Directional Haar Wavelet Frames on Triangles

Jens Krommweh[a], Gerlind Plonka[*,a]

[a]*Department of Mathematics, University of Duisburg-Essen, Campus Duisburg, 47048 Duisburg, Germany*

Abstract

Traditional wavelets are not very effective in dealing with images that contain orientated discontinuities (edges). To achieve a more efficient representation one has to use basis elements with much higher directional sensitivity. In recent years several approaches like curvelets and shearlets have been studied providing essentially optimal approximation properties for images that are piecewise smooth and have discontinuities along C^2-curves. While curvelets and shearlets have compact support in frequency domain, we construct directional wavelet frames generated by functions with compact support in time domain. Our Haar wavelet constructions can be seen as special composite dilation wavelets, being based on a generalized multiresolution analysis (MRA) associated with a dilation matrix and a finite collection of 'shear' matrices. The complete system of constructed wavelet functions forms a Parseval frame. Based on this MRA structure we provide an efficient filter bank algorithm. The freedom obtained by the redundancy of the applied Haar functions will be used for an efficient sparse representation of piecewise constant images as well as for image denoising.

Key words: Haar wavelet frames, non-separable wavelets, composite dilation wavelets, dual frames, sparse representation, image denoising
2000 MSC: 42C40, 41A30, 65T60, 68U10, 94A08

1. Introduction

Over the past few years, there has been a great interest in improved methods for sparse representations of higher dimensional data sets. Multiscale methods based on wavelets have been shown to provide successful schemes for data compression and denoising.

Indeed, wavelets are optimally efficient in representing functions with point singularities [23]. In addition, the multiresolution analysis (MRA) associated with wavelets results in fast algorithms for computing the wavelet coefficients [9, 23]. However, due to the missing rotation invariance of tensor product wavelets, the wavelet representation of $2D$-functions is not longer optimal.

Therefore, in recent years several attempts for improvement of wavelet systems in higher dimensions have been made, including complex wavelets [17], contourlets [11, 24], brushlets [8], curvelets [4, 5], bandelets [21], shearlets [14, 15, 22], and directionlets [27].

---

[*]Corresponding author
Email addresses: `jens.krommweh@uni-due.de` (Jens Krommweh), `gerlind.plonka@uni-due.de` (Gerlind Plonka)

Curvelets [4, 5] and shearlets [14, 15, 22] are examples of non-adaptive highly redundant function frames with strong anisotropic directional selectivity. For piecewise Hölder continuous functions of order 2 with discontinuities along C^2-curves, Candès and Donoho [5] proved that a best approximation f_M of a given function f with M curvelets satisfies

$$\|f - f_M\|^2 \leq C\, M^{-2}\, (\log_2 M)^3,$$

while a (tensor product) wavelet expansion only leads to an approximation error $\mathcal{O}(M^{-1})$ [23]. Up to the $(\log_2 M)^3$ factor, this curvelet approximation result is asymptotically optimal. A similar estimation has been achieved by Guo and Labate [14] for shearlet frames.

Instead of choosing a priori a basis or a frame to approximate f, one can rather adapt the approximation scheme to the image geometry. For example, one can construct an approximation f_M which is piecewise linear over an optimized triangulation including M triangles and satisfies $\|f - f_M\|^2 \leq C\, M^{-2}$. This requires adapting the triangulation to the edge geometry (see e.g. [10]). In [21], bandelet orthogonal bases and frames are introduced that adapt the geometric regularity of the image. Further, we want to mention the nonlinear edge adapted multiscale decompositions based on ENO schemes in [2, 7, 20] and the multidirectional edge adapted compression algorithm [1] based on an edge detection procedure.

In this paper we are especially interested in non-adaptive directional wavelet frames being compactly supported in time domain. The curvelet and shearlet systems constructed so far are tight frames of well-localized functions at different scales, positions and directions. The corresponding generating functions have compact support on triangles, 'parabolic' wedges or sheared wedges in frequency domain. In particular, for curvelet frames, there is no underlying multiresolution analysis supporting the efficient computation of curvelet representations. These circumstances can be seen as a certain drawback for the application of efficient wavelet filter banks based on these frames.

Therefore, we are strongly interested in directional wavelet frames generated by functions with compact support in time domain and providing an efficient algorithm based on the MRA structure.

The construction of directional Haar wavelet frames on triangles introduced in this paper is a first attempt in this direction. A different approach can also be found in [18]. Further constructions of piecewise constant wavelets in $L^2(\mathbb{R}^n)$ with $n \geq 2$ for other purposes are due to [13, 16, 25].

As in [18], our Haar wavelet constructions can be seen as special composite dilation wavelets (see [15]), being based on a generalized MRA associated with a dilation matrix and a collection of 'shear' matrices. But in contrast to [18] we shall use the dilation matrix $A = 2I$, leading to simple decomposition and reconstruction formulas. Since we use a finite collection of shear matrices, the considered MRA can also be understood as generated by a refinable function vector and the corresponding wavelet functions form a multiwavelet vector. Our results, presented in this paper, go far beyond the ideas given in [18], where Haar wavelet frames, based on the quincunx matrix as dilation matrix, are constructed, but without any consideration of frame properties, redundancy, and applications. Furthermore, we achieve a higher directional sensitivity considering eight directions instead of four, which produces an essential improvement in applications.

Due to their support, our scaling functions and wavelets are able to detect different directions. The complete system of constructed wavelet functions forms a Parseval frame. The freedom obtained by the redundancy of the applied Haar functions will be used for an efficient sparse representation of (piecewise constant) images as well as for image denoising.

In order to find a sparse representation of the image, we will apply a minimization of the vector of wavelet coefficients in the l^0-seminorm in each decomposition step. The considered l^0-minimization problem is related to the construction of M-term approximations of functions in a

redundant dictionary by greedy algorithms (see e.g. [6, 26]). We shall present a new and simple but efficient algorithm for finding frame representations with small l^0-seminorm that uses the known dependence relations in the frame and provides an optimal solution for piecewise constant images.

The piecewise constant directional Haar wavelet frame presented in this paper possesses some limitations yet. On the one hand, a construction of directional wavelet systems with small support in time domain and with higher smoothness is desirable. On the other hand, for image analysis one wishes to have as many different directions as possible. Unfortunately, this desire conflicts with a small redundancy of the wavelet system. We will discuss this issues and possible extensions of our approach in the Conclusions.

The paper is organized as follows. In Section 2, we introduce the space of scaling functions with compact support on triangles. Further, we present the canonical dual frames for the scaling spaces V_j. Section 3 is devoted to the construction of the directional Haar wavelet frame and corresponding decomposition and reconstruction formulas. In Section 4, we present the directional Haar wavelet filter bank based on the new wavelet frames on triangles. Further, we study the question of how to find a suitable orthogonal projection of a given digital image into the scaling space V_0 as well as the back projection after application of the filter bank algorithm. In Section 5, a new algorithm for sparse representation of images is presented, where we apply an l^0-minimization to the wavelet coefficients of the constructed frames. Finally, Section 6 is devoted to the application of the redundant directional Haar wavelet filter bank to image denoising and sparse image approximation. In particular, we shall compare its performance with curvelets and contourlets.

2. The space of scaling functions

We consider the domain $\Omega := [-1, 1]^2$ and divide it into 16 triangles with the same area, see Figure 1. We want to introduce a vector of characteristic functions on these 16 triangles. Let the first scaling function ϕ_0 be a characteristic function on the triangle

$$U_0 = \text{conv}\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\} := \{x \in \mathbb{R}^2 : 0 \leqslant x_2 \leqslant 1, \ 0 \leqslant x_1 \leqslant \tfrac{x_2}{2}\},$$

i.e.,

$$\phi_0(x) = \phi_0(x_1, x_2) = \chi_{U_0}(x_1, x_2) = \chi_{[0,1]}(\tfrac{2x_1}{x_2}) \cdot \chi_{[0,1]}(x_2).$$

The second scaling function ϕ_1 is given by

$$\phi_1(x) = \phi_1(x_1, x_2) = \chi_{U_1}(x_1, x_2) = \chi_{[1,2]}(\tfrac{2x_1}{x_2}) \cdot \chi_{[0,1]}(x_2),$$

where $U_1 = \text{conv}\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1 \end{pmatrix}\}$. Observe that, introducing the shear matrix $S := \begin{pmatrix} 1 & 1/2 \\ 0 & 1 \end{pmatrix}$, we have $\phi_1(x) = \phi_0(S^{-1}x)$. Let us apply the group $\mathcal{B} := \{B_i : i = 0, \ldots, 7\}$ of isometries of the square $[-1, 1]^2$ with

$$B_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ B_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ B_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \ B_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$B_4 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \ B_5 = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \ B_6 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ B_7 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

3

Then, for $i = 0, \ldots, 7$ we have

$$U_{2i} = \{B_i^{-1} x : x \in U_0\} = B_i^{-1} U_0, \qquad U_{2i+1} = \{B_i^{-1} x : x \in U_1\} = B_i^{-1} U_1,$$

and we define the further scaling functions ϕ_i by

$$
\begin{aligned}
\phi_{2i}(x) &:= \phi_0(B_i x) = \chi_{U_0}(B_i x) = \chi_{B_i^{-1} U_0}(x) = \chi_{U_{2i}}(x), \\
\phi_{2i+1}(x) &:= \phi_1(B_i x) = \chi_{U_1}(B_i x) = \chi_{B_i^{-1} U_1}(x) = \chi_{U_{2i+1}}(x), \qquad i = 0, \ldots, 7.
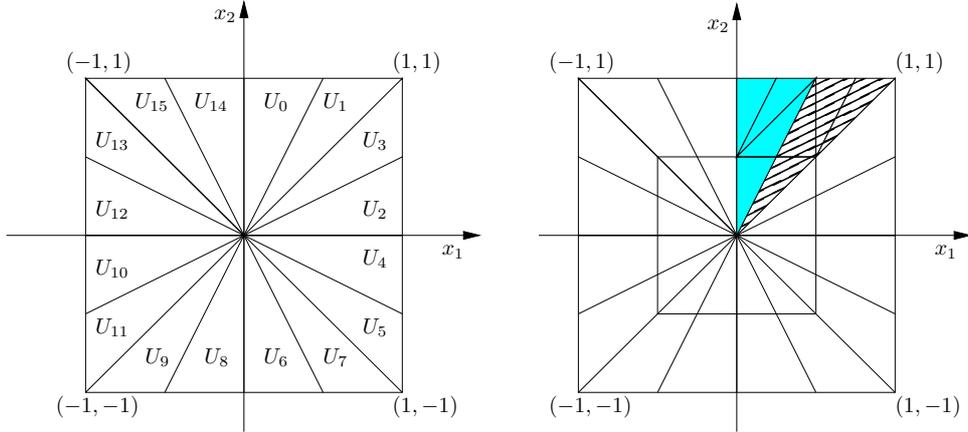\end{aligned}
$$

Figure 1: Construction of scaling functions ((a) coarsest level V_0, (b) refinements).

In the following, we consider the translated versions of ϕ_i with support in $[0,1]^2$ and put them into a vector Φ of length 16,

$$
\begin{aligned}
\Phi &:= (\phi_0, \ldots, \phi_3, \phi_4(\cdot - \tbinom{0}{1}), \ldots, \phi_7(\cdot - \tbinom{0}{1}), \\
&\quad \phi_8(\cdot - \tbinom{1}{1}), \ldots, \phi_{11}(\cdot - \tbinom{1}{1}), \phi_{12}(\cdot - \tbinom{1}{0}), \ldots, \phi_{15}(\cdot - \tbinom{1}{0}))^T.
\end{aligned}
\tag{2.1}
$$

We define now the sequence of spaces $\{V_j\}_{j \in \mathbb{Z}}$ given by

$$V_j := \mathrm{clos}_{L^2(\mathbb{R}^2)} \mathrm{span}\{\phi_{2i,j,k}, \phi_{2i+1,j,k}, : i = 0, \ldots, 7; k \in \mathbb{Z}^2\} \tag{2.2}$$

with

$$
\begin{aligned}
\phi_{2i,j,k}(x) &:= 2^j \phi_0(B_i(2^j x - k)), \\
\phi_{2i+1,j,k}(x) &:= 2^j \phi_1(B_i(2^j x - k)), \quad i = 0, \ldots, 7, k \in \mathbb{Z}^2,
\end{aligned}
$$

i.e., j denotes the scale, k the translation, and i the rotation/shearing. Note that these functions can be understood as scaling functions with composite dilations (see [15, 22]).

We show that $\{V_j\}_{j \in \mathbb{Z}}$ forms a generalized, stationary MRA of $L^2(\mathbb{R}^2)$, that can also be interpreted as a so-called AB-MRA with $A = 2I$ and $B \in \mathcal{B}$ as introduced in [15, 22].

Lemma 2.1. *The sequence $\{V_j\}_{j \in \mathbb{Z}}$ of subspaces of $L^2(\mathbb{R}^2)$ satisfies the following properties:*

1. $V_j \subset V_{j+1} \quad \forall j \in \mathbb{Z}$.

