Nonseparable Wavelet Packets

Lasse Borup and Morten Nielsen

Abstract. We introduce wavelet packets in the setting of a multiresolution analysis of $L^2(\mathbb{R}^d)$ generated by an arbitrary dilation matrix A satisfying $|\det A| = 2$, and note that the best-basis algorithm of wavelet packets in the standard one dimensional case can be generalized to this multidimensional setting. We then consider the application of the best basis algorithm in the two-dimensional case for a specific dilation matrix to image compression. The performance of the algorithm is compared to the separable wavelet algorithm for twelve test images.

§1. Introduction

Wavelet analysis was originally introduced in order to improve seismic signal processing by switching from short-time Fourier analysis to new algorithms better suited to detect and analyze abrupt changes in signals. It corresponds to a decomposition of phase space in which the trade-off between time and frequency localization has been chosen to provide better and better time localization at high frequencies in return for poor frequency localization. This makes the analysis well adapted to the study of transient phenomena and has proven a very successful approach to many problems in signal processing, numerical analysis, and quantum mechanics. Nevertheless, for stationary signals wavelet analysis is outperformed by short-time Fourier analysis. Wavelet packets were introduced by R. Coifman, Y. Meyer, and M. V. Wickerhauser [2] to improve the poor frequency localization of wavelet bases at high frequencies and thereby provide a more efficient decomposition of signals containing both transient and stationary components.

So far most work on wavelet packets has been done in one dimension or using separable wavelet packets in higher dimensions (i.e. tensor products of one dimensional wavelet packets). However, separable wavelet and wavelet packet bases both have several drawbacks for the application to fields like image analysis since they impose an unavoidable line structure on the plane. For example, the zero set of a separable wavelet packet at high frequencies will contain a large number (same order of magnitude as the frequency) of

Approximation Theory X
Charles K. Chui, Larry L. Schumaker, and Joachim Stoeckler (eds.), pp. 1–10.

ISBN 0-8265-xxxx-x.

horizontal and vertical lines that may create artifacts in the reconstructed image. Another potential problem is in the Fourier domain where separable two-dimensional wavelet packets have four characteristic peaks making it hard to selectively localize a unique frequency. R. Coifman and F. Meyer introduced the so-called Brushlets in [3] to remove the "uncertainty" in frequency localization, however the Brushlets are essentially Fourier transforms of smooth local trigonometric bases and are therefore no longer functions associated with a multiresolution structure.

In the present paper we will introduce wavelet packets on $\mathbb{R}^d$ associated with a multiresolution analysis generated by a general dilation matrix A with determinant ± 2. The (nonseparable) wavelets associated with such structures have been investigated in several recent papers [1] and they have also been applied to image analysis. As in the one dimensional case, wavelet packet analysis extends the wavelet expansion of a signal to a whole library of orthonormal expansions providing a more adaptable scheme to effectively represent the signal.

The structure of the paper is as follows. In Section 2 we introduce wavelet packets associated with the class of multiresolution analyses of $\mathbb{R}^d$ for which there are associated wavelet bases generated by only one wavelet. Such functions provide the same large number of orthonormal bases as wavelet packets in one-dimension do and should provide a good platform for doing image analysis using the well known best basis algorithm of Coifman and Wickerhauser. Section 3 contains the numerical results we obtained from applying the algorithm, in the two-dimensional case with a specific dilation matrix, to image compression. We use twelve "standard" test images to compare the effectiveness of the algorithm to the associated nonseparable wavelet decomposition at different compression rates.

§2. Non-stationary Wavelet Packets For A General Dilation Matrix

We begin by recalling some facts about multiresolution analyses associated with a general dilation matrix that we will use later in this section to define the wavelet packets we have in mind. The reader can find a more extensive discussion of the topic in [4].

