Available at
www.ElsevierMathematics.com
POWERED BY SCIENCE DIRECT®
Journal of Symbolic Computation 37 (2004) 187–207

Journal of Symbolic Computation

www.elsevier.com/locate/jsc

Two-dimensional wreath product group-based image processing

Richard Foote[a], Gagan Mirchandani[b], Daniel Rockmore[c,*]

[a]*Department of Mathematics and Statistics, University of Vermont, Burlington VT 05405, USA*
[b]*Department of Electrical and Computer Engineering, University of VT, Burlington VT 05405, USA*
[c]*Department of Mathematics, Dartmouth College, Hanover NH 03755, USA*

Received 4 December 2000; accepted 11 June 2002

Abstract

A theoretical foundation to the notion of 2D transform and 2D signal processing is given, focusing on 2D group-based transforms, of which the 2D Haar and 2D Fourier transforms are particular instances. Conditions for separability of these transforms are established. The theory is applied to certain groups that are wreath products of cyclic groups to give separable and inseparable 2D wreath product transforms and their filter bank implementations.

Keywords: Image processing; 2D; Transform

1. Introduction

Harmonic analysis is at the heart of much of signal and image processing. This is mainly the harmonic analysis of Abelian groups, and ultimately, after sampling, and quantizing, finite Abelian groups. Nevertheless, this general group theoretic viewpoint has proved fruitful, yielding a natural group-based multiresolution framework as well as some new and potentially useful nonAbelian examples.

The papers (Foote et al., 2000; Mirchandani et al., 2000) lay out a general finite group-based approach to signal and image processing and pay special attention to the use of certain wreath product groups for image processing. The theory explicated there is one dimensional in the sense that the input signal f (respectively its Fourier transform) is represented as a column vector, and all analysis and synthesis is effected as a matrix–vector multiply: $\mathcal{F} \cdot f$ where $\mathcal{F}$ is the Fourier matrix, (respectively its inverse) for some specified finite group.

* Corresponding author.
E-mail address: rockmore@cs.dartmouth.edu (D. Rockmore).

0747-7171/$ - see front matter © 2003 Published by Elsevier Ltd.
doi:10.1016/j.jsc.2002.06.004

This paper is intended to complement and extend this earlier work in the direction of a two-dimensional (2D) finite group-based theory, again with the intent of applying this work to image processing. Historically, the notion of a 2D transform almost always entails a representation of the signal f as a 2D array so that the transform is given by a matrix multiplication of the form AfB^t for suitable matrices A and B. The most familiar example is that in which $A = B$ is the Fourier matrix and the associated 2D transform is the usual 2D discrete Fourier transform.

The new contributions of this paper are twofold. On the one hand we give a theoretical foundation to the notion of 2D transform and 2D signal processing which extends naturally to higher dimensions. We also provide a theory of 2D *group transforms*, of which the 2D Haar transform and 2D Fourier transform are particular instances. These well-known 2D transforms are also particular instances of a 2D *wreath product transform* (WPT), whose explication is the second major new contribution. This extends the earlier work (ibid.) which focused exclusively on the 1D WPT. We include here examples and numerical experiments and indicate its potential use and application for image processing. The presence of an underlying (non-Abelian) group afforded by this approach provides additional structure that has heretofore been largely unexplored. In particular, there is a group-based convolution which gives rise to a wealth of new group-invariant filters for investigating classical problems such as filtering, pattern recognition, data compression, and noise reduction.

In this way, this paper is a part of the growing body of work devoted to developing and applying general group theoretic machinery to a deeper understanding of the foundations of signal processing (cf. Foote et al. (2000) for an extensive list of references). To cite but a few examples, this more general framework has already led to important breakthroughs in the design and implementation of Abelian FFT algorithms (Auslander et al., 1996) and is also at the heart of new initiatives directed towards the exploration and application of other important transforms in both classical signal processing problems as well as more novel nonAbelian transforms (Moura et al., 2001). This work combines in equal measure, symbolic computation and numerical experimentation. The former makes possible the latter, as the implementations require algorithms that allow us to work within the related wreath product groups. So, while the present paper focuses more on the numerical aspects, symbolic techniques underlie the results. This interplay of theoretical development, symbolic implementation, and numerical experimentation is a step towards an expanded view of what encompasses the discipline of signal processing.

The organization is as follows. We begin with a development some basic theory, define the concepts of a group-based 2D transform, work through realistic examples, and provide some experiments. Within this group theoretic context we define concepts of decomposable, separable, and inseparable transforms. These share various properties with the usual signal processing definitions. In particular, we construct a 2D separable WPT. This is the natural extension of the earlier work and also generalizes the familiar 2D DFT and 2D Haar transforms. These 2D transforms have a concomitant multiresolution spectral transform via a multichannel pyramid-structured filter bank. We describe a 2D quadtree scanning scheme and compare this separable 2D transform with the 1D WPT developed in Foote et al. (2000) (which we show to be inseparable). The associated 2D quadtree spectrum is also illustrated for a standard test image.

2. Product structures and multidimensional signal processing

We briefly recall some terminology and results from Foote et al. (2000). Let X be a finite set and let G be a finite group acting on X (on the left). Let $L(X)$ denote the vector space of all complex-valued functions on X. There is a natural choice of basis for $L(X)$ given by the set of delta functions indexed by elements $x \in X$. The *spatial domain* on the set X is $L(X)$ along with this choice of basis. There is an obvious correspondence between function expansions in terms of the delta functions and sample values. The action of G on X extends by linearity to an action on $L(X)$ according to

$$(\alpha f)(x) = f(\alpha^{-1}x) \qquad \text{for all } \alpha \in G \text{ and } x \in X.$$

In the language of representation theory, any vector space (for instance $L(X)$) which admits a group G acting as linear transformations is called a $\mathbb{C}G$-*module*, and its G-invariant subspaces are called $\mathbb{C}G$-submodules.