4

2. $\text{clos}_{L^2(\mathbb{R}^2)} \bigcup_{j\in\mathbb{Z}} V_j = L^2(\mathbb{R}^2)$.

3. $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$.

4. $\{\phi_{2i}(\cdot - k), \phi_{2i+1}(\cdot - k) : i = 0, \ldots, 7; k \in \mathbb{Z}^2\}$ *forms a frame of* V_0.

Proof. First observe, that the scaling functions are refinable, compare Figure 1(b). In particular, we have

$$
\begin{aligned}
\phi_0 &= \phi_0(2\cdot) + \phi_0\left(2\cdot - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\right) + \phi_1\left(2\cdot - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\right) + \phi_9\left(2\cdot - \left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)\right) \\
&= \tfrac{1}{2}\left(\phi_{0,1,\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)} + \phi_{0,1,\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)} + \phi_{1,1,\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)} + \phi_{9,1,\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)}\right) \\
\phi_1 &= \phi_1(2\cdot) + \phi_1\left(2\cdot - \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\right) + \phi_0\left(2\cdot - \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)\right) + \phi_8\left(2\cdot - \left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)\right) \\
&= \tfrac{1}{2}\left(\phi_{1,1,\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)} + \phi_{1,1,\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)} + \phi_{0,1,\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)} + \phi_{8,1,\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)}\right).
\end{aligned}
$$

The two-scale relations for the other scaling functions now simply follow as

$$
\begin{aligned}
\phi_{2i} &= \phi_0(B_i\,\cdot) \\
&= \phi_0(2B_i\,\cdot) + \phi_0\left(2B_i\,\cdot - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\right) + \phi_1\left(2B_i\,\cdot - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\right) + \phi_1\left(B_4\left(2B_i\,\cdot - \left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)\right)\right) \\
&= \tfrac{1}{2}\left(\phi_{2i,1,\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)} + \phi_{2i,1,B_i^{-1}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)} + \phi_{2i+1,1,B_i^{-1}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)} + \phi_{(2i+9)\bmod 16,1,B_i^{-1}\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)}\right)
\end{aligned}
$$

as well as

$$
\phi_{2i+1} = \tfrac{1}{2}\left(\phi_{2i+1,1,\left(\begin{smallmatrix}0\\0\end{smallmatrix}\right)} + \phi_{2i+1,1,B_i^{-1}\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)} + \phi_{2i,1,B_i^{-1}\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)} + \phi_{(2i+8)\bmod 16,1,B_i^{-1}\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)}\right).
$$

Hence, $V_0 \subset V_1$ holds. For $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^2$ we get the general refinement equations

$$
\begin{aligned}
\phi_{2i,j,k} &= \phi_0(B_i(2^j\,\cdot - k)) \\
&= \tfrac{1}{2}\Big(\phi_{2i,j+1,2k} + \phi_{2i,j+1,2k+B_i^{-1}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)} + \phi_{2i+1,j+1,2k+B_i^{-1}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)} + \\
&\quad + \phi_{(2i+9)\bmod 16,j+1,2k+B_i^{-1}\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)}\Big) \\
\phi_{2i+1,j,k} &= \phi_1(B_i(2^j\,\cdot - k)) \\
&= \tfrac{1}{2}\Big(\phi_{2i+1,j+1,2k} + \phi_{2i+1,j+1,2k+B_i^{-1}\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)} + \phi_{2i,j+1,2k+B_i^{-1}\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)} + \\
&\quad + \phi_{(2i+8)\bmod 16,j+1,2k+B_i^{-1}\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right)}\Big).
\end{aligned}
$$

Thus we have $V_j \subset V_{j+1}$ for all $j \in \mathbb{Z}$.

Secondly, since the spaces V_j defined in (2.2) contain the subspaces of Haar scaling functions, i.e. $V_j^H \subset V_j$ with

$$
V_j^H := \text{clos}_{L^2(\mathbb{R}^2)}\text{span}\,\{2^j \chi_{[0,1)^2}(2^j\,\cdot - k) : k \in \mathbb{Z}^2\},
$$

we find $\text{clos}_{L^2(\mathbb{R}^2)} \bigcup_{j\in\mathbb{Z}} V_j = L^2(\mathbb{R}^2)$. Further, the condition $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$ easily follows for a stationary sequence $\{V_j\}_{j\in\mathbb{Z}}$ (see e.g. [3, Corollary 4.14]).

Now, the last property remains to proved. The family of functions $\{\phi_{2i}(\cdot - k), \phi_{2i+1}(\cdot - k) : i = 0, \ldots, 7; k \in \mathbb{Z}^2\}$ does not generate a basis of V_0. Obviously, we have the following dependencies

(see Figure 2):

$$
\begin{aligned}
\phi_0 + \phi_1 &= \phi_{10}(\cdot - \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)) + \phi_{11}(\cdot - \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)), \\
\phi_2 + \phi_3 &= \phi_8(\cdot - \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)) + \phi_9(\cdot - \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)), \\
\phi_4 + \phi_5 &= \phi_{14}(\cdot - \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)) + \phi_{15}(\cdot - \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)), \\
\phi_6 + \phi_7 &= \phi_{12}(\cdot - \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)) + \phi_{13}(\cdot - \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)), \\
\phi_0 + \phi_1 + \phi_2 + \phi_3 &= \phi_4(\cdot - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)) + \phi_5(\cdot - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)) + \phi_6(\cdot - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)) + \phi_7(\cdot - \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)).
\end{aligned}
\tag{2.3}
$$

Figure 2: Redundancies of scaling functions.

Indeed, the space V_0 is already generated by the set of 11 functions $\{\phi_{2i} : i = 0, \ldots, 7\} \cup \{\phi_1, \phi_3, \phi_5\}$. The Gram matrix $G := \langle \Phi, \Phi \rangle \in \mathbb{R}^{16 \times 16}$ with Φ in (2.1) is given by

$$
G = \frac{1}{4}
\begin{pmatrix}
I_4 & G_1 & G_2 & G_1^T \\
G_1^T & I_4 & G_1 & G_2 \\
G_2 & G_1^T & I_4 & G_1 \\
G_1 & G_2 & G_1^T & I_4
\end{pmatrix}
\tag{2.4}
$$

with the identity matrix I_4 of size 4×4 and with

$$
G_1 =
\begin{pmatrix}
1/5 & 2/15 & 1/2 & 1/6 \\
7/15 & 1/5 & 1/6 & 1/6 \\
0 & 1/3 & 1/5 & 7/15 \\
1/3 & 1/3 & 2/15 & 1/5
\end{pmatrix},
\qquad
G_2 =
\begin{pmatrix}
0 & 0 & 2/3 & 1/3 \\
0 & 0 & 1/3 & 2/3 \\
2/3 & 1/3 & 0 & 0 \\
1/3 & 2/3 & 0 & 0
\end{pmatrix}.
$$

We observe that $rank(G) = 11$. The nonzero eigenvalues of G provide us with the frame constants of $\{\phi_i(\cdot - k) : i = 0, \ldots, 15; k \in \mathbb{Z}^2\}$; i.e., the inequality

$$
A\|f\|_{L^2(\mathbb{R}^2)}^2 \leqslant \sum_{i=0}^{15} \sum_{k \in \mathbb{Z}^2} |\langle f, \phi_i(\cdot - k) \rangle|^2 \leqslant B\|f\|_{L^2(\mathbb{R}^2)}^2
$$

is satisfied for all $f \in V_0$ with $A \approx 0.0745$ and $B = 1$. $\qquad \square$

Now, we look for a dual frame $\{\tilde{\phi}_i(\cdot - k) : i = 0, \ldots, 15; k \in \mathbb{Z}^2\}$ of V_0 such that

$$
f = \sum_{i=0}^{15} \sum_{k \in \mathbb{Z}^2} \langle f, \tilde{\phi}_i(\cdot - k) \rangle \phi_i(\cdot - k) = \sum_{i=0}^{15} \sum_{k \in \mathbb{Z}^2} \langle f, \phi_i(\cdot - k) \rangle \tilde{\phi}_i(\cdot - k) \qquad \forall f \in V_0.
\tag{2.5}
$$

The dual frame $\tilde{\Phi}$ of Φ can be computed by

$$\tilde{\Phi} = G^\dagger\,\Phi, \tag{2.6}$$

where $G^\dagger$ is the well-defined Moore-Penrose generalized inverse of the Gram matrix G. Indeed, the first part of (2.5) can be seen as follows. Since G is symmetric and positive semidefinite, there exists an orthogonal matrix $P \in \mathbb{R}^{16\times16}$ and a diagonal matrix $D = \mathrm{diag}\,(\lambda_1,\ldots,\lambda_{11},0,\ldots,0) \in \mathbb{R}^{16\times16}$ such that $G = P^T D P$. Hence, the pseudo-inverse $G^\dagger$ is given by $G^\dagger = P^T D^\dagger P$ where $D^\dagger = \mathrm{diag}\,(1/\lambda_1,\ldots,1/\lambda_{11},0,\ldots,0)$. In particular, V_0 is also generated by $P\Phi$, where

$$\langle P\Phi,\, P\Phi\rangle = P\,\langle \Phi, \Phi\rangle\,P^T = P\,G\,P^T = D,$$

i.e., the last five functions in the vector $P\Phi$ are zero functions. Now, for an arbitrary function $g = c^T P\Phi \in V_0$ (restricted to $[0,1]^2$), it follows that

$$\sum_{i=0}^{15}\langle g,\, \tilde{\phi}_i\rangle\,\phi_i = \langle c^T P\Phi,\, G^\dagger\Phi\rangle\Phi = c^T P\langle\Phi,\,\Phi\rangle G^\dagger\Phi = c^T P G G^\dagger\Phi = c^T D D^\dagger P\Phi = c^T P\Phi = g.$$

The second part of (2.5) follows similarly.

The pseudo inverse $G^\dagger$ has again block structure,

$$G^\dagger = \begin{pmatrix} \hat{G}_0 & \hat{G}_1 & \hat{G}_2 & \hat{G}_1^T \\ \hat{G}_1^T & \hat{G}_0 & \hat{G}_1 & \hat{G}_2 \\ \hat{G}_2 & \hat{G}_1^T & \hat{G}_0 & \hat{G}_1 \\ \hat{G}_1 & \hat{G}_2 & \hat{G}_1^T & \hat{G}_0 \end{pmatrix}$$

where

$$\hat{G}_0 = \begin{pmatrix} 3.75 & -2.58 & -0.08 & -0.82 \\ -2.58 & 3.28 & -0.82 & 0.37 \\ -0.08 & -0.82 & 3.75 & -2.58 \\ -0.82 & 0.37 & -2.58 & 3.28 \end{pmatrix}, \quad \hat{G}_1 = \begin{pmatrix} 0.36 & 0.68 & -1.74 & 0.94 \\ -1.16 & 0.36 & 0.94 & 0.10 \\ 1.48 & -0.44 & 0.36 & -1.16 \\ -0.44 & -0.35 & 0.68 & 0.36 \end{pmatrix},$$

and

$$\hat{G}_2 = \begin{pmatrix} -1.21 & 0.29 & -0.66 & 1.83 \\ 0.29 & -0.74 & 1.83 & -1.12 \\ -0.66 & 1.83 & -1.21 & 0.29 \\ 1.83 & -1.12 & 0.29 & -0.74 \end{pmatrix}.$$

(Here, we computed $G^\dagger$ with the common Maple procedure and its rational entries are rounded to two digits.)