Let A be a $(d \times d)$-matrix such that $A : \mathbb{Z}^d \to \mathbb{Z}^d$. If the eigenvalues of A all have absolute value strictly greater than 1 then we call A a dilation matrix. We can define a multiresolution analysis associated with such A:

Definition 1. *A multiresolution associated with a dilation matrix A is a sequence of closed subspaces $(V_j)_{j\in\mathbb{Z}}$ of $L^2(\mathbb{R}^d)$ satisfying*

(i) $V_j \subset V_{j+1}, \quad \forall j \in \mathbb{Z}$,
(ii) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R}^d)$ *and* $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$,
(iii) $f \in V_j \Leftrightarrow f(A\cdot) \in V_{j+1}, \quad \forall j \in \mathbb{Z}$,
(iv) *there exists a function $\phi \in V_0$ called a scaling function such that the system $\{\phi(\cdot - \gamma)\}_{\gamma\in\mathbb{Z}^d}$ is an orthonormal basis for V_0.*

The wavelet spaces W_j associated with such a multiresolution analysis are given by $W_j = V_{j+1} \cap V_j^{\perp}$, and one can easily check that $f \in W_j \Leftrightarrow$

$f(A\cdot) \in W_{j+1}$ and $L^2(\mathbb{R}^d) = \bigoplus_{j\in\mathbb{Z}} W_j$. A family of wavelets associated with the multiresolution analysis is a collection of s functions $\{\Psi^r\}_{r=1}^s$ for which $\{\Psi^r(\cdot-\gamma) : r = 1, \dots, s, \gamma \in \mathbb{Z}^d\}$ is an orthonormal basis for W_0. Suppose $|\det A| = q$. It turns out that the number of wavelets needed to generate such a basis for W_0 is exactly $q-1$. This makes the case $|\det A| = 2$ especially interesting since the wavelet basis is generated by only one function just as in the one-dimensional case.

Let $\{V_j\}_{j\in\mathbb{Z}}$ be a multiresolution analysis of $L^2(\mathbb{R}^d)$ associated with a dilation matrix A satisfying $|\det A| = 2$. Suppose (Φ, Ψ) is an associated scaling function/wavelet pair. Then there exist $2\pi\mathbb{Z}^d$-periodic functions m_0 and m_1 such that

$$\hat{\Phi}(\xi) = m_0(D\xi)\hat{\Phi}(D\xi)$$
$$\hat{\Psi}(\xi) = m_1(D\xi)\hat{\Phi}(D\xi),$$

with $D = (A^*)^{-1}$. Since $|\det A| = 2$ we can find $\Gamma \in \mathbb{Z}^d$ such that $\Gamma + \mathbb{Z}^d/A^*\mathbb{Z}^d$ satisfies $\mathbb{Z}^d = A^*\mathbb{Z}^d \cup (\Gamma + A^*\mathbb{Z}^d)$. Then it is easy to check that the matrix

$$\begin{bmatrix} m_0(\xi) & m_0(\xi + 2\pi D\Gamma) \\ m_1(\xi) & m_1(\xi + 2\pi D\Gamma) \end{bmatrix}$$

is unitary for a.a. $\xi \in \mathbb{R}^d$. This observation leads to the following definition. We let A and Γ be related as above.

Definition 2. *Let m_0 and m_1 be $2\pi\mathbb{Z}^d$ periodic functions for which*

$$\begin{bmatrix} m_0(\xi) & m_0(\xi + 2\pi D\Gamma) \\ m_1(\xi) & m_1(\xi + 2\pi D\Gamma) \end{bmatrix}$$

is unitary a.e. Then we call (m_0, m_1) a pair of orthogonal quadrature filters associated with (A, Γ).

We can now define the natural generalization of wavelet packets to the setting of a multiresolution analysis associated with a dilation matrix A with $|\det A| = 2$.

Definition 3. *Let $\{(m_0^{(p)}, m_1^{(p)})\}_{p=1}^\infty$ be a sequence of orthogonal quadrature filters associated with (A, Γ). We define the basic nonstationary wavelet packets $\{w_n\}_{n=0}^\infty$ by $w_0 = \Phi$, $w_1 = \Psi$, and for $2^k \le n < 2^{k+1}$ with binary expansion $n = \sum_{j=1}^{k+1} \epsilon_j 2^{j-1}$, we let*

$$\hat{w}_n(\xi) = \left[\prod_{j=1}^{k+1} m_{\epsilon_j}^{(k-j+2)}(D^j\xi)\right]\hat{\Phi}(D^{k+1}\xi).$$

Let us state two most important facts about the wavelet packets from the above definition. The two theorems below show how to extract orthonormal bases from the wavelet packet construction above, and thus gives us some new tools to signal and image processing. We have included the proofs for convenience. However, the reader should notice that everything works exactly as in the one-dimensional case, only the multiresolution structure matters.