When X has cardinality n and the elements of X are ordered and indexed sequentially, then $L(X)$ may be viewed as the n-dimensional vector space of all complex-valued 1D signals of length n. In this paper we are interested in the special case when $X = X_1 \times X_2$. It is then natural to view X (after ordering X_1 and X_2) as an $n \times m$ array whose points may be thought of as representing the pixel positions in a 2D rectangular array. Each element f of the spatial domain $L(X)$ then can be thought of as an image whose color or intensity at position i, j is $f(x_i, y_j)$.

Within the usual signal processing literature (see Pratt, 1991; Gonzalez and Woods, 1993) the decomposition of the index set X as a cartesian product $X_1 \times X_2$ appears to characterize the notion of a 2D transform. This type of decomposition is of course possible whenever the size of the index set (domain) is composite, but is only useful in cases when the decomposition is natural. Obvious example include images (2D), movies, (3D time-indexed images) and volumes (3D), or even 4D time-indexed volumes, such as occur in functional magnetic resonance imaging. The subject of *tactical designs* from which designed experiments arise give an example of a situation in which the index set is a subset of a product space, but not an entire product space (see Scheffé, 1999).

The advantages of this decomposition of the signal domain are, by and large, in the opportunities which are made possible for decomposing the range of the accompanying signal space, $L(X_1 \times X_2)$. The product structure on the domain then implies a natural tensor product structure for the signals (which again, is *possible* as long as the total dimension of the space is composite). This tensor product structure can be exploited both for processing speed (i.e. efficiency of computation) as well as for related psychophysical reasons (e.g. for enabling vertical and horizontal edge detection).

A product decomposition of the domain also permits the possibility of a group action by a cartesian product of (possibly distinct) groups. This is the setting which gives rise to the multidimensional forms of the DFT and FFT. We now give a general formulation of these ideas.

Definition 1. We say the action of G on $X_1 \times X_2$ is *decomposable* if G is a direct product, $G \cong G_1 \times G_2$, such that G_i acts on X_i for $i = 1, 2$. More generally, for fixed n and m we say the action of G on a set X of cardinality nm is *decomposable* if there are sets

X_1, X_2 of cardinalities n, m acted on by groups G_1, G_2 respectively, such that (X, G), and $(X_1 \times X_2, G_1 \times G_2)$ are isomorphic group actions.

The above notion of decomposability is essentially equivalent to specifying that X be a set of indeterminates $x_{i,j}$, $1 \leq i \leq n$, $1 \leq j \leq m$, and letting G_1 be any subgroup of the symmetric group S_n and G_2 be any subgroup of S_m. Then $G = G_1 \times G_2$ acts on X by permuting rows and columns of the array X as usual:

$$(\sigma, \tau)(x_{i,j}) = x_{\sigma^{-1}(i),\tau^{-1}(j)}.$$

The latter perspective suggests that we think of X as an array or image whose rows and columns are permuted by G, and its spectrum will be computed by scanning X along its rows or columns. However, we shall use the Cartesian product formulation to conjure a less rectilinear decomposition of arrays that are nonetheless decomposable in *some* scanning scheme. We shall exploit this more "coordinate free" viewpoint when we define a 2D quadtree scanning scheme in Section 4.

Example 1. Let $Z_n \times Z_m = \langle \sigma \rangle \times \langle \tau \rangle$ be the direct product of cyclic groups, and let $X_1 \times X_2 = \{0, \ldots, n-1\} \times \{0, \ldots, m-1\} = \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. We think of $X_1 \times X_2$ as corresponding to the usual grid points on the $n \times m$ lattice. In this way to any function $f \in L(X_1 \times X_2)$, we associate a matrix, given by the corresponding function values. Then $(\sigma^i, \tau^j)(r, s) = (\overline{i+r}, \overline{j+s})$ with the bar indicating the appropriate modular arithmetic. Thus the group action performs independent cyclic shifts of left and right indices, i.e. it cyclically shifts the rows and columns of the arrays. This group action gives rise to the familiar (decomposable) 2D DFT.

Assume now that n and m are relatively prime. Then $Z_n \times Z_m$ is a cyclic group of order nm with (σ, τ) as a generator (see Proposition 6 in Section 5.2 of Dummit and Foote, 1999). By the Chinese remainder theorem (Section 7.6 of Dummit and Foote, 1999), for each $(r, s) \in X_1 \times X_2$ there is a unique integer k in the set $Y = \{0, 1, \ldots, nm-1\}$ such that k is congruent to $r \mod n$ and to $s \mod m$, i.e. $(\overline{k}, \overline{k}) = (r, s)$. Fix this one-to-one correspondence between Y and $X_1 \times X_2$, and let a generator μ for Z_{nm} act on Y by $\mu(k) = k + 1 \mod nm$. Then the group actions (Y, Z_{nm}) and $(X_1 \times X_2, Z_n \times Z_m)$, where $\mu \leftrightarrow (\sigma, \tau)$, are seen to be isomorphic. In particular, the action of the cyclic group Z_{nm} on Y by cyclic shifts is decomposable.

Example 2. We shall see in Section 4 that the familiar 2D separable Haar transform on $2^n \times 2^n$ images, as described in Section 5.9 of Jain (1989), arises from the decomposable group action of $G_1 \times G_2$ on rows and columns of the space of square images, where G_1 and G_2 are both Sylow 2-subgroups of the symmetric group on 2^n points. In Section 4 we describe these groups explicitly as (wreath product) automorphism groups of trees, and generalize this example to other wreath product groups.