3. Construction of a tight directional wavelet frame and reconstruction formulas

Let us now consider the wavelet spaces W_j satisfying the condition $V_j + W_j = V_{j+1}$ for all $j \in \mathbb{Z}$. The locality and refinability of generating functions $\phi_i, i = 0, 1$, imply to consider the wavelet functions for ϕ_0

$$\psi_0^1 := \tfrac{1}{2}\left(\phi_{0,1,\binom{0}{0}} + \phi_{0,1,\binom{0}{1}} - \phi_{1,1,\binom{0}{1}} - \phi_{9,1,\binom{1}{2}}\right),$$

$$\psi_0^2 := \tfrac{1}{2}\left(\phi_{0,1,\binom{0}{0}} - \phi_{0,1,\binom{0}{1}} - \phi_{1,1,\binom{0}{1}} + \phi_{9,1,\binom{1}{2}}\right),$$

$$\psi_0^3 := \tfrac{1}{2}\left(\phi_{0,1,\binom{0}{0}} - \phi_{0,1,\binom{0}{1}} + \phi_{1,1,\binom{0}{1}} - \phi_{9,1,\binom{1}{2}}\right),$$

see Figure 3, and for ϕ_1

$$\psi_1^1 \;:=\; \tfrac{1}{2}\left(\phi_{1,1,\binom{0}{0}} - \phi_{1,1,\binom{1}{1}} + \phi_{0,1,\binom{1}{1}} - \phi_{8,1,\binom{1}{2}}\right),$$

$$\psi_1^2 \;:=\; \tfrac{1}{2}\left(\phi_{1,1,\binom{0}{0}} - \phi_{1,1,\binom{1}{1}} - \phi_{0,1,\binom{1}{1}} + \phi_{8,1,\binom{1}{2}}\right),$$

$$\psi_1^3 \;:=\; \tfrac{1}{2}\left(\phi_{1,1,\binom{0}{0}} + \phi_{1,1,\binom{1}{1}} - \phi_{0,1,\binom{1}{1}} - \phi_{8,1,\binom{1}{2}}\right).$$

The wavelet functions ψ_i^1, ψ_i^2, ψ_i^3 have the same support as ϕ_i, $i = 0, 1$. All further wavelet functions can be obtained by rotation/reflection of these six functions, namely

$$\psi_{2i}^r := \psi_0^r(B_i\cdot) \quad\text{and}\quad \psi_{2i+1}^r := \psi_1^r(B_i\cdot), \quad\text{for}\quad i = 0,\ldots,7,\ r = 1,2,3.$$

Now, we are able to define the wavelet spaces

$$W_j := \operatorname{clos}_{L^2(\mathbb{R}^2)} \operatorname{span}\{\psi_{i,j,k}^r : i = 0,\ldots,15;\ r = 1,2,3;\ k \in \mathbb{Z}^2\},$$

where $\psi_{2i,j,k}^r := 2^j \psi_0^r(B_i(2^j \cdot -k)), \psi_{2i+1,j,k}^r := 2^j \psi_1^r(B_i(2^j \cdot -k))$. The above refinement equations for ψ_0^r and ψ_1^r $(r = 1,2,3)$ directly imply the relations

$$
\begin{aligned}
\psi_{2i,j,k}^1 \;=\;& \tfrac{1}{2}\Big(\phi_{2i,j+1,2k} + \phi_{2i,j+1,2k+B_i^{-1}\binom{0}{1}} - \phi_{2i+1,j+1,2k+B_i^{-1}\binom{0}{1}} \\
& -\phi_{(2i+9)\bmod 16,j+1,2k+B_i^{-1}\binom{1}{2}}\Big), \qquad\qquad\qquad (3.1)\\
\psi_{2i+1,j,k}^1 \;=\;& \tfrac{1}{2}\Big(\phi_{2i+1,j+1,2k} - \phi_{2i+1,j+1,2k+B_i^{-1}\binom{1}{1}} + \phi_{2i,j+1,2k+B_i^{-1}\binom{1}{1}} \\
& -\phi_{(2i+8)\bmod 16,j+1,2k+B_i^{-1}\binom{1}{2}}\Big),
\end{aligned}
$$

and analogous relations for $\psi_{i,j,k}^2$ and $\psi_{i,j,k}^3$. Thus we have $W_j \subset V_{j+1}$ for all $j \in \mathbb{Z}$. Obviously, the wavelet functions $\psi_{i,j,k}^r$, $r = 1, 2, 3$, possess the same compact support as the corresponding scaling functions $\phi_{i,j,k}$ for all i, j, k.

Reconstruction formulas can now be derived as follows (see also Figure 3),

$$\phi_{0,j+1,2k} \;=\; \tfrac{1}{2}\left(\phi_{0,j,k} + \psi_{0,j,k}^1 + \psi_{0,j,k}^2 + \psi_{0,j,k}^3\right),$$

$$\phi_{0,j+1,2k+\binom{0}{1}} \;=\; \tfrac{1}{2}\left(\phi_{0,j,k} + \psi_{0,j,k}^1 - \psi_{0,j,k}^2 - \psi_{0,j,k}^3\right),$$

$$\phi_{0,j+1,2k+\binom{1}{0}} \;=\; \tfrac{1}{2}\left(\phi_{9,j,k+\binom{1}{1}} - \psi_{9,j,k+\binom{1}{1}}^1 + \psi_{9,j,k+\binom{1}{1}}^2 - \psi_{9,j,k+\binom{1}{1}}^3\right),$$

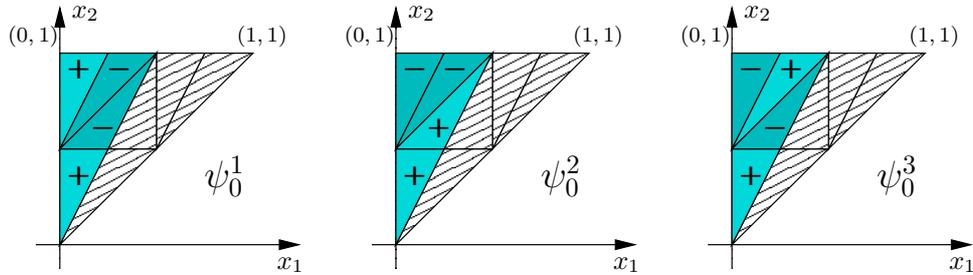$$\phi_{0,j+1,2k+\binom{1}{1}} \;=\; \tfrac{1}{2}\left(\phi_{1,j,k} + \psi_{1,j,k}^1 - \psi_{1,j,k}^2 - \psi_{1,j,k}^3\right),$$

Figure 3: Construction of directional wavelets ψ_0^1, ψ_0^2, and ψ_0^3.

as well as

$$\phi_{1,j+1,2k} = \tfrac{1}{2}\left(\phi_{1,j,k} + \psi^1_{1,j,k} + \psi^2_{1,j,k} + \psi^3_{1,j,k}\right),$$

$$\phi_{1,j+1,2k+\binom{1}{0}} = \tfrac{1}{2}\left(\phi_{8,j,k+\binom{1}{1}} - \psi^1_{8,j,k+\binom{1}{1}} + \psi^2_{8,j,k+\binom{1}{1}} - \psi^3_{8,j,k+\binom{1}{1}}\right),$$

$$\phi_{1,j+1,2k+\binom{0}{1}} = \tfrac{1}{2}\left(\phi_{0,j,k} - \psi^1_{0,j,k} - \psi^2_{0,j,k} + \psi^3_{0,j,k}\right),$$

$$\phi_{1,j+1,2k+\binom{1}{1}} = \tfrac{1}{2}\left(\phi_{1,j,k} - \psi^1_{1,j,k} - \psi^2_{1,j,k} + \psi^3_{1,j,k}\right).$$

The reconstruction formulas for the rotated and reflected functions follow accordingly. Hence, we indeed have

$$V_j + W_j = V_{j+1}.$$

Now we can prove the essential tight frame property of the system

$$\Psi_D := \{\psi^r_{i,j,k} : i = 0,\ldots,15;\ r = 1,2,3;\ j \in \mathbb{Z};\ k \in \mathbb{Z}^2\}$$

generating $L^2(\mathbb{R}^2)$.

Theorem 3.1. *The directional Haar wavelet system Ψ_D forms a Parseval frame of $L^2(\mathbb{R}^2)$, i.e.,*

$$\|f\|^2_{L^2(\mathbb{R}^2)} = \sum_{\psi \in \Psi_D} |\langle f, \psi \rangle|^2 \qquad \forall\, f \in L^2(\mathbb{R}^2).$$

Proof. Firstly, we consider the following subspaces V_j^0, V_j^1, V_j^2, and V_j^3 of V_j given by

$$V_j^\nu := \mathrm{clos}_{L^2(\mathbb{R}^2)}\mathrm{span}\,\{\phi_{2\nu,j,k}, \phi_{2\nu+1,j,k}, \phi_{2\nu+8,j,k}, \phi_{2\nu+9,j,k} : k \in \mathbb{Z}^2\}, \quad \nu = 0,1,2,3. \tag{3.2}$$

From the observations in Section 2 it follows that the sequences $\{V_j^\nu\}_{j\in\mathbb{Z}}$ themselves already form a multiresolution of $L^2(\mathbb{R}^2)$, and moreover, the generating functions form an orthogonal basis of V_j^ν, where $\|\phi_{i,j,k}\|^2_{L^2(\mathbb{R}^2)} = \tfrac{1}{4}$ for arbitrary $i = 0,\ldots,15; j \in \mathbb{Z}; k \in \mathbb{Z}^2$. Now, taking the subspaces W_j^0, W_j^1, W_j^2, and W_j^3 of W_j in the same manner, i.e.,

$$W_j^\nu := \mathrm{clos}_{L^2(\mathbb{R}^2)}\mathrm{span}\,\{\psi^r_{2\nu,j,k}, \psi^r_{2\nu+1,j,k}, \psi^r_{2\nu+8,j,k}, \psi^r_{2\nu+9,j,k} : r = 1,2,3;\ k \in \mathbb{Z}^2\},$$

for $\nu = 0,1,2,3$, we find that $W_j^\nu \perp V_j^\nu$, and for each $\nu = 0,1,2,3$, this generating system is even an orthogonal basis of W_j^ν. Hence each of the function sets

$$\Psi^\nu := \{\psi^r_{2\nu,j,k}, \psi^r_{2\nu+1,j,k}, \psi^r_{2\nu+8,j,k}, \psi^r_{2\nu+9,j,k} : r = 1,2,3;\ j \in \mathbb{Z};\ k \in \mathbb{Z}^2\}, \quad \nu = 0,1,2,3,$$

forms an orthogonal basis of $L^2(\mathbb{R}^2)$, and the Parseval identity implies

$$\|f\|^2_{L^2(\mathbb{R}^2)} = \sum_{\psi \in \Psi^\nu} \frac{|\langle f, \psi \rangle|^2}{\langle \psi, \psi \rangle} = 4 \sum_{\psi \in \Psi^\nu} |\langle f, \psi \rangle|^2$$

for all $f \in L^2(\mathbb{R}^2)$, $\nu = 0,1,2,3$, such that the complete system Ψ_D forms a tight frame of $L^2(\mathbb{R}^2)$ with

$$\|f\|^2_{L^2(\mathbb{R}^2)} = \sum_{\psi \in \Psi_D} |\langle f, \psi \rangle|^2.$$

$\square$

Remark 3.2.   1. *An alternative proof of the tight frame property is given by one of the authors in [19] with arguments in the frequency domain.*

  2. *Note the important fact that the directional wavelet frame consists of four orthogonal bases, i.e., it can be interpreted as a redundant dictionary. In our applications we will exploit this fact in order to get an efficient implementation.*

4. Directional Haar wavelet filter bank

Let $J := \{0,\ldots,2N_1-1\} \times \{0,\ldots,2N_2-1\}$ be the index set of a digital image $a = (a_k)_{k \in J}$ with $2N_1 \times 2N_2$ pixels. The function $f \in L^2(\Omega)$ with $\Omega = [0,N_1] \times [0,N_2]$,

$$f(x_1,x_2) = \sum_{k_1=0}^{2N_1-1} \sum_{k_2=0}^{2N_2-1} a_{k_1,k_2} \cdot \chi_{[0,1)^2}(2x_1-k_1,2x_2-k_2) = \sum_{k \in J} a_k \cdot \chi_{[0,1)^2}(2x-k), \qquad (4.1)$$

can be seen as the corresponding 'L^2-version' of the discrete image a. Here, $\chi_{[0,1)^2}$ denotes the characteristic function on $[0,1)^2$ and we assume that $N_1 = n_1 \cdot 2^{j_0}$, $N_2 = n_2 \cdot 2^{j_0}$ with some $n_1, n_2 \in \mathbb{N}$ and a fixed $j_0 \in \mathbb{N}$. We want to apply our redundant Haar wavelet frame constructed above for an efficient analysis of f. Let us shortly describe the procedure, before going into a detailed analysis of the single steps.

First we compute an *orthogonal projection* f_0 (resp. f_j) of f into the space V_0 defined in (2.2) or into a coarser space V_j with $j < 0$ (see Subsection 4.1). Then we apply the directional Haar wavelet filter bank generated by the decomposition and reconstruction formulas for $\phi_i, \psi_i^1, \psi_i^2, \psi_i^3$, $i = 0,\ldots,15$, in order to decompose f_j into $f_{j-1} \in V_{j-1}$ and $g_{j-1} \in W_{j-1}$ as usual. Using the fact that our constructed frame can be split into four bases, the decomposition can be done by a fourfold application of the fast wavelet transform (FWT), see Subsection 4.2. If we use the directional wavelet frame for image denoising, we do not reduce the redundancies because redundant information is desirable with denoising. By contrast, if we apply the filter bank algorithm to find a sparse image representation, we have to reduce redundancies. This issue will be considered in detail in Section 5.

4.1. Orthogonal projection of f into V_0

In order to apply the directional Haar wavelet frames constructed in the preceding sections, we need a suitable projection f_0 of a given function f of the form (4.1) into the scaling space V_0 defined in (2.2). For this projection we require two conditions. Firstly, the redundancy introduced by this projection, i.e., the ratio between the number of coefficients determining f_0 and the $4N_1N_2$ coefficients determining f should be as small as possible. Secondly, there should be no loss of information, i.e., we desire that f can be perfectly reconstructed from f_0.