Theorem 4. *The basic wavelet packets*

$$\{w_n(\cdot - k) | 0 \le n < 2^j, k \in \mathbb{Z}^d\}$$

form a basis for V_j. *Furthermore,*

$$\{w_n(\cdot - k) | n \in \mathbb{N}_0, k \in \mathbb{Z}^d\}$$

form an orthonormal basis for $L^2(\mathbb{R}^d)$.

Proof: Let $\Omega_n = \overline{Span}\{w_n(\cdot - k)\}_{k \in \mathbb{Z}^d}$, and define $\delta f(x) = \sqrt{2} f(Ax)$. Using the QMF-condition it is not hard to verify that $\delta\Omega_n = \Omega_{2n} \oplus \Omega_{2n+1}$ (see eg. [4] p. 112). Thus,

$$\begin{aligned} \delta\Omega_0 \ominus \Omega_0 &= \Omega_1 \\ \delta^2\Omega_0 \ominus \delta\Omega_0 &= \delta\Omega_1 = \Omega_2 \oplus \Omega_3 \\ \delta^3\Omega_0 \ominus \delta^2\Omega_0 &= \delta\Omega_2 \oplus \delta\Omega_3 = \Omega_4 \oplus \Omega_5 \oplus \Omega_6 \oplus \Omega_7 \\ &\vdots \\ \delta^k\Omega_0 \ominus \delta^{k-1}\Omega_0 &= \Omega_{2^{k-1}} \oplus \Omega_{2^{k-1}+1} \oplus \cdots \oplus \Omega_{2^k-1}. \end{aligned}$$

By telescoping the above equalities we finally get the wanted result

$$\delta^k\Omega_0 \equiv \delta^k V_0 = V_k = \Omega_0 \oplus \Omega_1 \oplus \cdots \oplus \Omega_{2^k-1},$$

and $\cup_{k \ge 0} V_k$ is dense in $L^2(\mathbb{R}^d)$ by the definition of a multiresolution analysis. □

The above theorem can be generalized considerably. The following construction gives us a whole library of orthonormal bases each with different time-frequency properties. The collection of bases will be used for the best basis algorithm.

Theorem 5. *Let* $\{w_n\}$ *be a family of non-stationary wavelet packets associated with the dilation matrix* A. *For every partition* P *of* $\mathbb{N}_0$ *into sets of the form* $I_{nj} = \{n2^j, \ldots, (n+1)2^j - 1\}$ *with* $n, j \in \mathbb{N}_0$, *the family*

$$\{2^{j/2} w_n(A^j \cdot -k)\}_{k \in \mathbb{Z}^d, I_{nj} \in P}$$

is an orthonormal basis for $L^2(\mathbb{R}^d)$.

Proof: An argument similar to the one in Theorem 4 shows that

$$\delta^k\Omega_n = \Omega_{2^k n} \oplus \Omega_{2^k n+1} \oplus \cdots \oplus \Omega_{2^k(n+1)-1}.$$

Moreover, the functions $\{2^{j/2} w_n(A^j \cdot -q)\}_{q \in \mathbb{Z}^d}$ span the space $\delta^j\Omega_n$ and

$$\sum_{I_{nj} \in P} \delta^j \Omega_n = \bigoplus_{q \ge 0} \Omega_q = L^2(\mathbb{R}^d),$$

which proves the theorem. □

2.1. The Best Basis Algorithm

One of the main advantages of wavelet packets in the nonseparable setting, described above, is that one can use the implementation of the best-basis algorithm of Coifman and Wickerhauser for the one dimensional wavelet packet case without any modification whatsoever. We will not describe the algorithm here but instead refer the reader to [2].