3. Representation theory and spectral transforms

Before focusing on the wreath product groups we refine of the notion of a decomposable actions to that of a separable group-based transformation.

Continuing in the above notation, let $G = G_1 \times G_2$ act decomposably on $X = X_1 \times X_2$. With no essential loss of generality we also impose the condition that G acts transitively on X, or equivalently, that G_i acts transitively on X_i, for $i = 1, 2$. In this setting, if H_i is the stabilizer of some point in X_i for $i = 1, 2$ then by basic permutation group theory we obtain the following identifications:

$$L(X) \cong L(G/H) \cong L(G_1/H_1 \times G_2/H_2) \cong L(X_1 \times X_2) \cong L(X_1) \otimes L(X_2)$$

where the tensor product of vector spaces is over $\mathbb{C}$.

By the basic theory of tensors (see Dummit and Foote, 1999), if $e_1, \ldots, e_n$ is any basis of $L(X_1)$ and $f_1, \ldots, f_m$ is any basis of $L(X_2)$, then the simple tensors $e_i \otimes f_j$ form a basis of the nm-dimensional space $L(X_1) \otimes L(X_2)$. In the special case when e_i is the δ-function supported on $x_i \in X_1$ and f_j is the δ-function on $y_j \in X_2$, then $e_i \otimes f_j$ is the δ-function supported on $(x_i, y_j) \in X$. In this regard, when viewing the elements of $L(X)$ as 2D images, the e_is index the rows and the f_js index the columns. The action of $G_1 \times G_2$ on $L(X_1) \otimes L(X_2)$ is defined on simple tensors by $(\sigma, \tau)(e_i \otimes f_j) = (\sigma e_i) \otimes (\tau f_j)$ and extended by linearity to an action on all tensors. In this way the matrix of the action of each group element (σ, τ) on $L(X)$ with respect to the basis $e_i \otimes f_j$ is the tensor or Kronecker product of the matrices representing σ on $L(X_1)$ and τ on $L(X_2)$ with respect to the bases $\{e_i\}$ and $\{f_j\}$.

Assume $L(X_1)$ and $L(X_2)$ have a decomposition into irreducible subspaces under the actions of G_1 and G_2 respectively as

$$L(X_1) \cong V_1 \oplus V_2 \oplus \cdots \oplus V_r \qquad \text{and} \qquad L(X_2) \cong W_1 \oplus W_2 \oplus \cdots \oplus W_s. \tag{3.1}$$

Since tensor products distribute over direct sums, we have a vector space decomposition into G-invariant subspaces (i.e. into $\mathbb{C}G$-submodules)

$$L(X_1 \times X_2) \cong \bigoplus_{i,j} V_i \otimes W_j. \tag{3.2}$$

Since the irreducible representations of a direct product are the tensor products of irreducible representations of the direct factors (see Curtis and Reiner, 1981), the decomposition in (3.2) exhibits $L(X_1 \times X_2)$ as a direct sum of irreducible G-invariant subspaces.

Fix some choice of bases for the irreducible subspaces V_i for all i, and let A be the change of basis (transformation or matrix) from the original (e.g. δ-function) basis on $L(X_1)$ to a new basis on $L(X_1)$ consisting of the union of bases for the irreducible subspaces V_i in (3.1). We call A a *group-based Fourier transform*, and for each $h \in L(X_1)$, Ah is called the *spectrum* or *group-based Fourier decomposition* of h. Likewise let B be a group-based Fourier transform on $L(X_2)$. The decomposition (3.2) into irreducible subspaces shows that $A \otimes B$ is thus a group-based Fourier transform on $L(X_1 \times X_2)$ with respect to $G_1 \times G_2$. These representation theoretic observations motivate the following generalization of the notion of decomposable.

Definition 2. For fixed integers $n, m \geq 2$ we say the $\mathbb{C}G$-module L is *separable* if there are $\mathbb{C}G$-modules V and W of dimensions n, m respectively, such that L is isomorphic to $V \otimes_{\mathbb{C}} W$ (as $\mathbb{C}G$-modules), and there are decompositions $V \cong V_1 \oplus \cdots \oplus V_r$ and

$W \cong W_1 \oplus \cdots \oplus W_s$ of V and W into irreducible $\mathbb{C}G$-submodules such that $V_i \otimes W_j$ is irreducible for all i and j.

We will occasionally use the term *"2D separable"* synonymously with "separable" when it seems desirable to emphasize concepts applied to 2D image processing. The condition in the definition that the module be a tensor product is essentially specifying that the module be "group-based 2D"; and we shall see that the condition that the products of the irreducible components also be irreducible ensures that the group-based spectral transform is a separable linear transformation (i.e. of the form AXB^t). A representation is called *separable* if a $\mathbb{C}G$-module affording it is separable. A module (or representation) that is not separable is called *inseparable*. Note that a module L may be separable with respect to some degrees n, m but not others whose product is $\dim L$. Also note that the action of G is inherent in the notion of separability, and L need not necessarily be a permutation module for G. Indeed, if G acts trivially on L (e.g. if G is the trivial group) then, as observed earlier, L is separable for any nontrivial factorization nm of its dimension.

A representation is separable if and only if its character is a product of two characters α and β of degrees n and m such that when these are written as sums of irreducible characters $\alpha = \alpha_1 + \cdots + \alpha_r$ and $\beta = \beta_1 + \cdots + \beta_s$, then the character $\alpha_i \beta_j$ is irreducible for all i and j.