We are interested in the orthogonal projection of f in (4.1) into the space V_0 of the form

$$f_0 = \sum_{k \in J_1} (c_k^0)^T \Phi(\cdot - k), \qquad (4.2)$$

where $J_1 := \{0,\ldots,N_1-1\} \times \{0,\ldots,N_2-1\}$, $c_k^0 = (c_{0,k}^0,\ldots,c_{15,k}^0)^T \in \mathbb{R}^{16}$ and where the support of all functions in Φ is contained in $[0,1]^2$ (see (2.1)). Since the basis functions $\chi_{[0,1)^2}(2x-k)$ in (4.1) as well as the scaling functions in $\Phi(\cdot - k)$ have small compact support, we can look at the

projection problem locally. We restrict ourselves to the case $k = 0$ and consider the area $[0,1)^2$. Hence we need a projection of

$$f|_{[0,1)^2}(x) = a_{\binom{0}{0}}\chi_{[0,1)^2}(2x) + a_{\binom{1}{0}}\chi_{[0,1)^2}\left(2x - \binom{1}{0}\right) + a_{\binom{0}{1}}\chi_{[0,1)^2}\left(2x - \binom{0}{1}\right) + a_{\binom{1}{1}}\chi_{[0,1)^2}\left(2x - \binom{1}{1}\right)$$

to $f_0|_{[0,1)^2}(x) = (c_0^0)^T \Phi(x)$. Obviously, such a projection provides the redundancy factor 4. Using the dual canonical frame $\tilde{\Phi} = G^\dagger \Phi$ defined in (2.6), the coefficient vector $c_0^0 \in \mathbb{R}^{16}$ is now given by

$$\begin{aligned}
c_0^0 &= \langle \tilde{\Phi}, f \rangle = G^\dagger \langle \Phi, f \rangle \\
&= a_{\binom{0}{0}} G^\dagger \langle \Phi, \chi_{[0,1)^2}(2\cdot) \rangle + a_{\binom{1}{0}} G^\dagger \langle \Phi, \chi_{[0,1)^2}\left(2 \cdot - \binom{1}{0}\right) \rangle \\
&\quad + a_{\binom{0}{1}} G^\dagger \langle \Phi, \chi_{[0,1)^2}\left(2 \cdot - \binom{0}{1}\right) \rangle + a_{\binom{1}{1}} G^\dagger \langle \Phi, \chi_{[0,1)^2}\left(2 \cdot - \binom{1}{1}\right) \rangle.
\end{aligned}$$

The vectors

$$\langle \Phi, \chi_{[0,1)^2}(2 \cdot -l) \rangle = \tfrac{1}{4} \int\limits_{[0,1)^2} \Phi\left(\tfrac{y+l}{2}\right) dy, \qquad l \in \left\{\binom{0}{0}, \binom{1}{0}, \binom{0}{1}, \binom{1}{1}\right\},$$

in $\mathbb{R}^{16}$ can now easily be computed, and we find

$$\begin{aligned}
M &:= \left(\langle \Phi, \chi_{[0,1)^2}(2\cdot) \rangle, \langle \Phi, \chi_{[0,1)^2}\left(2 \cdot - \binom{1}{0}\right) \rangle, \langle \Phi, \chi_{[0,1)^2}\left(2 \cdot - \binom{0}{1}\right) \rangle, \langle \Phi, \chi_{[0,1)^2}\left(2 \cdot - \binom{1}{1}\right) \rangle \right) \\
&= \frac{1}{16} \begin{pmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 3 & 1 & 0 & 2 & 0 & 2 & 3 & 1 & 0 & 0 \\
0 & 0 & 3 & 1 & 0 & 2 & 0 & 2 & 3 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\
3 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 3 & 1 & 0 & 2 & 0 & 2 \\
0 & 2 & 0 & 2 & 3 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 3 & 1
\end{pmatrix}^T.
\end{aligned}$$

Hence, the coefficient vector c_0^0 is found by

$$c_0^0 = G^\dagger M \left(a_{\binom{0}{0}}, a_{\binom{1}{0}}, a_{\binom{0}{1}}, a_{\binom{1}{1}} \right)^T.$$

Generally, the coefficient vectors c_k^0 in (4.2) are obtained for all $k \in J_1$ by

$$c_k^0 = G^\dagger M \left(a_{2k}, a_{2k+\binom{1}{0}}, a_{2k+\binom{0}{1}}, a_{2k+\binom{1}{1}} \right)^T.$$

Next, we will show that this projection f_0 in (4.2) contains the full information of f. In other words, f can be perfectly reconstructed from f_0. This can be seen as follows. Consider the subspace V_0^0 of V_0 containing only the $\mathbb{Z}^2$-translates of ϕ_0, ϕ_1, ϕ_8, and ϕ_9, we obtain an orthogonal projection of f in (4.1) into the subspace V_0^0 by

$$f_0^0 = \sum_{k \in J_1} c_{0,k}^0 \, \phi_{0,0,k} + c_{1,k}^0 \, \phi_{1,0,k} + c_{8,k}^0 \, \phi_{8,0,k+\binom{0}{1}} + c_{9,k}^0 \, \phi_{9,0,k+\binom{1}{1}},$$

where, with the same arguments as above,

$$\begin{pmatrix} c_{0,k}^0 \\ c_{1,k}^0 \\ c_{8,k}^0 \\ c_{9,k}^0 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 0 & 3 & 0 \\ 1 & 0 & 1 & 2 \\ 0 & 3 & 0 & 1 \\ 2 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{2k} \\ a_{2k+\binom{1}{0}} \\ a_{2k+\binom{0}{1}} \\ a_{2k+\binom{1}{1}} \end{pmatrix}. \tag{4.3}$$
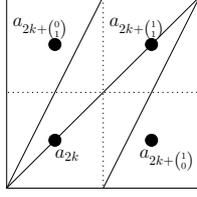
Figure 4: Computation of the coefficients $c_{0,k}^0, c_{1,k}^0, c_{8,k}^0, c_{9,k}^0$ from four image pixels.

Here, the coefficient matrix contains the $0., 1., 8.$ and $9.$ row of M, because these rows of M provide the coefficients of $\phi_0, \phi_1, \phi_{8,0,\binom{1}{1}}$, and $\phi_{9,0,\binom{1}{1}}$, see Figure 4. Further, since $\phi_0, \phi_1, \phi_{8,0,\binom{1}{1}}, \phi_{9,0,\binom{1}{1}}$ are orthogonal, the corresponding Gramian matrix has the form $\frac{1}{4}I_4$.

Since the coefficient matrix in (4.3) is invertible, we can reconstruct f from f_0^0. But V_0^0 is a subspace of V_0, and in particular it follows that f_0^0 is also found as the orthogonal projection of f_0 into V_0^0. Hence, taking

$$
\begin{aligned}
c_{r,k}^0 &= \langle f_0, \phi_{r,0,k} \rangle = \langle (c_k^0)^T \Phi(\cdot - k), \phi_{r,0,k} \rangle = (c_k^0)^T g_r, & r = 0, 1, \\
c_{r+8,k}^0 &= \langle f_0, \phi_{r+8,0,k+\binom{1}{1}} \rangle = \langle (c_k^0)^T \Phi(\cdot - k), \phi_{r+8,0,k+\binom{1}{1}} \rangle = (c_k^0)^T g_{r+8}, & r = 0, 1,
\end{aligned}
$$

where g_r is the r-th column vector of the Gramian matrix G, we obtain for all $k \in J_1$

$$
\begin{pmatrix} c_{0,k}^0 \\ c_{1,k}^0 \\ c_{8,k}^0 \\ c_{9,k}^0 \end{pmatrix} = (g_0, g_1, g_8, g_9)^T c_k^0 = G'^T c_k^0,
$$

where $G' \in \mathbb{R}^{16 \times 4}$ contains the $0., 1., 8.$ and $9.$ column of G, and with (4.3) we have the reconstruction formula

$$
\begin{pmatrix} a_{2k} \\ a_{2k+\binom{1}{0}} \\ a_{2k+\binom{0}{1}} \\ a_{2k+\binom{1}{1}} \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & -3 & -3 & 9 \\ 1 & -3 & 5 & 1 \\ 5 & 1 & 1 & -3 \\ -3 & 9 & 1 & -3 \end{pmatrix} G'^T c_k^0.
$$

4.2. Directional wavelet filter bank algorithm

Let now $f_0 \in V_0$ be given as in (4.2). We want to derive an efficient algorithm for the decomposition of f_0 into $f_{-1} \in V_{-1}$ and $g_{-1} \in W_{-1}$ and for the reconstruction $f_{-1} + g_{-1}$. With $J_2 := \{0, \dots, \frac{N_1}{2} - 1\} \times \{0, \dots, \frac{N_2}{2} - 1\}$ we can write

$$
\begin{aligned}
f_0 &= \sum_{k \in J_2} (c_{2k}^0)^T \Phi(\cdot - 2k) + (c_{2k+\binom{1}{0}}^0)^T \Phi(\cdot - (2k + \binom{1}{0})) \\
&\quad + (c_{2k+\binom{0}{1}}^0)^T \Phi(\cdot - (2k + \binom{0}{1})) + (c_{2k+\binom{1}{1}}^0)^T \Phi(\cdot - (2k + \binom{1}{1})).
\end{aligned}
$$

Again, we derive the decomposition of f_0 locally. On $[0,2]^2$, there are 64 basis functions of V_0, namely the 16 components of Φ (as given in (2.1)) with a support inside $[0,1]^2$ and the components of $\Phi(\cdot - \binom{1}{0})$, $\Phi(\cdot - \binom{0}{1})$, and $\Phi(\cdot - \binom{1}{1})$. For the filter bank algorithm, we want to apply the

12

knowledge that V_j is composed by the subspaces V_j^0, V_j^1, V_j^2, and V_j^3 as given in the proof of Theorem 3.1 (see (3.2)), i.e.,

$$V_j = V_j^0 + V_j^1 + V_j^2 + V_j^3.$$

Therefore, we reorder the frame functions in $(\Phi^T, \Phi(\cdot - \binom{1}{0}))^T, \Phi(\cdot - \binom{0}{1}))^T, \Phi(\cdot - \binom{1}{1})^T)$ according to the subspaces V_0^ν, $\nu = 0, 1, 2, 3$. In $[0,2)^2$ we hence consider the function vectors

$$\Phi_0^0 \; := \; \Big(\phi_{0,0,\binom{0}{0})}, \phi_{9,0,\binom{1}{2})}, \phi_{0,0,\binom{0}{1})}, \phi_{1,0,\binom{0}{1})}, \phi_{1,0,\binom{0}{0})}, \phi_{8,0,\binom{1}{2})}, \phi_{0,0,\binom{1}{1})}, \phi_{1,0,\binom{1}{1})},$$

$$\phi_{8,0,\binom{2}{2})}, \phi_{1,0,\binom{1}{0})}, \phi_{8,0,\binom{2}{1})}, \phi_{9,0,\binom{2}{1})}, \phi_{9,0,\binom{2}{2})}, \phi_{0,0,\binom{1}{0})}, \phi_{8,0,\binom{1}{1})}, \phi_{9,0,\binom{1}{1})} \Big)^T$$

as well as

$$\Phi_0^\nu := \Phi_0^0(B_\nu \cdot), \quad \nu = 1, 2, 3.$$

The order of functions in Φ_0^ν is taken for simplifying the refinement relations and the application of the corresponding wavelet filter bank. See Figure 5 for the new order of frame functions in the four directions related to the four subspaces V_0^ν.
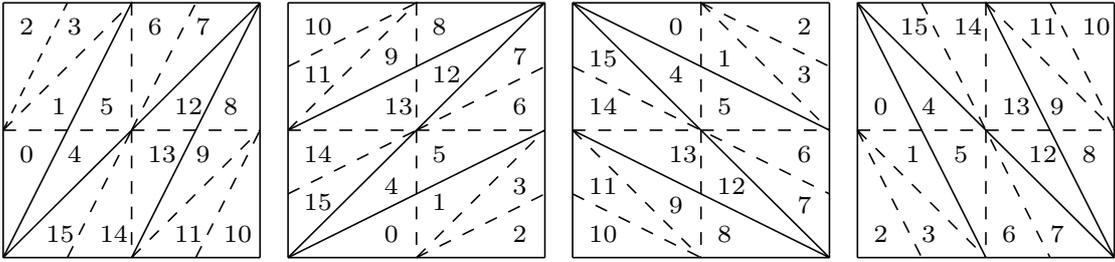

Figure 5: Order of components in the function vectors Φ_0^ν for $\nu = 0$ (left), $\nu = 1$ (middle left), $\nu = 2$ (middle right), and $\nu = 3$ (right). The figures indicate the supports of the i-th component of Φ_0^ν for $i = 0, \dots, 15$.