2.2. A Family of Nonseparable Wavelets

The nonseparable wavelets we are going to use in the following section are examples of the family of wavelets constructed by E. Belogay and Y. Wang in [1]. Using the dilation matrix

$$A = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix},$$

they constructed a family of nonseparable wavelets with arbitrary smoothness. We denote the corresponding family of coefficient masks of size $4(r-1) \times 2$, $r > 0$ by $\mathcal{C}_r$.

§3. Implementation

In this section we are going to test the performance of the nonseparable wavelet packets constructed using Belogay and Wang's family of wavelets. We will compare the performance with a wavelet transform using the same wavelets and with a standard separable wavelet packets construction using Daubechies univariate wavelets. The tests are made on the following 12 standard images

image	pixel size	image	pixel size
Lena	512×512	Plane	512×512
Barbera	512×512	Milkdrop	512×512
Goldhill	512×512	Kilauea	416×512
Baboon	512×512	Boats	512×512
Bridge	512×512	France	480×672
Peppers	512×512	Library	352×448

All the images are given with 8-bit gray-scale values. In Section 3.1 we describe in details the test on the image "Lena", followed by the results from the tests on the other images. In the following "Daub#" means Daubechies 1D filter of length #.

We have implemented the nonseperable best basis algorithm in matlab, using the same construction as in the wavelet toolbox "Uvi_wave" (see http://www.tsc.uvigo.es/~wavelets/uvi_wave.html). Since we have just one scalingfunction and mother wavelet in our nonseparable construction, only minor modification of the one dimensional best basis algorithm is needed. To get perfect reconstruction of a transformed image, we have chosen to use circular transformations, i.e. we represent an image as a periodic function. When we make a best basis search we only allow transformations down to a certain level k_{max}. If, say a particular image is of size $M \times N$ and $m \times n$ is the size of a given filter, then k_{max} must satisfy $M2^{-k_{max}} > m$ and $N2^{-k_{max}} > n$. This restriction only ensures that it is enough to copy the image once at the boundary. As cost function in the best basis algorithm we use Shannon's entropy function (other cost functions such as the ℓ_p-norm, $p \leq 1$, have been used as well, but the overall picture was the same).

3.1. Test Results

To introduce the method used to test the nonseparable wavelet packets, we begin by describing a comparison of $\mathcal{C}_5$ and Daub20 when used on the image "Lena". On this image we do a best basis search, using $\mathcal{C}_5$ and Daub20, followed by a hard threshold keeping only a certain percent of the coefficients. Then we perform the inverse transformation using only the remaining coefficients. Figure 1 (left) displays the signal to noise ratio (SNR) for different threshold values, using the best basis wavelet packets transformation (best-WPT) with $\mathcal{C}_5$ and Daub20 respectively . The SNR is given by

$$\mathrm{SNR} := 10 \log_{10} \left(\frac{\|x^{exact} - x^{thres}\|_{L^2}^2}{\|x^{exact}\|_{L^2}^2} \right),$$

where x^{exact} is the original image and x^{thres} is the reconstructed image after threshold.

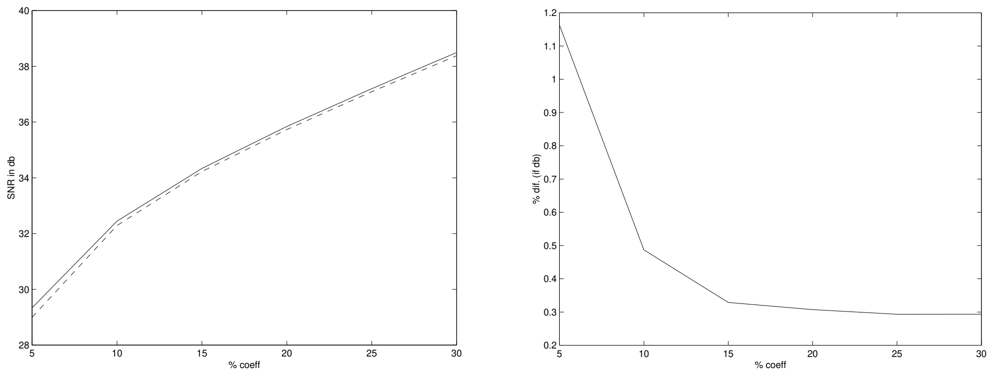

Fig. 1. Left: Error due to threshold using Best-WPT with $\mathcal{C}_5$ (solid) and Daub20 (dashed). Right: Relative improvement using $\mathcal{C}_5$ instead of Daub20.