Our earlier discussion showed the following:

Proposition 3.1. *If $G = G_1 \times G_2$ acts decomposably on $X = X_1 \times X_2$, then $L(X)$ is a separable $\mathbb{C}G$-module.*

A group action need not be decomposable in order that the associated representation be separable. For example, separable modules can be constructed by taking G to be any subgroup of a direct product of Abelian groups $G_1 \times G_2$ acting on $X_1 \times X_2$ (or on any modules V, W); the "Abelian" assumption ensures that irreducible $\mathbb{C}G$-modules are 1D, hence their tensor products are also irreducible. Another particularly interesting example comes from the computation of DFTs and FFTs for crystallography. The crystallographic symmetry groups or space groups are non-decomposable, yet their action on three-space has an associated invariant decomposition as a triple tensor product which allows for great savings when computing the Fourier transform (Auslander et al., 1988; An et al., 1990). As another example, the largest Mathieu group M_{24} has irreducible representations of degrees 23 and 45 whose product is irreducible, i.e. in this case $r = s = 1$ (cf. Conway et al., 1985).

We may also view the separable spectrum of an $n \times m$ image $h \in L(X_1 \times X_2)$ as a matrix-valued $(r \times s)$-array

$$\widehat{H} = \begin{bmatrix} \widehat{H}_{11} & \widehat{H}_{12} & \dots & \widehat{H}_{11} \\ \widehat{H}_{21} & \widehat{H}_{22} & \dots & \widehat{H}_{11} \\ \dots & \dots & \dots & \dots \\ \widehat{H}_{r1} & \widehat{H}_{r2} & \dots & \widehat{H}_{rs} \end{bmatrix} \tag{3.3}$$

or as an $n \times m$ array in an especially coherent format. The representation $\widehat{H}$ is very important for future directions (see Section 6) such as for harmonic analysis of group-based convolution and correlation. Let basis functions for the irreducible G_1-invariant subspace

V_i (of dimension k_i) be indexed as $\widehat{e}_{n_i}, \widehat{e}_{n_i+1}, \ldots, \widehat{e}_{n_i+k_i-1}$, and let basis functions for the G_2-invariant subspace W_j (of dimension l_j) be $\widehat{f}_{m_j}, \widehat{f}_{m_j+1}, \ldots, \widehat{f}_{m_j+l_j-1}$. Then the coefficients for the $G_1 \times G_2$-invariant irreducible subspace $V_i \otimes W_j$ spanned by the $\widehat{e}_p \otimes \widehat{f}_q$ form a $k_i \times l_j$ submatrix of the $n \times n$ spectrum array whose upper left-hand entry appears in position n_i, m_j. Examples of this are given in Section 5.

A significant computational feature of separable representations is that their group-based Fourier transforms may be computed with generally fewer operations than inseparable transforms of the same degrees. In the notation of Definition 2, let $\{e_i\}$ and $\{f_j\}$ denote bases of V and W respectively, and let A and B denote change of bases to the direct sums of irreducible subspaces for V and W respectively. Then each $h \in L$ may be written as

$$h = \sum_{i=1}^{n} \sum_{j=1}^{m} h(x_i, y_j) e_i \otimes f_j = \sum_{i=1}^{n} e_i \otimes w_i,$$

where $w_i = \sum_{j=1}^{m} h(x_i, y_j) f_j$. For each i let

$$Bw_i = \sum_{j=1}^{m} b_{i,j} \widehat{f}_j$$

where $\widehat{f}_1, \ldots, \widehat{f}_m$ is a group-based Fourier basis for W. Since $\{\widehat{f}_j\}$ is a basis of W, $\{e_i \otimes \widehat{f}_j\}$ is a basis of L. Since the linear transformation $A \otimes B$ is the composition of the transformations $A \otimes 1$ and $1 \otimes B$ (in any order),

$$\begin{aligned}
(A \otimes B)(h) &= (A \otimes 1)\left(\sum_{i=1}^{n} e_i \otimes Bw_i\right) = (A \otimes 1)\left(\sum_{i=1}^{n} e_i \otimes \sum_{j=1}^{m} b_{i,j} \widehat{f}_j\right) \\
&= (A \otimes 1)\left(\sum_{i=1}^{n} \sum_{j=1}^{m} b_{i,j} e_i \otimes \widehat{f}_j\right) \\
&= (A \otimes 1)\left(\sum_{j=1}^{m} v_j \otimes \widehat{f}_j\right) = \sum_{j=1}^{m} Av_j \otimes \widehat{f}_j
\end{aligned}$$

where $v_j = \sum_{i=1}^{n} b_{i,j} e_i$. In other words, writing the coefficients of each element h of L as a matrix whose i, j entry is the coefficient of $e_i \otimes f_j$, the group-based Fourier transform on h may be found by first transforming W into irreducibles by applying B to each row w_i of h. This results in the set of vectors Bw_i whose coefficients, $b_{i,j}$, are placed in the ith row of an $n \times m$ matrix. Denote the columns of this matrix by $v_1, \ldots, v_m$. Then $A \otimes B$ is computed by now applying A to each column of the latter matrix, i.e. by the matrix multiplication $A[b_{i,j}]$. In summary we have the following:

Proposition 3.2. *If $L \cong V \otimes W$ is a separable $\mathbb{C}G$-module and A, B are matrices representing separable group-based Fourier transformations on V and W respectively (as in Definition 2), then each h in L may be represented by an $n \times m$ matrix H such that the* *separable group-based Fourier transformation on H is AHB^t, where B^t is the transpose of B.*

Example 1. The decomposable action of $Z_n \times Z_m$ on $n \times m$ arrays described in Example 1 of the preceding section gives the familiar 2D DFT (see Section 5.5 of Jain, 1989). The $n \times m$ 2D separable DFT on an image $[u(i, j)]$ is defined by

$$\widehat{u}(k,l) = \sum_{p=1}^{n}\sum_{q=1}^{m} u(p,q) W_n^{k(p-1)} W_m^{l(q-1)}$$

where $W_N = e^{2\pi i/N}$, $i = \sqrt{-1}$. In matrix notation this transform becomes $[\widehat{u}(k,l)] = D_n[u(i, j)]D_m^t$, where D_N is the matrix of the familiar N-point DFT on signals of length N. When $n = m$ and the n-point DFT is normalized by multiplying by $1/n$, the resulting (symmetric) matrix U_n is unitary, and the normalized 2D separable DFT for $n \times n$ images becomes $U_n H U_n$. Since $U_n^{-1} = U_n^*$, where the star denotes complex conjugation, the (normalized) inverse 2D separable DFT on $n \times n$ arrays is $U_n^* Y U_n^*$.