By corresponding reordering of the coefficients of f_0 in the coefficient vectors c_{2k+l}^0, $l \in \{\binom{0}{0}, \binom{1}{0}, \binom{0}{1}, \binom{1}{1}\}$, we obtain

$$f_0|_{[0,2)^2} = (d_0^0)^T \Phi_0^0 + (d_0^1)^T \Phi_0^1 + (d_0^2)^T \Phi_0^2 + (d_0^3)^T \Phi_0^3.$$

Since Φ_0^ν contains the 16 frame functions in $[0,2)^2$ that correspond only to the direction ν, we can separately apply the decomposition formulas (3.1). For every $\nu = 0, 1, 2, 3$, we find with

$$\Psi_{-1}^0 \; := \; \Big(\phi_{0,-1,\binom{0}{0})}, \psi_{0,-1,\binom{0}{0})}^1, \psi_{0,-1,\binom{0}{0})}^2, \psi_{0,-1,\binom{0}{0})}^3, \phi_{1,-1,\binom{0}{0})}, \psi_{1,-1,\binom{0}{0})}^1, \psi_{1,-1,\binom{0}{0})}^2, \psi_{1,-1,\binom{0}{0})}^3,$$

$$\phi_{8,-1,\binom{1}{1})}, \psi_{8,-1,\binom{1}{1})}^1, \psi_{8,-1,\binom{1}{1})}^2, \psi_{8,-1,\binom{1}{1})}^3, \phi_{9,-1,\binom{1}{1})}, \psi_{9,-1,\binom{1}{1})}^1, \psi_{9,-1,\binom{1}{1})}^2, \psi_{9,-1,\binom{1}{1})}^3 \Big)^T$$

and

$$\Psi_{-1}^\nu := \Psi_{-1}^0(B_\nu \cdot), \qquad \nu = 0, 1, 2, 3,$$

the relation

$$\Psi_{-1}^\nu = A \Phi_0^\nu, \qquad \nu = 0, 1, 2, 3,$$

13

where the orthogonal matrix $A \in \mathbb{R}^{16 \times 16}$ is a tensor product matrix of the form

$$A = (I_4 \otimes B) = \begin{pmatrix} B & & & \\ & B & & \\ & & B & \\ & & & B \end{pmatrix} \quad \text{with} \quad B := \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Observe that Ψ_{-1}^{ν} generates the directional subspace $V_{-1}^{\nu} \oplus W_{-1}^{\nu}$, $\nu = 0, 1, 2, 3$. Now we denote with

$$\Phi_0 := (\Phi_0^{0^T}, \Phi_0^{1^T}, \Phi_0^{2^T}, \Phi_0^{3^T})^T \tag{4.4}$$

the (reordered) vector of functions generating the frame in V_0 (restricted to $[0, 2)^2$) and with $\Psi_{-1} := (\Psi_{-1}^{0^T}, \Psi_{-1}^{1^T}, \Psi_{-1}^{2^T}, \Psi_{-1}^{3^T})^T$ the generating frame in $V_{-1} \oplus W_{-1}$ (restricted to $[0, 2)^2$). Both, Φ_0 and Ψ_{-1} are now function vectors of length 64.

Using $A = A^T = A^{-1}$, the representation of f with generating functions from V_0 resp. $V_{-1} \oplus W_{-1}$ can be given in the form

$$f_0|_{[0,2)^2} = ((d_0^0)^T, (d_0^1)^T, (d_0^2)^T, (d_0^3)^T) \, \Phi_0 = D_{0,0}^T \, \Phi_0 = D_{0,0}^T \, (I_4 \otimes A) \Psi_{-1},$$

where $D_{0,0}^T := ((d_0^0)^T, (d_0^1)^T, (d_0^2)^T, (d_0^3)^T) \in \mathbb{R}^{64}$. For the complete image f_0 it follows the decomposition

$$f_0 = \sum_{k \in J_2} D_{0,k}^T \, \Phi_0(\cdot - 2k) = \sum_{k \in J_2} D_{0,k}^T \, (I_4 \otimes A) \Psi_{-1}(\cdot - 2k), \tag{4.5}$$

with $D_{0,k}^T := ((d_{0,k}^0)^T, (d_{0,k}^1)^T, (d_{0,k}^2)^T, (d_{0,k}^3)^T) \in \mathbb{R}^{64}$. Summing up, the decomposition algorithm for the constructed directional wavelets on triangles has the following form.

Algorithm 1a (Decomposition by directional Haar wavelet filter bank)

1. Input: Initial image obtained by orthogonal projection of f into V_0,

$$f_0 = \sum_{k \in J_1} (c_{0,k}^0)^T \, \Phi(\cdot - k), \qquad c_{0,k}^0 = \langle f, \tilde{\Phi}(\cdot - k) \rangle.$$

2. Reorder the frame functions (resp. corresponding coefficients) by directions. Let P_1 be the permutation matrix for reordering of generating functions in V_0,

$$(\Phi^T, \Phi(\cdot - \begin{pmatrix} 1 \\ 0 \end{pmatrix})^T, \Phi(\cdot - \begin{pmatrix} 0 \\ 1 \end{pmatrix})^T, \Phi(\cdot - \begin{pmatrix} 1 \\ 1 \end{pmatrix})^T) P_1 = \Phi_0^T.$$

Now, for each $k \in J_2$, reorder the coefficient vectors,

$$D_{0,k}^T = ((d_{0,k}^0)^T, (d_{0,k}^1)^T, (d_{0,k}^2)^T, (d_{0,k}^3)^T) = ((c_{0,2k}^0)^T, (c_{0,2k+\binom{1}{0}}^0)^T, (c_{0,2k+\binom{0}{1}}^0)^T, (c_{0,2k+\binom{1}{1}}^0)^T) P_1,$$

such that

$$f_0 = \sum_{k \in J_2} D_{0,k}^T \, \Phi_0(\cdot - 2k).$$

3. Decompose $f_0 \in V_0$ into $f_{-1} \in V_{-1}$ and $g_{-1} \in W_{-1}$ using the relation $\Phi_0 = (I_4 \otimes A) \, \Psi_{-1}$. Compute the corresponding coefficients by

$$D_{-1,k}^T = D_{0,k}^T \, (I_4 \otimes A).$$

14

4. Using the definition of Ψ_{-1}, reorder the coefficients in the vectors

$$D_{-1,k}^T = ((d_{-1,k}^0)^T, (d_{-1,k}^1)^T, (d_{-1,k}^2)^T, (d_{-1,k}^3)^T), \qquad k \in J_2,$$

in order to obtain f_{-1} and g_{-1}.

5. Iterative application: Apply the same procedure to the low pass part f_{-1}, while the high pass part g_{-1} is stored.

As previously described, the decomposition algorithm of $f \in V_0$ into $f_{-1} \in V_{-1}$ and $g_{-1} \in W_{-1}$ only involves some permutations and some additions/subtractions. Since the transformation matrix $(I_4 \otimes A)$ is orthogonal, the algorithm is numerically stable. The reconstruction procedure then easily follows by reversing the steps of Algorithm 1a.

5. Sparse image representations in wavelet spaces

At present, a usual approach to find a sparse representation of f_0 in a redundant dictionary is the orthogonal matching pursuit (OMP) (see e.g. [12, 26] and references therein). OMP is an iterative greedy algorithm that selects at each step the dictionary element best correlated with the residual part of the signal. Then, a new approximation of the signal is produced by a projection on the dictionary elements that have already been selected. Unfortunately, because of the large coherence of the considered dictionary of Haar wavelet functions, the OMP algorithm does not provide satisfying sparse representations of $f_0 \in V_0$ in our case.

In order to get a sparse representation of images, we need to reduce the existing (fourfold) redundancy by exploiting our explicit knowledge about it. The idea is as follows. After decomposing a given image $f_j \in V_j$ into $f_{j-1} \in V_{j-1}$ and $g_{j-1} \in W_{j-1}$ we aim to exploit the redundancy of the frames generating V_{j-1} and W_{j-1} and try to find a representation of f_{j-1} and g_{j-1} that contains as many zero coefficients as possible. The procedure will be applied after each decomposition step. Finally, using a threshold procedure to remove remaining small frame coefficients, we obtain a suitable sparse approximation of the image, where, due to the frame construction, different directions of the image are well adapted.

Again, we use the decomposition $V_0 = V_0^0 + V_0^1 + V_0^2 + V_0^3$, and recall that V_0 is spanned by $\{\Phi_0(\cdot - 2k) : k \in \mathbb{Z}^2\}$, where Φ_0 is the function vector of length 64 defined in (4.4), and $V_0^\nu = \text{span}\,\{\Phi_0^\nu(\cdot - 2k) : k \in \mathbb{Z}^2\}$ for $\nu = 0, 1, 2, 3$, see Figure 5.

Considering (4.5), we note that the representations of f_0 in V_0 as well as in $V_{-1} + W_{-1}$ are not uniquely determined. The dependence relations (2.3) in V_0 imply with unit vectors $e_k := (\delta_{k,l})_{l=0}^{15}$ the equations

$$
\begin{aligned}
(e_0 + e_4)^T \Phi_0^\nu &- (e_{14} + e_{15})^T \Phi_0^{\nu+1} = 0, \\
(e_2 + e_3)^T \Phi_0^\nu &- (e_{10} + e_{11})^T \Phi_0^{\nu+1} = 0, \\
(e_9 + e_{13})^T \Phi_0^\nu &- (e_1 + e_5)^T \Phi_0^{\nu+1} = 0, \\
(e_6 + e_7)^T \Phi_0^\nu &- (e_8 + e_{12})^T \Phi_0^{\nu+1} = 0, \\
(e_0 + e_4)^T \Phi_0^{\nu+1} &- (e_{14} + e_{15})^T \Phi_0^\nu = 0, \\
(e_2 + e_3)^T \Phi_0^{\nu+1} &- (e_{10} + e_{11})^T \Phi_0^\nu = 0, \\
(e_9 + e_{13})^T \Phi_0^{\nu+1} &- (e_1 + e_5)^T \Phi_0^\nu = 0, \\
(e_6 + e_7)^T \Phi_0^{\nu+1} &- (e_8 + e_{12})^T \Phi_0^\nu = 0,
\end{aligned}
$$

15

for $\nu = 0, 1$, and

$$
\begin{aligned}
(e_0 + e_4)^T (\Phi_0^0 + \Phi_0^1) &- (e_9 + e_{13})^T \Phi_0^2 - (e_2 + e_3)^T \Phi_0^3 = 0, \\
(e_9 + e_{13})^T \Phi_0^0 + (e_2 + e_3)^T \Phi_0^1 &- (e_6 + e_7)^T (\Phi_0^2 + \Phi_0^3) = 0, \\
(e_2 + e_3)^T \Phi_0^0 + (e_9 + e_{13})^T \Phi_0^1 &- (e_0 + e_4)^T (\Phi_0^2 + \Phi_0^3) = 0, \\
(e_6 + e_7)^T (\Phi_0^0 + \Phi_0^1) &- (e_2 + e_3)^T \Phi_0^2 - (e_9 + e_{13})^T \Phi_0^3 = 0.
\end{aligned}
$$

These 20 relations directly provide a matrix $U \in \mathbb{R}^{20 \times 64}$ containing these dependencies, such that

$$
U \left((\Phi_0^0)^T, (\Phi_0^1)^T, (\Phi_0^2)^T, (\Phi_0^3)^T \right)^T = U \, \Phi_0 = 0 \quad \text{on } [0, 2)^2. \tag{5.1}
$$

Hence we obtain from (4.5) for arbitrary vectors $g_k \in \mathbb{R}^{20}$, $k \in J_2$, a redundant representation of $f_0 \in V_0$ of the form

$$
\begin{aligned}
f_0 &= \sum_{k \in J_2} D_{0,k}^T \, \Phi_0(\cdot - 2k) = \sum_{k \in J_2} \left(D_{0,k}^T + g_k^T U \right) \Phi_0(\cdot - 2k) \\
&= \sum_{k \in J_2} \left(D_{0,k}^T + g_k^T U \right) (I_4 \otimes A) \Psi_{-1}(\cdot - 2k).
\end{aligned}
$$

We aim to represent f_0 in $V_{-1} + W_{-1}$ with the smallest possible number of nonzero wavelet coefficients. Observe that because of the local supports of Φ_0 resp. Ψ_{-1}, this problem can be considered separately for each $k \in J_2$. Thus, for each $k \in J_2$ we have to determine a vector $g_k \in \mathbb{R}^{20}$ such that the l^0-seminorm

$$
\left\| \left(D_{0,k}^T + g_k^T U \right) (I_4 \otimes A) \right\|_0
$$

is minimized, where the l^0-seminorm of a vector simply counts the number of its nonzero components. This minimization leads to a large amount of vanishing wavelet frame coefficients.

The modified decomposition algorithm has the following form.

Algorithm 1b (Decomposition with redundancy reduction)

1. Input: Initial image in V_0 by orthogonal projection of f into V_0,

$$
f_0 = \sum_{k \in J_1} (c_{0,k}^0)^T \, \Phi(\cdot - k), \qquad c_{0,k}^0 = \langle f_0, \tilde{\Phi}(\cdot - k) \rangle.
$$

2. Reorder the basis functions (resp. corresponding coefficients) by directions (see step 2 of Algorithm 1a),

$$
f_0 = \sum_{k \in J_2} D_{0,k}^T \, \Phi_0(\cdot - 2k).
$$

3. Add redundancies in V_0 and apply the transform to $V_{-1} + W_{-1}$:

$$
f_0 = \sum_{k \in J_2} (D_{0,k}^T + g_k^T U) \, \Phi_0(\cdot - 2k) = \sum_{k \in J_2} (D_{0,k}^T + g_k^T U)(I_4 \otimes A) \, \Psi_{-1}(\cdot - 2k).
$$

4. For each $k \in J_2$ compute $g_k \in \mathbb{R}^{20}$ such that the l^0-seminorm

$$
\left\| (D_{0,k}^T + g_k^T U)(I_4 \otimes A) \right\|_0 = \left\| D_{-1,k}^T + g_k^T U(I_4 \otimes A) \right\|_0
$$

with $D_{-1,k}^T = D_{0,k}^T(I_4 \otimes A)$ becomes minimal.

