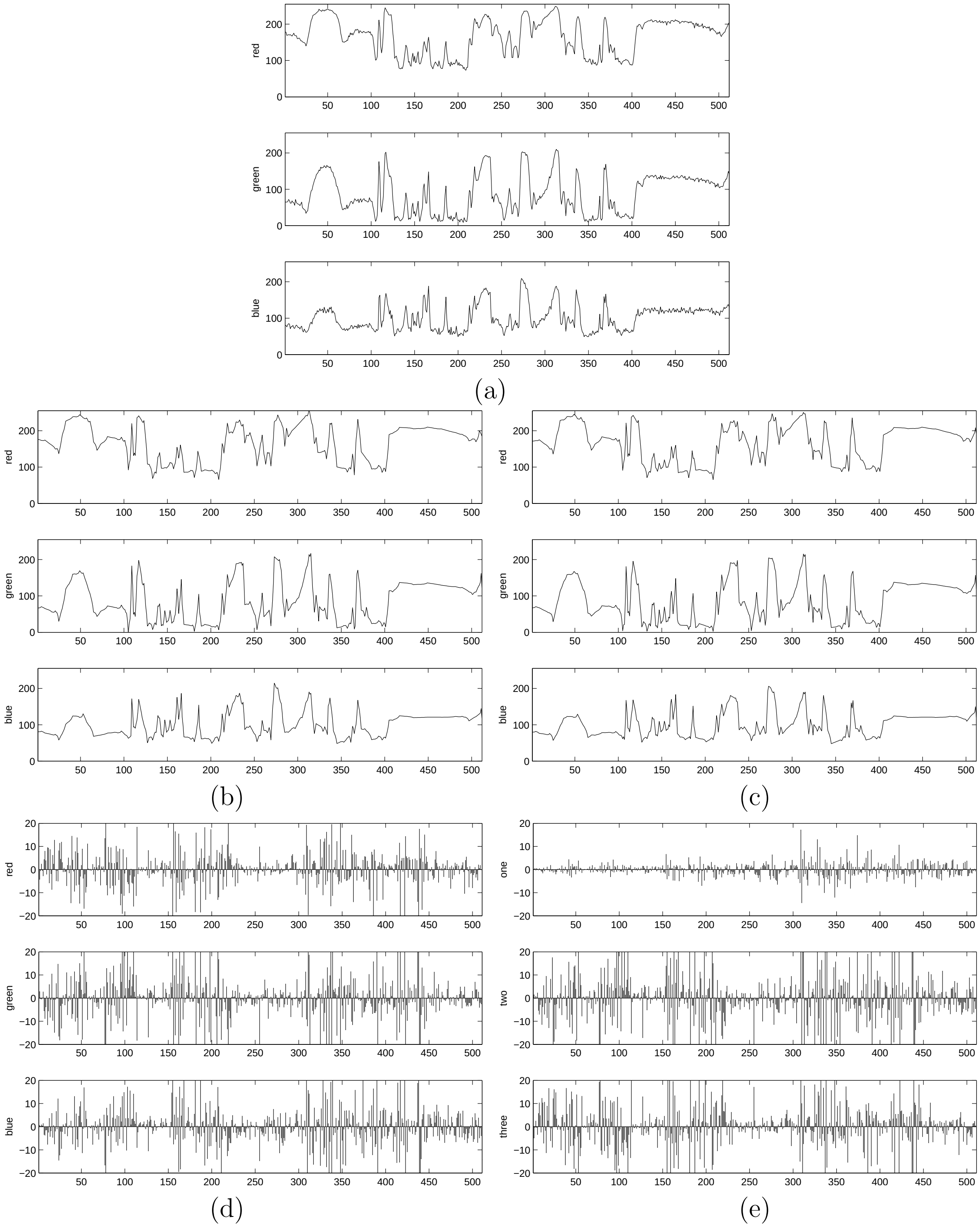

FIGURE 3. Comparison of usual DWT and scalar VDWT on 1-D data extracted from the color Lena image: (a) Original Data; (b) Reconstruction via usual DWT; (c) Reconstruction via scalar VDWT; (d) Wavelet coefficients using usual DWT; (e) Wavelet coefficients using scalar VDWT.