Example 2. The decomposable action of the permutation groups described of Example 2 in Section 2 leads to the 2D separable Haar transform, as described in Section 9.5 of Strang and Nguyen (1996) or Section 5.9 of Jain (1989). In these books the Haar transform is exhibited both as a transform and in matrix format (explicitly, for small degrees). This 2D Haar transform is a special case of the 2D WPC transform described in Section 4.

Once again, it is useful and important to relate these concepts to the usual ones of signal and image processing. A generic 2D transform for data $f(x, y)$ is written as

$$T(u,v) = \sum_{y\in Y}\sum_{x\in X} f(x,y)g(x,y,u,v)$$

where g is the transform *kernel*. The kernel is called *separable* if there is a factorization $g(x, y, u, v) = g_1(x, u)g_2(y, v)$ and called *symmetric* if $g_1 = g_2$. The separable case permits T to be written as a matrix–vector product transform $T = G_1 f G_2^t$ (t denotes transpose) for suitable matrices G_1 and G_2. In the symmetric case G_2 is replaced by G_1. Notice that without separability, this 2D transform would be computed as a simple matrix–vector multiplication with indexing that would realize the data as a column vector and the kernel as a matrix.

Remark. The increased efficiency in computing a tensor product transform encourages the recognition of such structure when possible. This is part of the motivation behind the SPIRAL (Signal Processing algorithms Implementation Research for Adaptable Libraries) project, a multidisciplinary effort targeted at the automatic generation of tensor product factorizations and code computing signal processing transforms (Moura et al., 2001).

In fact, even if a tensor product decomposition exists, it need not be unique, and the different factorizations may entail very different implementations whose efficiency can be architecture dependent. The paper (Auslander et al., 1996) shows how different tensor product formulations for the Abelian FFT are related and characterized by certain cohomological invariants. The relationship among computer architecture, tensor product structure, and execution time is also investigated.

The computational savings due to tensor product structure suggest that it even can be useful to recognize if a transform may only be approximated by a tensor product. In such situations accuracy may be traded off against efficiency. In Pitsianis (1997), Pitsianis studies the problem of finding the tensor product of equal dimension, closest in Frobenius norm to a given matrix. This problem is both studied in an unconstrained (arbitrary tensor product) and constrained (tensor product with further specified structure) manner. A code generation system which specifies efficient implementations of the tensor product transform is also presented.

4. Two-dimensional wreath product group-based transforms

The wreath product groups discussed in this section arise as automorphism groups of spherically homogeneous rooted trees (SHRTs). For clarity of notation and with an eye to specific applications to image processing, such as those in Section 5, we shall restrict attention to particular families of these trees. For general SHRTs the nomenclature, the group actions and associated representations, and the application of these to 1D signal processing are developed fully in Foote et al. (2000). Based on this foundation, the treatment herein will make it transparent to the interested reader how to extend 1D results for general SHRTs to the 2D setting.

Let $\mathcal{Q}_n$ denote the tree which has a root node (at level zero), four nodes descending from it (these are at level 1), four nodes descending from each node at level 1 (these 16 nodes are at level 2), and so on with each node having four children, terminating in the 4^n nodes at level n (called the leaves of the tree). We call $\mathcal{Q}_n$ the *quadtree* of degree n. Let $Z(n)$ be the group of automorphisms of $\mathcal{Q}_n$ that fixes the root node and at each level allows independent cyclic permutations of the four children of each vertex at that level. Thus $Z(n)$ has the structure of an iterated *wreath product of cyclic groups* of order 4:

$$Z(n) \cong Z_4 \wr Z_4 \wr \cdots \wr Z_4 \quad (n \text{ factors})$$

and hence is called a *WPC group*. The order of $Z(n)$ is $4^{(4^n-1)/3}$.

Another important family consists of the binary SHRTs, $\mathcal{B}_n$, with $n+1$ levels, where each node has two children; the corresponding WPC group of automorphisms, also denoted by $Z(n)$, is the iterated wreath product of cyclic groups of order 2.

In each family the WPC group $Z(n)$ transitively permutes the set X of leaves of the tree (i.e. the nodes at level n). This action affords a group representation on the spatial domain $L(X)$ which is studied extensively in Foote et al. (2000) and Mirchandani et al. (2000). Explicit WPC group-based (fast) Fourier transforms, filter bank algorithms for computing (multiresolution) spectra, and many examples are exhibited therein. In particular, when $n = 1$, the WPC group is just the cyclic group of order 4 (quadtree) or order 2 (binary tree), and the WPC group-based transform is the 4- or 2-point DFT respectively. Also, for arbitrary n, the group-based Fourier transform obtained from the WPC group action on the binary tree $\mathcal{B}_n$ is the familiar Haar transform on signals of length 2^n, and discrete Haar wavelets at various scales are WPC group-based Fourier basis functions spanning the irreducible group invariant subspaces of $L(X)$.