Figure 1 (right) shows the relative improvement using $\mathcal{C}_5$ instead of Daub20, measured in SNR frame i.e.

$$100 \cdot \frac{\mathrm{SNR}^{\mathcal{C}_5} - \mathrm{SNR}^{\mathrm{Daub20}}}{\mathrm{SNR}^{\mathcal{C}_5}} \%.$$

As can be seen from the plot, the most improvement is given at the highest compression rates.

To see if there is any visual difference between the result of using $\mathcal{C}_5$ and Daub20 we plot the reconstructed image at compression rate 1:50 using the two methods (see figure 2). Notice that the left image seems less blurred than the right.

Fig. 2. Reconstruction of "Lena" with a compression rate of 1:50 using $\mathcal{C}_5$ (left) and Daub20 (right).

3.1.1. Remaining Results

We now display the test results corresponding to the other test images. Shannon's entropy function is still used as cost function. We first test the improvement of using best wavelet packets transform instead of using a normal wavelet transform (WT).

The following table gives the relative improvement using best-WPT with $\mathcal{C}_5$ instead of a WT still using $\mathcal{C}_5$. The top row shows the percentage of kept coefficients before reconstruction.

% coeff	1	2	5	10	15	20
Lena	2.146	2.209	2.058	1.400	0.990	0.770
Barbera	8.091	10.598	10.956	9.301	7.550	6.071
Goldhill	3.439	4.153	4.309	3.984	3.704	3.509
Baboon	2.093	3.029	4.418	4.931	4.823	4.502
Bridge	3.173	2.939	2.586	2.432	2.290	2.176
Peppers	5.099	4.167	2.034	1.114	0.896	0.732
Plane	3.502	3.978	2.642	1.968	1.548	1.269
Milkdrop	6.306	4.763	2.209	1.516	1.288	1.187
Kilauea	3.162	2.161	0.733	-0.156	-0.741	-1.208
Boats	3.484	3.872	4.326	4.073	3.904	3.637
France	4.163	4.874	3.320	0.795	0.618	0.233
Library	4.078	3.744	3.118	2.729	2.038	1.431

Notice that the best-WPT perform better on all twelve image at the highest compression rates, compared to the WT. Especially there is a big difference on "Barbera" possibly due to the textures. In figure 3 a 1:20 compressed version of "Barbera" is shown using a best-WPT (left) and a WT (right). Notice the different performances in reconstructing the pattern, say on the trousers.