16

5. For each $k \in J_2$ let $\tilde{D}^T_{-1,k} = D^T_{-1,k} + \tilde{g}^T_k U(I_4 \otimes A)$ be this minimized coefficient vector, where

$$\tilde{g}_k := \arg \min_{g_k \in \mathbb{R}^{20}} \left\| D^T_{-1,k} + g^T_k U(I_4 \otimes A) \right\|_0.$$

   Compute the sparse representation

$$f_0 = \sum_{k \in J_2} \tilde{D}^T_{-1,k} \, \Psi_{-1}(\cdot - 2k),$$

   and determine $f_{-1} \in V_{-1}$ and $g_{-1} \in W_{-1}$ from this representation.

6. Iterative application: Apply the same procedure to the low pass part f_{-1}, while the high pass part g_{-1} is stored.

Let us now focus on the local minimization problem in step 4,

$$\arg \min_{g_k \in \mathbb{R}^{20}} \{ \| D_{-1,k} + (I_4 \otimes A) \, U^T g_k \|_0 \} = \arg \min_{g_k \in \mathbb{R}^{20}} \{ \| D_{-1,k} + R \, g_k \|_0 \}, \tag{5.2}$$

that has to be solved for each $k \in J_2$, and where $D_{-1,k} \in \mathbb{R}^{64}$ as well as the matrix $R := (I_4 \otimes A) U^T \in \mathbb{R}^{64 \times 20}$ are given. Observe that R has full rank 20. A naive approach to the problem is to consider all possibilities to take 20 linearly independent rows of R to build a matrix $C_k \in \mathbb{R}^{20 \times 20}$ and to solve the system

$$C_k \, g_k = -(D_{-1,k})_q,$$

where the vector $(D_{-1,k})_q \in \mathbb{R}^{20}$ is a subvector of $D_{-1,k}$ obtained by taking the 20 components of $D_{-1,k}$ that correspond to the 20 rows of R generating C_k. Then, the vector

$$D_{-1,k} + R \, g_k = D_{-1,k} - R \, C_k^{-1} (D_{-1,k})_q$$

contains at least 20 zeros. All vectors $D_{-1,k} + R \, g_k$ obtained in this manner need to be compared with respect to their l^0-seminorm. Obviously, such a procedure is inefficient for our purposes. Unfortunately, the idea of replacing the l^0-seminorm by the l^1-norm does not work in our case. For example, for constant parts of the image, i.e. $f_0 = c$ on $[0,2)^2$, the coefficient vectors in the two representations

$$f_0 = c \, (1^T_{16}, 0^T_{16}, 0^T_{16}, 0^T_{16}) \, \Phi_0 = c \, (w^T, 0^T_{16}, 0^T_{16}, 0^T_{16}) \, \Psi_{-1}$$

and

$$f_0 = \frac{c}{4} \, (1^T_{16}, 1^T_{16}, 1^T_{16}, 1^T_{16}) \, \Phi_0 = \frac{c}{4} \, (w^T, w^T, w^T, w^T) \, \Psi_{-1}$$

of f_0 have the same l^1-norm while their l^0-seminorm strongly differs. Here, $1_{16} := (1, \ldots, 1)^T \in \mathbb{R}^{16}$, 0_{16} denotes the zero vector of length 16, and

$$w^T := (2, 0, 0, 0, 2, 0, 0, 0, 2, 0, 0, 0, 2, 0, 0, 0) \in \mathbb{R}^{16}. \tag{5.3}$$

We are interested in a simple algorithm to find a nearly optimal local representation of the image in the sense of (5.2), that also uses our explicit knowledge about the special properties of the redundant function system and especially on the dependence relations of the system collected in the matrix U. Therefore, we propose a new method, that can be proved to be optimal for piecewise

17

constant images. Further, we will show that the sparse representation found by our algorithm does not depend on the representation of the signal to start with.

The algorithm is based on the following idea. We consider the local orthogonal projections of the image f_0 into the subspaces $V_{-1}^{\nu} \oplus W_{-1}^{\nu}$, $\nu = 0, 1, 2, 3$, that represent four different directions, see Figure 5. For a fixed $k = (k_1, k_2)^T \in J_2$, we recall that for each $\nu \in \{0, 1, 2, 3\}$, the generating functions in $\Psi_{-1}^{\nu}(\cdot - 2k)$ form an orthogonal basis of $V_{-1}^{\nu} \oplus W_{-1}^{\nu}$ (restricted to $Q_k := [2k_1, 2k_1 + 2) \times [2k_2, 2k_2 + 2)$).

We consider the 64 coefficients obtained altogether in the four projections, and select the smallest coefficients first. Those functions in our frame being connected with the smallest coefficients are not so "important" for the representation of f and will be pushed to be zero by using a suitable vector g_k.

The complete algorithm is described as follows. It needs to be applied for all $k \in J_2$.

Algorithm 2 (Sparse local representation in $V_{-1} + W_{-1}$)

Input: $D_{0,k} \in \mathbb{R}^{64}$ as in Algorithm 1b, such that $f_0|_{Q_k} = D_{0,k}^T \Phi_0(\cdot - 2k)$, and $R := (I_4 \otimes A) U^T \in \mathbb{R}^{64 \times 20}$.

1. Compute the local orthogonal projections of $f_0|_{Q_k} = D_{0,k}^T \Phi_0(\cdot - 2k)$ into the subspaces $V_{-1}^{\nu} \oplus W_{-1}^{\nu}$ ($\nu = 0, 1, 2, 3$).
   Start with the ansatz $f_{-1}^{\nu} := (h_k^{\nu})^T \Psi_{-1}^{\nu}(\cdot - 2k)$ in Q_k for the local orthogonal projection of f_0 into $V_{-1}^{\nu} \oplus W_{-1}^{\nu}$. The orthogonality of basis functions in Ψ_{-1}^{ν} implies

$$h_k^{\nu} := 4 \langle \Psi_{-1}^{\nu}(\cdot - 2k), f_0 \rangle \in \mathbb{R}^{16}.$$

   Introducing the vector $h_k := ((h_k^0)^T, (h_k^1)^T, (h_k^2)^T, (h_k^3)^T)^T \in \mathbb{R}^{64}$, we obtain

$$\begin{aligned} h_k &= 4 \langle \Psi_{-1}(\cdot - 2k), f_0 \rangle = 4 \langle (I_4 \otimes A) \Phi_0(\cdot - 2k), D_{0,k}^T \Phi_0(\cdot - 2k) \rangle \\ &= 4 (I_4 \otimes A) \langle \Phi_0, \Phi_0 \rangle D_{0,k}. \end{aligned}$$

   Let again P_1 be the permutation matrix for reordering of generating functions in V_0,

$$(\Phi^T, \Phi(\cdot - (\begin{smallmatrix}1\\0\end{smallmatrix}))^T, \Phi(\cdot - (\begin{smallmatrix}0\\1\end{smallmatrix}))^T, \Phi(\cdot - (\begin{smallmatrix}1\\1\end{smallmatrix}))^T) P_1 = \Phi_0^T,$$

   then with the Gram matrix G of Φ given in (2.4) we obtain $\langle \Phi_0, \Phi_0 \rangle = P_1^T (I_4 \otimes G) P_1$. Thus

$$h_k = 4 (I_4 \otimes A) P_1^T (I_4 \otimes G) P_1 D_{0,k}.$$

   The vector h_k now contains all local coefficients of the four orthogonal projections of f_0 into $V_{-1}^{\nu} \oplus W_{-1}^{\nu}$ ($\nu = 0, 1, 2, 3$).

2. Arrange the components of $h_k \in \mathbb{R}^{64}$ from lowest absolute value to highest value and compute the corresponding permutation $(p_1, \ldots, p_{64})$ of indices $(1, \ldots, 64)$. If some values in h_k have the same absolute value then take that with the smallest index first.

3. Compute an invertible matrix $C_k \in \mathbb{R}^{20 \times 20}$ by choosing 20 rows of R as follows.
   (a) The first row of C_k is the p_1-th row of R.
   (b) The second row of C_k is the p_2-th row of R, if it is linearly independent from the p_1-th row of R. Otherwise, consider the p_3-th row of R etc. In general, proceed as follows for $i = 1, 2, \ldots$: if the p_i-th row of R is linearly independent from the rows being already chosen in C_k then take this row as a further row of C_k. Otherwise, go further to the p_{i+1}-th one.

18

(c) Stop this procedure if 20 linear independent rows of R are found and C_k is completely determined. Since $rank(R) = 20$ the procedure comes to an end.

4. Let $q = (q_1, \ldots, q_{20})$ be the vector of indices of rows from R taken in C_k. Solve the linear system

$$C_k \, g_k = -(D_{-1,k})_q,$$

where $(D_{-1,k})_q$ contains the components with indices $q_1, \ldots, q_{20}$ of $D_{-1,k} = (I_4 \otimes A)D_{0,k}$ in this order. With the resulting vector g_k we determine the desired sparse coefficient vector $\tilde{D}^T_{-1,k} = D^T_{-1,k} + g_k^T R^T$ and find the new sparse local representation $\tilde{D}^T_{-1,k}\Psi_{-1}$ of f in $V_{-1} + W_{-1}$.

Step 4 of Algorithm 2 implies that in the new representation of $f_0 \in V_0$ given by $\tilde{D}^T_{-1,k}$ at least 20 wavelet coefficients vanish, namely those corresponding to the indices $(q_1, \ldots, q_{20})$. Finally we show two important properties of the proposed algorithm.

Lemma 5.1. *The sparse local representation of $f_0 \in V_0$ in $V_{-1} + W_{-1}$ obtained by Algorithm 2 is uniquely determined, i.e., it does not depend on the initial redundant representation of f_0 in V_0.*

Proof. Since the components of Ψ^ν_{-1} form a basis of $V^\nu_{-1} \oplus W^\nu_{-1}$ for each $\nu = 0, 1, 2, 3$, we observe that the local projections of f_0 into $V^\nu_{-1} \oplus W^\nu_{-1}$ are uniquely determined and do not depend on the initial representation of f_0 in V_0. Hence the matrix C_k computed in step 3 of Algorithm 2 is uniquely determined, too. Taking the parameter vector $g_k = -C_k^{-1}(D_{-1,k})_q$ as given in step 4 of Algorithm 2, the obtained new local representation of f_0 in $V_{-1} + W_{-1}$,

$$f_0 = \left[D^T_{-1,k} - (D_{-1,k})_q^T (C_k^{-1})^T R^T \right] \Psi_{-1},$$

contains by construction 20 zero coefficients corresponding to the indices $(q_1, \ldots, q_{20})$, i.e., the components $\psi_{q_1}, \ldots, \psi_{q_{20}}$ of the function vector $\Psi_{-1} = (\psi_\mu)^{64}_{\mu=1}$ are not longer involved in the representation of f_0. Let $f_0 = \hat{D}^T_{-1,k}\Psi_{-1}$ be a second representation of f_0 in $V_{-1} + W_{-1}$, then there exists a vector $b \in \mathbb{R}^{20}$ with $\hat{D}^T_{-1,k} = D^T_{-1,k} + b^T R^T$. Algorithm 2 provides now the representation

$$
\begin{aligned}
f_0 &= \left[\hat{D}^T_{-1,k} - (\hat{D}_{-1,k})_q^T (C_k^{-1})^T R^T \right] \Psi_{-1} \\
&= \left[D^T_{-1,k} + b^T R^T - [(D_{-1,k})_q^T + b^T R_q^T](C_k^{-1})^T R^T \right] \Psi_{-1} \\
&= \left[D^T_{-1,k} - (D_{-1,k})_q^T (C_k^{-1})^T R^T \right] \Psi_{-1},
\end{aligned}
$$

since $b^T R_q^T (C_k^{-1})^T R^T = b^T C_k^T (C_k^{-1})^T R^T = b^T R^T$ by construction. Here $b^T R_q^T \in \mathbb{R}^{20}$ denotes the subvector of $b^T R^T \in \mathbb{R}^{64}$ with components indexed by $(q_1, \ldots, q_{20})$. $\square$

We can show that the procedure in Algorithm 2 provides optimal results if the function f is locally constant.