The quadtree $\mathcal{Q}_n$ arises in a natural context in image processing. The leaves of this tree may be used to index pixel positions in a $2^n \times 2^n$ array scanned in the quadtree nested grid fashion as follows: Subdivide the array into its four $2^{n-1} \times 2^{n-1}$ quadrants, and likewise subdivide each of these, and so on. Then for any $k \in \{0, 1, \ldots, n\}$ we obtain 4^{n-k} subarrays in a (quadtree) nested grid at level $n-k$, each of which is a $2^k \times 2^k$ matrix. Each such submatrix is associated to a node of the tree at level $n-k$. The four arrays descending from each node (i.e. the four quadrants of each subimage) are then cyclically permuted clockwise (fixing their orientation) by generators of the Z_4 factors at the corresponding level in the WPC group. Note that even though the spatial domain has a natural 2D structure, this nested grid decomposition effectively scans images into a space of 1D vectors upon which the WPC group acts. Evidently this group action is indecomposable because $Z(n)$ is a 2-group with a cyclic center, hence it cannot be a nontrivial direct product. We prove the stronger result: this action on 2D arrays is inseparable.

Proposition 4.1. *For all $n \geq 2$ the 4^n-dimensional representation of the WPC group $Z(n)$ on $2^n \times 2^n$ images indexed by the quadtree $\mathcal{Q}_n$ is inseparable for any nontrivial factorization of 4^n.*

Proof. By definition of $Z(n)$, each nonidentity element induces a nontrivial permutation of the leaves of the quadtree, so the representation of $Z(n)$ on $L(X)$ is *faithful*. Let π be the character of this representation, so that $\pi(g)$ equals the number of leaves fixed by g for any $g \in Z(n)$ (see Example 3 on p. 832 of Dummit and Foote, 1999). Also, the degree of the representation is $\pi(1) = 4^n$.

By way of contradiction assume the representation is separable. By the remarks following Definition 2 this means π is a product $\alpha\beta$ of nonlinear characters. We compare the possible values of $\alpha\beta$ and π on elements in the center of $Z(n)$. An element in the center of $Z(n)$ commutes with all elements of $Z(n)$, hence must induce the same permutation on each block of four leaves descending from a common node at level $n-1$. Thus the center of $Z(n)$ is generated by an element w of order 4 that cyclically permutes the four elements in each such block (with respect to a fixed labeling determined by the first block), i.e. w is a product of 4^{n-1} disjoint 4-cycles. Let $z = w^2$ be the unique element of order 2 in the center, so that z also fixes no leaves, i.e. $\pi(z) = 0$.

If both α and β have nontrivial kernels, then since $Z(n)$ is a 2-group with cyclic center, Theorem 1(2) in Chapter 6 of Dummit and Foote (1999) shows that z is in the kernel of both α and β, hence is in the kernel of π. This in turn forces $\pi(z) = \pi(1) = 4^n$, a contradiction. We may therefore choose notation so that α is faithful.

By Theorem 3.1 of Foote et al. (2000), $Z(n)$ contains a normal subgroup B, the kernel of the action of $Z(n)$ on the nodes at level $n-1$, which is a direct product of 4^{n-1} cyclic groups of order 4. Thus every faithful representation of $Z(n)$ has dimension at least 4^{n-1}, so α has degree at least 4^{n-1}. Since the product of the degrees of α and β is 4^n, α must have degree either 4^{n-1} or $2 \cdot 4^{n-1}$, and β must have degree 4 or 2 respectively.

If α has degree 4^{n-1}, it must be irreducible—otherwise, as above, z would be in the kernel of each of its constituents, hence in the kernel of α. By Schur's lemma z is represented by a scalar matrix, hence $\alpha(z) \neq 0$. Now β has degree 4. By the same reasoning, either z is in the kernel of β or β is irreducible (with $n = 2$); in either situation $\beta(z)$ is likewise nonzero. This is a contradiction because $0 = \pi(z) = \alpha(z)\beta(z)$.

It remains to consider when α has degree $2 \cdot 4^{n-1}$ and β has degree 2. Arguing as above, we must have $\beta(z) \neq 0$ and $\alpha(z) = 0$. By Schur's lemma, α is reducible, hence it follows that $\alpha = \alpha_1 + \alpha_2$ where α_1 is a faithful irreducible character of degree 4^{n-1}. If β were irreducible, then by the remarks following the definition of separable, $\alpha_1\beta$ would be an irreducible constituent of π of degree $2 \cdot 4^{n-1}$; this would contradict the fact that every irreducible constituent of π has degree at most 4^{n-1} by Theorem 4.8 of Foote et al. (2000). Thus β must be a sum, $\beta_1+\beta_2$, of two degree 1 characters. Hence every irreducible constituent γ of α gives two irreducible constituents $\gamma\beta_1$ and $\gamma\beta_2$ of the same degree of π. This contradicts the fact that π has exactly three irreducible constituents of degree 4^{n-1} by Theorem 4.8 of Foote et al. (2000) and so completes the proof. □

At first blush, there are obvious 2D separable WPC group-based transforms for images of certain sizes. For trees with sets of leaves X_1 and X_2, let G_1 and G_2 be WPC groups acting on X_1 and X_2 respectively. Then $G_1 \times G_2$ acts decomposably on $L(X_1 \times X_2)$. Indeed, we recover the 2D separable Haar transform as a special case of this: when the trees are both binary and the rows and columns of a space of arrays are indexed by X_1 and X_2 respectively, and so are permuted by $G_1 \times G_2$. Likewise the separable 2D DFT is a special case of an action by a direct product of WPC groups, when both trees have only one nonzero level but are not restricted to being binary or quadtrees. In general when G_1 and G_2 are WPC groups permuting the rows and columns of arrays indexed by the leaves of trees in this manner we shall refer to the resulting 2D transform (or spectrum) as the *2D separable rectilinear WPC transform (spectrum, respectively)*.