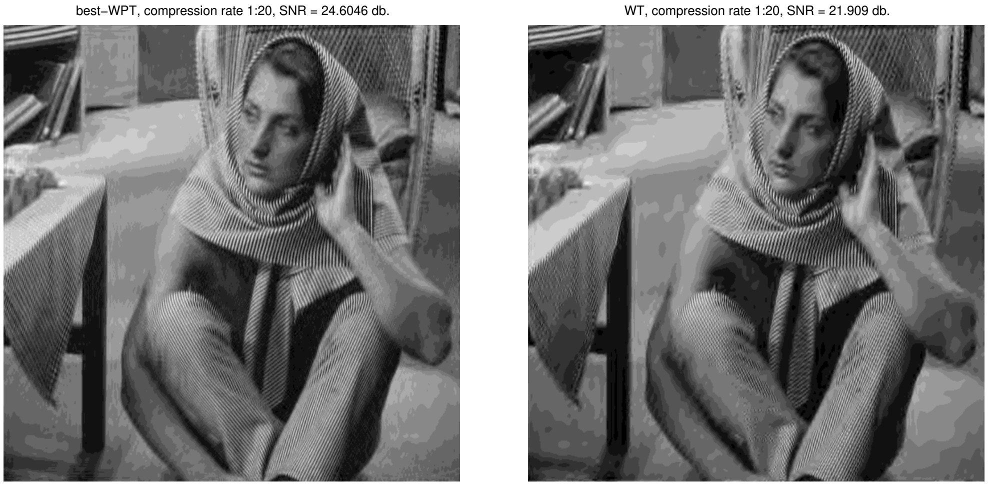

Fig. 3. The image "Barbera" after a 1:20 compression using Best-WPT (left) and WT (right) respectively both using $\mathcal{C}_5$.

As an apparent paradox the WT perform better on "Kilauea" at compression rate 1:10. This is because the algorithm choose the best basis relative to a specific cost function different from the SNR-measure.

We now return to the test earlier described in details on "Lena". The next table gives the relative improvements using best-WPT with $\mathcal{C}_5$ instead of Daub20.

% coeff	1	2	5	10	15	20
Lena	1.996	1.904	1.163	0.487	0.329	0.307
Barbera	1.004	0.774	0.332	-0.043	-0.307	-0.512
Goldhill	0.889	0.920	0.756	0.494	0.422	0.356
Baboon	0.596	0.743	0.619	0.382	0.266	0.153
Bridge	0.492	0.431	0.375	0.317	0.333	0.336
Peppers	4.311	2.777	1.173	0.690	0.545	0.477
Plane	2.009	2.174	1.725	1.209	0.900	0.737
Milkdrop	4.025	3.083	1.457	0.798	0.584	0.484
Kilauea	1.594	5.418	10.196	7.025	2.766	2.965
Boats	1.347	1.785	1.916	1.480	1.091	0.802
France	1.614	3.094	5.691	8.341	10.081	11.237
Library	1.596	2.049	3.204	3.853	3.948	3.987

Notice again that the overall best improvements are given at the highest compression rates. A particularly good performance is on the images "Kilauea" and "France". This might be because both of these images contains text. Figure 4 displays the result of a 3:50 compression of "Kilauea". Observe that the text seems sharper on the left image than on the right.

On the images "Peppers" and "Milkdrop" we also experience a relative large improvement when keeping only about 1% of the transformed coefficients. This might be due to the relative simple structure of both images. The difference is hardly visible on the compressed images though (see figure 5).

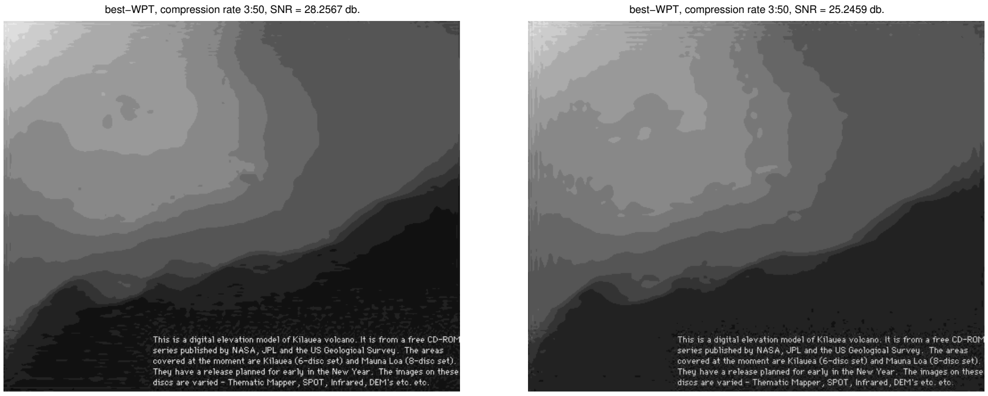

Fig. 4. The image "Kilauea" after a 3:50 compression using Best-WPT, $\mathcal{C}_5$ (left) and daub20 (right) respectively.