and reduced at each coarser scale by a factor of $\sqrt{2}$. The DWT and VDWT were limited to four scales and no thresholding was performed on the low-pass coefficients. In the two-dimensional experiments the usual DWT in each channel is compared to two different VDWT schemes, one a scalar orthogonalization and the other a polynomial orthogonalization. The filters for the scalar VDWT were generated as in Theorem 3.1.2 using

$$P(\xi,\omega) = \begin{pmatrix} -0.679565 & -0.324521 & -0.657934 \\ 0.517183 & 0.424139 & -0.743390 \\ 0.520301 & -0.845454 & -0.120393 \end{pmatrix},$$

while the polynomial orthogonalization implements $P(\xi,\omega)$ of the form used in Theorem 3.2.1,

$$P(\xi,\omega) = A + Be^{2\pi i2\xi} + Ce^{2\pi i2\omega} + De^{2\pi i2(\xi+\omega)},$$

where D is the zero matrix,

$$A = \begin{pmatrix} 0.058960 & -0.139956 & 0.046731 \\ -0.342804 & 0.813725 & -0.271699 \\ -0.129218 & 0.306729 & -0.102416 \end{pmatrix}, \quad B = \begin{pmatrix} 0.763185 & 0.125941 & -0.585726 \\ 0.063603 & 0.010496 & -0.048814 \\ 0.179498 & 0.029621 & -0.137760 \end{pmatrix},$$

and

$$C = \begin{pmatrix} -0.090153 & -0.083173 & -0.135351 \\ -0.184657 & -0.170359 & -0.277233 \\ 0.448743 & 0.413996 & 0.673716 \end{pmatrix}.$$

The results of the compression experiments are presented in Table 2. The famous Lena image (512×512 color version) was used for two of the experiments, one experiment comparing the usual DWT, the scalar VDWT, and the polynomial VDWT for a smaller threshold and the second adopting a larger threshold. In each of these experiments the scalar VDWT was best in terms of both compression ratio and SNR, while the polynomial VDWT yielded a performance between the usual DWT and scalar VDWT. Finally, the same scalar and polynomial VDWTs were applied to the 512×512 color Peppers image. The scalar VDWT still provided the best performance, followed by the polynomial VDWT, and the usual DWT.

Picture	Method	Threshhold	Comp. Ratio	SNR
Lena	D4, none	15	9.36	30.64
Lena	D4, scalar	15	10.96	30.93
Lena	D4, poly.	15	9.88	30.64
Lena	D4, none	50	28.77	26.14
Lena	D4, scalar	50	34.62	26.75
Lena	D4, poly.	50	30.58	26.37
Pepper	D4, none	15	10.71	31.41
Pepper	D4, scalar	15	12.14	32.06
Pepper	D4, poly.	15	10.83	31.66

TABLE 2. Performance of thresholding using orthogonal wavelet frames.

The results of these simple experiments suggest that orthogonalization can lead to a benefit in the representation of multi-channel data. It is a little surprising that the greater flexibility present with polynomial orthogonalization did not yield superior results to the scalar case.

One possible explanation for this fact is simply that the choice of orthogonalization was not optimized for the image, which is a natural area for future work with orthogonal wavelet frames. The fact that our fixed choices for the scalar and polynomial orthogonalizations led to improved compression for both the Lena and Peppers images supports the idea that optimization may lead to even better results. Another important aspect of future work would be the inclusion of the quantization and encoding components of the compression scheme.

Acknowledgements

The first author was supported in part by a SPRIG grant from the College of Liberal Arts and Sciences, Iowa State University. The third author was supported in part by NSF grant DMS-0355573. The authors are indebted to the editor and referees for greatly improving the presentation of the paper.