Lemma 5.2. *Let f be constant on the square $Q_k = [2k_1, 2k_1 + 2) \times [2k_2, 2k_2 + 2)$ for some $k = (k_1, k_2)^T \in J_2$. Then Algorithm 2 provides an optimal representation of f in $V_{-1} + W_{-1}$.*

Proof. If $f \equiv c$ on Q_k with some constant $c \in \mathbb{R}$, it can be represented in V_0 by

$$f = c \cdot (1_{16}^T, 0_{16}^T, 0_{16}^T, 0_{16}^T) \, \Phi_0.$$

According to Lemma 5.1 we can reduce our considerations to this representation of f. We apply Algorithm 2 and show that the resulting coefficient vector is a sparsest possible one. Computing the orthogonal projections of f into $V_{-1}^\nu \oplus W_{-1}^\nu$, $\nu = 0, 1, 2, 3$, as in step 1 of Algorithm 2, the vector $h_k = c\,(w^T, w^T, w^T, w^T)$ with $w^T \in \mathbb{R}^{16}$ from (5.3) is obtained. Applying step 2 of Algorithm 2, the vector h_k yields the permutation

$$p = (2, 3, 4, 6, 7, 8, 10, 11, 12, \ldots, 62, 63, 64, 1, 5, 9, \ldots, 57, 61).$$

Without loss of generality let the rows of U be determined by the dependence relations given in Subsection 5.1 in the order as mentioned there. By a simple computation according to step 3 of Algorithm 2 it can be observed that the 48 rows with indices $(2, 3, 4, \ldots, 62, 63, 64)$ of $R = (I_4 \otimes A)U^T$ contain exactly 15 linearly independent rows that will be used to determine the first 15 rows of C_k. The last 5 rows of C_k are determined by the indices $(1, 9, 17, 33, 41)$ of R.

Now, we a apply the next step of Algorithm 2, step 4, and we obtain with

$$D_{-1,k} = c\,(I_4 \otimes A)(1_{16}^T, 0_{16}^T, 0_{16}^T, 0_{16}^T) = c\,(w^T, 0_{16}^T, 0_{16}^T, 0_{16}^T)$$

the vector $(D_{-1,k})_q = 2c(e_{16} + e_{17})$, since only the indices $q_{16} = 1$ and $q_{17} = 9$ determining the 16th and the 17th row of C_k correspond to nonzero values in $D_{-1,k}$. Here $e_\mu := (\delta_{\mu,\nu})_{\nu=1}^{20}$ denote the unit vectors of length 20. Using

$$w_1^T := (2, 0, 0, 0, 2, 0, 0, 0, 0_8^T) \in \mathbb{R}^{16}, \qquad w_2^T := w^T - w_1^T,$$

the 16th and 17th row of the matrix $(R\,C_k^{-1})^T$ have the form $\frac{1}{2}(w_1^T, -w_2^T, 0_{16}^T, 0_{16}^T)$ and $\frac{1}{2}(w_2^T, w_2^T, 0_{16}^T, -w^T)$ and we obtain

$$
\begin{aligned}
f &= \left[D_{-1,k}^T - (D_{-1,k})_q^T (C_k^{-1})^T R^T \right] \Psi_{-1} \\
&= c \cdot \left[(w^T, 0_{16}^T, 0_{16}^T, 0_{16}^T) - (w_1^T, -w_2^T, 0_{16}^T, 0_{16}^T) - (w_2^T, w_2^T, 0_{16}^T, -w^T) \right] \Psi_{-1} \\
&= c \cdot (0_{16}^T, 0_{16}^T, 0_{16}^T, w^T) \Psi_{-1}.
\end{aligned}
$$

This is an optimal sparse representation of f. $\qquad\square$

6. Numerical results

In this section, we use the described algorithms for image denoising and image approximation. Both applications are based on the efficient multiscale decomposition using the proposed directional Haar wavelet filter bank and a suitable wavelet shrinkage.

6.1. Image denoising

We consider a Gaussian noise with standard deviation $\sigma = 20$ that is added to the 256×256-synthetic image (Figure 6(a)) and to the pepper image of the same size (Figure 7(a)). We apply the directional Haar wavelet filter bank algorithm (Algorithm 1a) with a global hard-thresholding after a complete decomposition of the image. Choosing the shrinkage parameter $\lambda = \sigma\sqrt{\log(N)/2}$, where N denotes the number of pixels we obtain good denoising results (see Figure 6(b)) because the directional edges of the geometrical figures are well detected. Figures 6(c),(d) show that our method outperforms curvelets [5] as well as contourlets [11] for piecewise constant images. Even for natural images the denoising result is acceptable albeit not excellent, see Figure 7. However, it is of the same PSNR scale as with curvelets and it outperforms contourlets again. For the computation of the curvelet and the contourlet transform we have used the toolboxes from www.curvelet.org and www.ifp.uiuc.edu/~minhdo/software/.
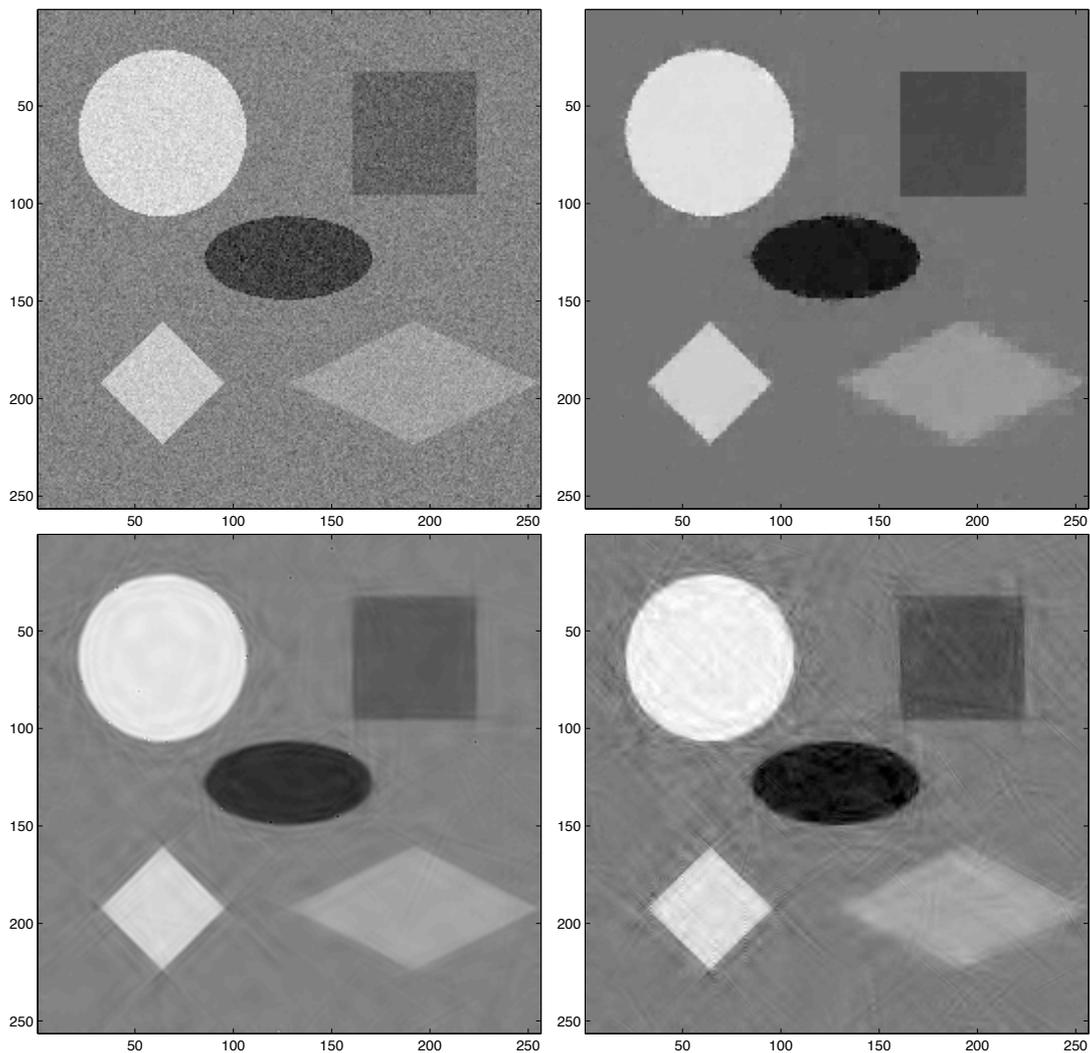
Figure 6: Piecewise constant image denoising. (a) noisy image, PSNR 22.08, (b) denoised image by our method, PSNR 33.12, (c) denoised image by curvelets, PSNR 32.18, (d) denoised image by contourlets, PSNR 29.91.

6.2. Image approximation

While for image denoising redundancy information is helpful, for image approximation it is not. Therefore, in order to get a sparse image representation we reduce the redundancies by applying Algorithm 1b. By doing this, the redundancy decreases from 4 to 2.74, i.e., for the pepper image of size 256×256 we get $179,780$ nonzero coefficients after decomposition. The same scale of redundancy occurs with the discrete curvelet transform. There we get 185,344 curvelet coefficients, what corresponds with a redundancy of 2.83. In comparison, the contourlet transform achieves a remarkable low redundancy of 1.3, i.e., the decomposition leads to $86,016$ coefficients. Now, in Figures 8 and 9 we keep the 1311 (resp. 6554) largest coefficients and reconstruct the images.
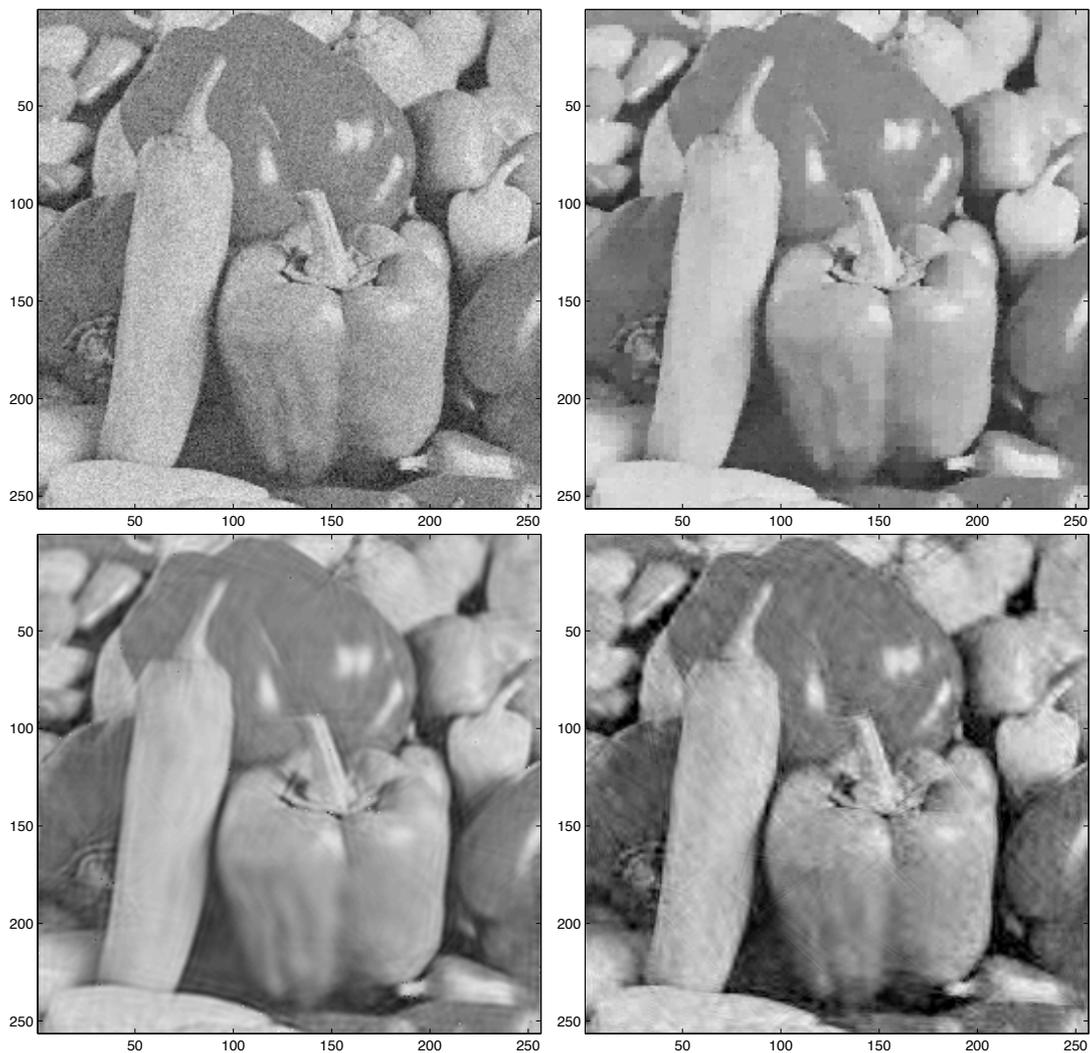
Figure 7: 'Pepper' image denoising. (a) noisy image, PSNR 22.08, (b) denoised image by our method, PSNR 28.00, (c) denoised image by curvelets, PSNR 28.14, (d) denoised image by contourlets, PSNR 26.16.

7. Conclusions

Certain drawbacks of the existing non-adaptive directional wavelet constructions, like curvelets, are the global support of curvelet elements in time domain and the missing MRA structure leading to rather complex algorithms for the digital curvelet transform. Therefore we desire to construct wavelet frame functions with small support in time domain, simple structure, based on a multiresolution analysis and leading to efficient filter bank algorithms, high directionality, low redundancy, and suitable smoothness.