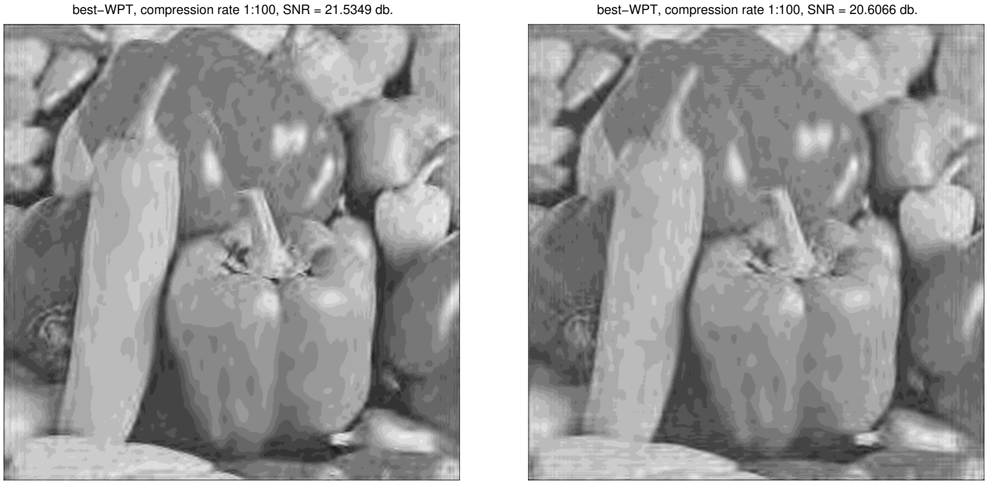

Fig. 5. The image "Peppers" after a 1:100 compression using Best-WPT, $\mathcal{C}_5$ (left) and daub20 (right) respectively.

§4. Conclusion

We have introduced wavelet packets in the setting of a multiresolution analysis associated with a dilation matrix with determinant 2. This setting allows us to use the well known wavelet packets analysis from the one dimensional case. The wavelet packets analysis associated with this setting is in general nonseparable, and thus different from what can be obtained by taking tensor product of one dimensional wavelet packets. The construction is especially interesting in two dimensions, since it makes it possible to perform image analysis using methods very similar to one dimensional methods. Due to this equivalence it is easier to index function expansions in this setting than in a separable one, thus simplifying implementation and image analysis.

We have tested the new algorithm on 12 standard images. The nonseparable best wavelet packets transform (best-WPT) has been compared to a separable best-WPT, to give an impression of the performance of the new algorithm. The choice of separable wavelets has solely been made such that the two algorithms have comparable complexity. We notice that the new algorithm shows some advantages when compressing images with a simple structure or images containing text. We have also tested the best-WPT against a wavelet transform (both in the nonseparable case). In almost all the tests the best-WPT gave very positive results. On the test image "Barbera" the relative improvement exceeded 10% at certain compression rates. In a few tests the WT performed better though. This is due to an inherent problem with the best basis algorithm; the algorithm finds the optimal choice with respect to the cost function and not the L^2-norm. Finally we should notice that any dilation matrix with determinant 2 and corresponding MRA would fit into this scheme.

References

1. Eugene Belogay and Yang Wang. Arbitrarily smooth orthogonal nonseparable wavelets in $\mathbb{R}^2$. SIAM J. Math. Anal. **30**(3):678–697 (electronic), 1999.
2. Ronald R. Coifman, Yves Meyer, Steven Quake, and M. Victor Wickerhauser. Signal processing and compression with wavelet packets. In *Progress In Wavelet Analysis And Applications (Toulouse, 1992)*, pages 77–93. Frontières, Gif, 1993.
3. François G. Meyer and Ronald R. Coifman. Brushlets: a tool for directional image analysis and image compression. Appl. Comput. Harmonic Anal. **4**(2):147–187, 1997.
4. P. Wojtaszczyk. *A Mathematical Introduction To Wavelets.* Cambridge University Press, Cambridge, 1997.

Morten Nielsen
Dept. of Mathematics
University of South Carolina
Columbia, SC 29208
nielsen@math.sc.edu

Lasse Borup
Dept. of Mathematical Sciences
Aalborg University
Aalborg, DK-9220
Denmark
lasse@math.auc.dk