References

1. R. Balan, *Weyl-Heisenberg super frames*, preprint, 1999.
2. R. Balan and Z. Landau, *Topologies of Weyl-Heisenberg sets*, preprint, 2002.
3. S. Bildea, D. Dutkay, and G. Picioroaga, *MRA super-wavelets*, New York J. Math. **11** (2005), 1–19 (electronic).
4. M. Bownik, *A characterization of affine dual frames in $L^2(\mathbb{R}^n)$*, Appl. Comp. Harmonic Anal. **8** (2000), no. 2, 203–221.
5. C. Chui and Q. Jiang, *Balanced multi-wavelets in $\mathbb{R}^s$*, Math. Comp. **74** (2005), no. 251, 1323–1344 (electronic).
6. C. Chui and J. Lian, *Construction of compactly supported symmetric and antisymmetric orthonormal wavelets with scale = 3*, Appl. Comput. Harmon. Anal. **2** (1995), no. 1, 21–51.
7. ———, *A study of orthonormal multi-wavelets*, Appl. Numer. Math. **20** (1996), no. 3, 273–298, Selected keynote papers presented at 14th IMACS World Congress (Atlanta, GA, 1994).
8. I. Daubechies, A. Grossmann, and Y. Meyer, *Painless nonorthogonal expansions*, J. Math. Phys. **27** (1986), no. 5, 1271–1283.
9. I. Daubechies, B. Han, A. Ron, and Z. Shen, *Framelets: MRA-based constructions of wavelet frames*, Appl. Comput. Harmonic Anal. **14** (2003), no. 1, 1–46.
10. R. Duffin and A. Schaeffer, *A class of nonharmonic Fourier series*, Trans. Amer. Math. Soc. **72** (1952), 341–366.
11. J. Fowler and L. Hua, *Wavelet transforms for vector fields using omnidirectionally balanced multiwavelets*, IEEE Trans. Signal Process. **50** (2002), no. 12, 3018–3027.
12. J. Geronimo, D. Hardin, and P. Massopust, *Fractal functions and wavelet expansions based on several scaling functions*, J. Approx. Theory **78** (1994), no. 3, 373–401.
13. D. Han and D. Larson, *Frames, bases and group representations*, vol. 147, Mem. Amer. Math. Soc., no. 697, AMS, Providence, RI, September 2000.
14. E. Hernandez and G. Weiss, *An introduction to wavelets*, CRC Press, Boca Raton, 1998.
15. A. Ron and Z. Shen, *Affine systems in $L^2(\mathbb{R}^d)$: the analysis of the analysis operator*, J. Funct. Anal. **148** (1997), no. 2, 408–447.
16. Z. Shen, *Refinable function vectors*, SIAM J. Math. Anal. **29** (1998), no. 1, 235–250 (electronic).
17. P. P. Vaidyanathan, *Multirate systems and filter banks*, Prentice Hall, 1993.
18. E. Weber, *Orthogonal frames of translates*, Appl. Comput. Harmonic Anal. **17** (2004), no. 1, 69–90.
19. X. Xia, J. Geronimo, D. Hardin, and B. Suter, *Design of prefilters for discrete multiwavelet transforms*, IEEE Trans. Signal Process. **44** (1994), no. 1, 25–35.
20. X. Xia and B.W. Suter, *Vector-valued wavelets and vector filter banks*, IEEE Trans. Signal Process. **44** (1994), no. 3, 508–518.

Department of Mathematics, Iowa State University, 396 Carver Hall, Ames, IA 50011
Current address: Department of Mathematics, Rose-Hulman Institute of Technology, Terre Haute, IN 47803
E-mail address: bhatt@rose-hulman.edu

Department of Mathematics and Computer Science, Saint Louis University, St. Louis, MO 63103
E-mail address: brody@slu.edu

Department of Mathematics, Iowa State University, 396 Carver Hall, Ames, IA 50011
E-mail address: esweber@iastate.edu