We now introduce a variant of this family of 2D group-based transforms—essentially differing from the rectilinear transforms just described by a different scanning scheme—and explore their properties. As described above, $2^n \times 2^n$ images may be given a nested grid structure indexed by the quadtree $\mathcal{Q}_n$ and acted upon (in a 1D manner) by the WPC group $Z(n)$. We consider this spatial domain of 2D images imbued with the same quadtree nested grid structure, but acted on in a decomposable 2D fashion by certain *subgroups* of the full WPC group $Z(n)$. The group-based Fourier transforms resulting from the action of these subgroups thus give a "refinement" of the Fourier transforms for the whole group. Again, for clarity we focus on decomposing $2^n \times 2^n$ arrays under the action of the direct product of two isomorphic WPC subgroups of $Z(n)$; it is straightforward to generalize to nonisomorphic WPC subgroups, even with permutation representations of different degrees.

Assume $n = 2k$ and view the WPC group $Z(n)$ as acting on the space of all $2^n \times 2^n$ images in the quadtree fashion. Let G_1 be the subgroup of $Z(n)$ that permutes the $2^k \times 2^k$ blocks in the nested grid among themselves but does not alter the relative position or orientation of the entries within each of these subarrays (e.g. does not rotate the subarrays). The subgroup of $Z(n)$ that stabilizes each of these subarrays is a direct product of groups of type $Z(k)$. Let G_2 be the diagonal subgroup in this direct product stabilizer that maps each of these subarrays to themselves and acts simultaneously as the same relative permutation on each subarray. For example, there is an element of G_2 that simultaneously rotates each subarray clockwise by 90°; but no element of G_2 rotates one array while fixing the others. Since elements of G_2 act the same way on each array, they commute with the elements of G_1. It follows that

$$G = G_1 \times G_2 \cong Z(k) \times Z(k).$$

(The subgroup structure and actions of WPC groups on sub- and quotient trees is detailed in Theorem 3.1 of Foote et al. (2000).) This subgroup determines a "square" *2D separable quadtree decomposition of* $2^{2k} \times 2^{2k}$ *images*, where the "rows" are the $2^k \times 2^k$ subarrays in the nested grid decomposition (each of which is scanned in a quadtree fashion to obtain a vector of length 2^k), and each column consists of one entry chosen from the same relative position in each of these subarrays. Moreover, because G is a subgroup of $Z(n)$, each G-invariant irreducible subspace in the spectral decomposition may be contained in a $Z(n)$-invariant irreducible component, so the 2D G-based spectral decomposition refines the 1D $Z(n)$-based decomposition.

In light of Proposition 3.2 we may summarize these calculations as follows:

Proposition 4.2. *Let L be the spatial domain of all $2^{2k} \times 2^{2k}$ images and let $G = G_1 \times G_2$ with $G_1 \cong G_2 \cong Z(k)$ acting on L via a transitive permutation representation. Let A be a $2^{2k} \times 2^{2k}$ WPC group-based Fourier transform matrix for the WPC action of $Z(k)$ on the space of 1D vectors of length 2^{2k}.*

(1) *If G_1 and G_2 act by permuting the rows and columns of images respectively via the usual WPC action of $Z(k)$ on vectors of length 2^{2k}, then the rectilinear 2D separable WPC group-based spectrum of an image H is AHA^t.*

(2) *If $G_1 \times G_2$ is a subgroup of the WPC group $Z(2k)$ acting in the quadtree fashion on images, then the 2D separable quadtree spectrum of an image H is $A(\theta(H))A^t$, where θ is the permutation of the entries of a $2^{2k} \times 2^{2k}$ matrix that transforms the $2^k \times 2^k$ blocks in the quadtree nested grid decomposition of H into the rows of a matrix $\theta(H)$. In this situation, each G-invariant irreducible subspace of L may be contained in one of the $Z(2k)$-irreducible components.*

Note that because the permutation actions of G_1 and G_2 are the same, the transformation matrix B appearing in Proposition 3.2 equals A. These 1D WPC group-based transforms are described in detail in Foote et al. (2000), with explicit matrices given for small k; numerous examples of 1D quadtree spectra are also depicted. The permutation $\theta : H \mapsto \theta(H)$ is the mathematical formulation of utilizing a different scanning scheme (in this case the quadtree scheme).

In Section V.A of Foote et al. (2000) we also described how to display the spectrum of each $2^n \times 2^n$ image in another $2^n \times 2^n$ array in such a way that Fourier coefficients from the group invariant irreducible subspaces formed square submatrices of the spectral array—we called this depiction the *quadtree spectrum* of an image. This quadtree spectrum is obtained from the 1D WPC spectrum of length 4^{2k} by simply reversing the quadtree scanning scheme, i.e. by applying θ^{-1}. This gives a more precise geometric realization of the last sentence of Proposition 4.2.

Corollary 4.1. *In case (2) of Proposition 4.2, $Y = \theta^{-1}(A(\theta(H)A^t)$ is the 2D separable quadtree spectrum of H with entries rearranged so that the G-invariant subspaces of Y appear as submatrices of Y, and each of these submatrices lies within an irreducible component of the 1D WPC group-based quadtree spectrum $\mathcal{Q}(H)$ (described in Foote et al., 2000).*

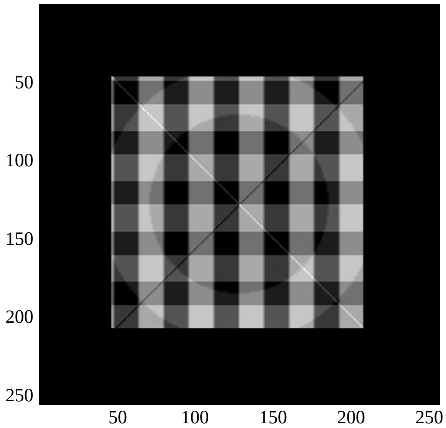

Fig. 1. Test image.