The proposed directional Haar wavelet frames on triangles can be seen as a first step in this direction. However, our approach has some limitations. First, one would like to have continuous (or smooth) frame functions instead of piecewise constants. Unfortunately, the usage of box splines on a multi-directional mesh seems not to be advantageous due to their fast increasing support. This results in large corresponding filters and filter bank algorithms with high complexity. Therefore,
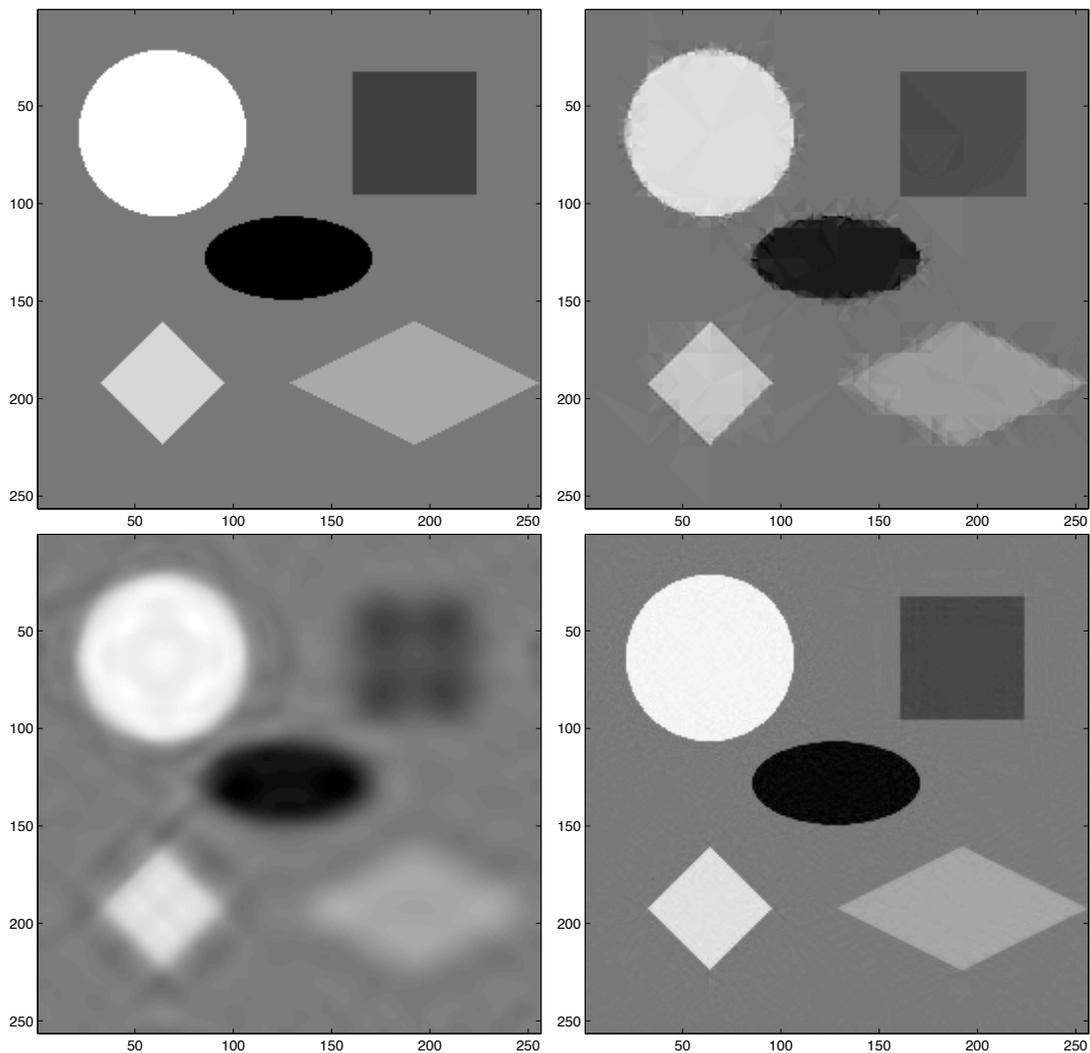
22

Figure 8: Piecewise constant image approximation with 2 % coefficients. (a) original image, (b) approximation by our method, PSNR 31.82, (c) approximation by curvelets, PSNR 26.77, (d) approximation by contourlets, PSNR 34.13.

we are currently investigating a multiwavelet approach using piecewise linear wavelet functions with small support on directional triangles.

One may also ask for an extension of the proposed Haar wavelet filter bank to more directions using thinner support triangles. But a further splitting of the considered triangles leads to the problem that vertices of these triangles may not lie in the set $2^{-j}\mathbb{Z}^2$. One way out of this limitation of directionality may be the choice of different dilation matrices.

An important question is, how to relate high directionality with small redundancy of the wavelet dictionary. This problem may be only solvable by a locally adaptive choice of directional frame functions. In particular, for the proposed Haar wavelet dictionary one may think about a local optimization procedure in order to activate only frame functions that correspond to certain locally important directions. This topic is also subject of further research.
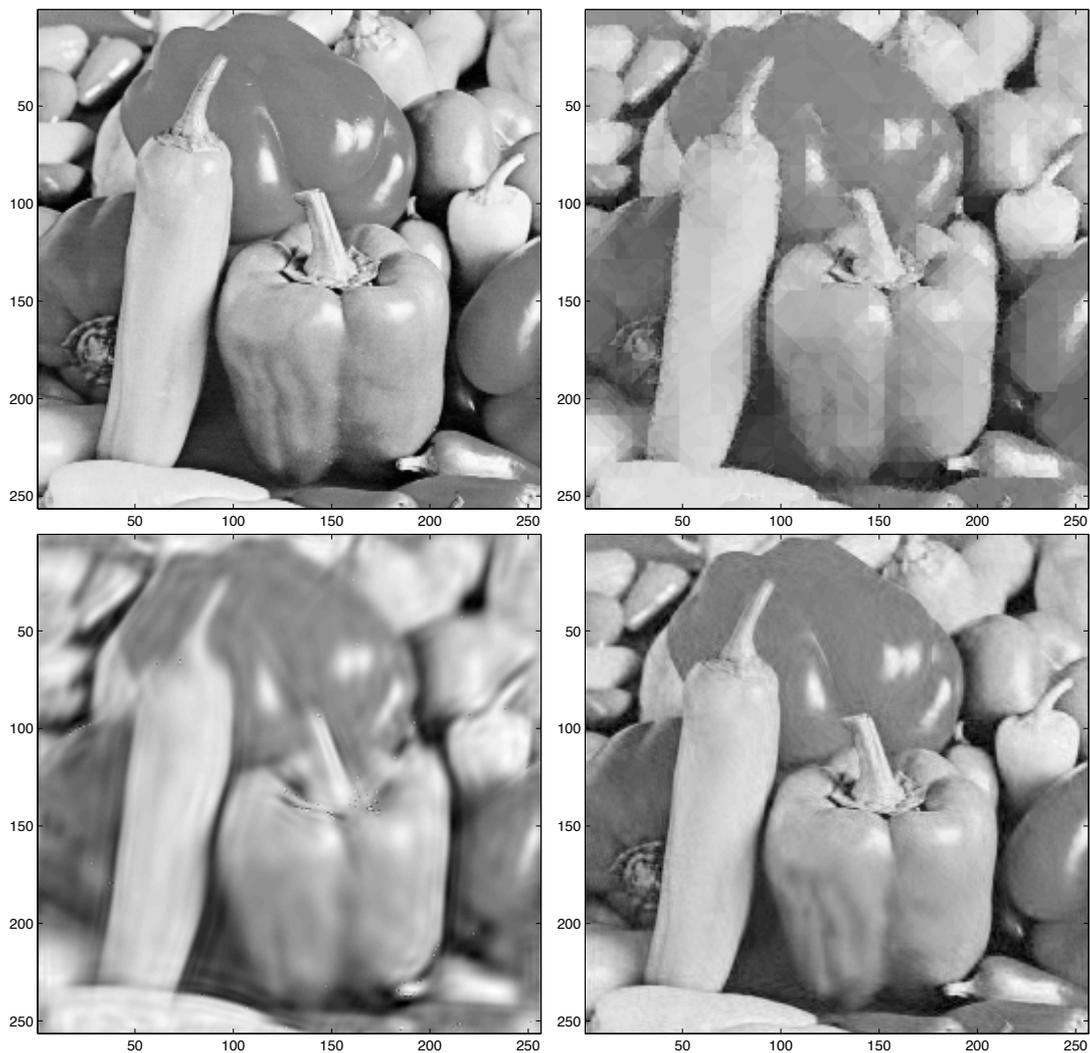
Figure 9: 'Pepper' image approximation with 10 % coefficients. (a) original image, (b) approximation by our method, PSNR 27.80, (c) approximation by curvelets, PSNR 25.51, (d) approximation by contourlets, PSNR 32.12.

References

[1] F. Arandiga, J. Baccou, M. Doblas, & J. Liandrat, Image compression based on a multi-directional map-dependent algorithm, Appl. Comput. Harmon. Anal. **23** (2007), 181–197.

[2] F. Arandiga, A. Cohen, R. Donat, N. Dyn, & B. Matei, Approximation of piecewise smooth functions and images by edge-adapted (ENO-EA) nonlinear multiresolution techniques, Appl. Comput. Harmon. Anal. **24** (2008), 225–250.

[3] C. de Boor, R.A. DeVore, & A. Ron, On the construction of multivariate (pre)wavelets, Constr. Approx. **9** (1993), 123–166.

24

[4] E.J. Candès & D.L. Donoho, Curvelets – a surprisingly effective nonadaptive representation for objects with edges, in: Curves and Surfaces, C. Rabut, A. Cohen and L.L. Schumaker (Eds.), Vanderbilt University Press, Nashville, 2000, 105–120.

[5] E.J. Candès & D.L. Donoho, New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities, Comm. Pure and Appl. Math. **56** (2004), 216–266.

[6] S.S. Chen, Basis Pursuit, PhD thesis, Stanford University, 1995.

[7] A. Cohen & B. Matei, Compact representation of images by edge adapted multiscale transforms. in Proc. IEEE Int. Conf. on Image Proc. (ICIP), 2001, Thessaloniki, Greece, 8–11.

[8] R.R. Coifman & F.G. Meyer, Brushlets: a tool for directional image analysis and image compression, Appl. Comput. Harmon. Anal. **5** (1997), 147–187.

[9] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.

[10] L. Demaret, N. Dyn, & A. Iske, Image compression by linear splines over adaptive triangulations. Signal Processing **86** (2006), 1604–1616.

[11] M.N. Do & M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Process. **14**(12) (2005), 2091–2106.

[12] D.L. Donoho & M. Elad, Optimally sparse representation in general (non-orthogonal) dictionaries via l^1 minimization, Proceedings of the National Academy of Sciences of the USA, **100** (2003), 2197–2202.

[13] K. Gröchenig & W.R. Madych, Multiresolution Analysis, Haar bases, and self-similar tilings of $\mathbb{R}^n$, IEEE Trans. Inform. Theory **38** (1992), 556–568.

[14] K. Guo & D. Labate, Optimally sparse multidimensional representation using shearlets, SIAM J. Math. Anal. **39** (2007), 298–318.

[15] K. Guo, W.-Q. Lim, D. Labate, G. Weiss, & E. Wilson, Wavelets with composite dilations and their MRA properties, Appl. Comput. Harmon. Anal. **20** (2006), 231–249.

[16] Y. Hur & A. Ron, New constructions of piecewise-constant wavelets, Electronic Trans. Numerical Analysis **25** (2006), 138–157.

[17] N. Kingsbury, Complex wavelets with shift invariant analysis and filtering of signals, Appl. Comput. Harmon. Anal. **10** (2001), 234–253.

[18] I. Krishtal, B. Robinson, G. Weiss, & E. Wilson, Some simple Haar-type wavelets in higher dimensions, J. Geom. Anal. **17** (2007), 87–96.

[19] J. Krommweh, Tight frame characterization of multiwavelet vector functions in terms of the polyphase matrix, Int. J. Wavelets Multiresolut. Inf. Process., **7**(1) (2009) 9–21.

[20] B. Matei, Smoothness characterization and stability in nonlinear multiscale framework: theoretical results, Asymptotic Anal. **41** (2005), 277–309.

[21] E. Le Pennec & S. Mallat, Sparse geometric image representations with bandelets, IEEE Trans. Image Process. **14** (2005), 423–438.

[22] W.-Q. Lim, Wavelets with Composite Dilations, PhD thesis, Washington University, St. Louis, MO, 2006.

[23] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, San Diego, CA, 1998.

[24] D.D. Po & M.N. Do, Directional multiscale modeling of images using the contourlet transform, IEEE Trans. Image Process. **15** (2006), 1610–1620.

[25] D. Rosça, Piecewise constant wavelets on triangulations obtained by 1-3 splitting, Int. J. Wavelets Multiresolut. Inf. Process. **6**(2) (2008), 209–222.

[26] J.A. Tropp, Greed is good: algorithmic results for sparse approximation, IEEE Trans. Information Theory **50**(10) (2004), 2231–2242.

[27] V. Velisavljević, B. Beferull-Lozano, M. Vetterli, & P.L. Dragotti, Directionlets: Anisotropic multidirectional representation with separable filtering, IEEE Trans. Image Process. **15**(7) (2006), 1916–1933.