The matrix Y in Corollary 4.1 is called the *2D separable quadtree spectrum* of a $2^{2k} \times 2^{2k}$ image. Unlike the 1D WPC quadtree spectrum for the group $Z(2k)$, the $Z(k) \times Z(k)$ 2D spectral components need not be square subarrays in the 2D separable quadtree spectrum. Figs. 1 and 2 illustrate the 2D rectilinear and quadtree spectra of a standard $2^8 \times 2^8$ test image. (In these figures the magnitudes of the complex spectral coefficients are plotted. The evident rectangular subarrays that form a pattern in the spectra each represent a block of Fourier coefficients coming from a common irreducible component of the spectra.)

Finally, we record another important computational feature of the 2D separable transform. By Theorem 4.9 (and following) in Foote et al. (2000), the 1D WPC transform A for the $Z(k)$ group action is a block transform that may be computed by a pyramid-structured multichannel DFT filter bank. In other words, we may write A as a sequence of linear transforms: $A = A_k A_{k-1} \cdots A_1$, where A_i is the ith stage in the pyramid filter bank (decomposing the projection of images into the irreducible components spanned by discrete wavelet basis functions at “scale” 2^i). Correspondingly, the 2D separable quadtree transform $C : H \to Y$ in Corollary 4.1 decomposes as a sequence $C = C_k C_{k-1} \cdots C_1$. The iterative process

$$H \longmapsto H_1 = C_1(H) \longmapsto H_2 = C_2(H_1) \longmapsto \cdots \longmapsto H_k = C_k(H_{k-1}) = Y$$

where each C_i is a separable linear transformation, gives a fast (2D group-based) Fourier transform for computing the 2D separable WPC spectrum of H. The number of operations required for this transform on an $N \times N$ image H is $O(N^2 \log N)$, by the same reasoning as for the separable Haar transform in Jain (1989, Section 5.2). The pyramid structured filter bank implementation of this fast algorithm is discussed in the next section.

5. The 2D wreath product transform and multirate filter banks

In this section we describe the effects of the 2D separable WPC group-based transform $Y = CHC^t$, where H is a $2^{2k} \times 2^{2k}$ image and C is the analysis matrix associated with the

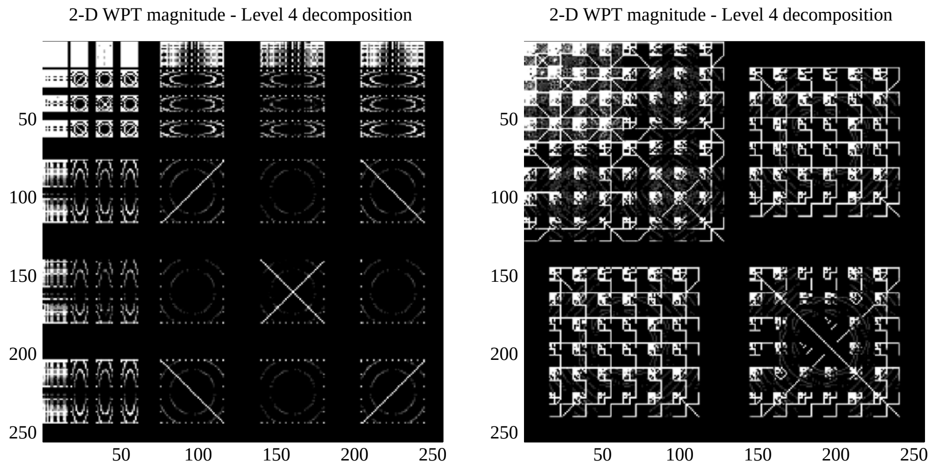

Fig. 2. The rectilinear and quadtree 2D amplitude spectra of the test image.

quadtree spectral decomposition obtained from the action of $Z(k)$ on 1D signals of length 2^{2k} indexed by the quadtree $\mathcal{Q}_k$. The group action is by the direct product of WPC groups $Z(k) \times Z(k)$ permuting rows and columns of images, hence this is the rectilinear 2D WPC transform.

A filter bank is a set of convolution operators, while a decimated filter bank has the output of the filter banks downsampled by a factor. These operations may be iterated to obtain a multi-level decimated filter bank. The direct representation of the WPC transform leads to a uniform 4-channel decimated multi-level analysis filter bank which is seen to progressively decompose the lowpass approximation into 16 subbands. After the first level of decomposition, the CHC^t formulation leads not only to the decomposition of the lowpass approximation at the first level, but also to operations on related subbands, but only in one dimension. Consequently at level two and beyond, along with the decomposition of the lowpass approximation, we see additional decompositions that give “thin” and “thick” versions of the original image. That is, the one dimension decompositions generate rectangular version approximations of the original image. In the 1D *WPC transform (WPT)*, C was generated by the iterative application of the 4-channel decimated filter bank to the lowpass approximation (see Foote et al., 2000, p. 116). We determine the 16 2D 4×4 subband filters and their frequency response and show that they filter specific 2D complex exponentials which corresponds to directional filtering of edges. These 2D 4×4 filters are the same as those for analysis with the 2D 4×4 DFT. Finally, we observe the effect of the 2D WPT on an image and interpret the magnitude and phase of the spectrum in terms of local directional edge detection characteristics of the filters.

5.1. Filter bank decomposition

Following standard terminology and development in the literature (cf. Strang and Nguyen, 1996; Vetterli and Kovacevic, 1995) we write the transform as $Y^t = C(CH)^t$

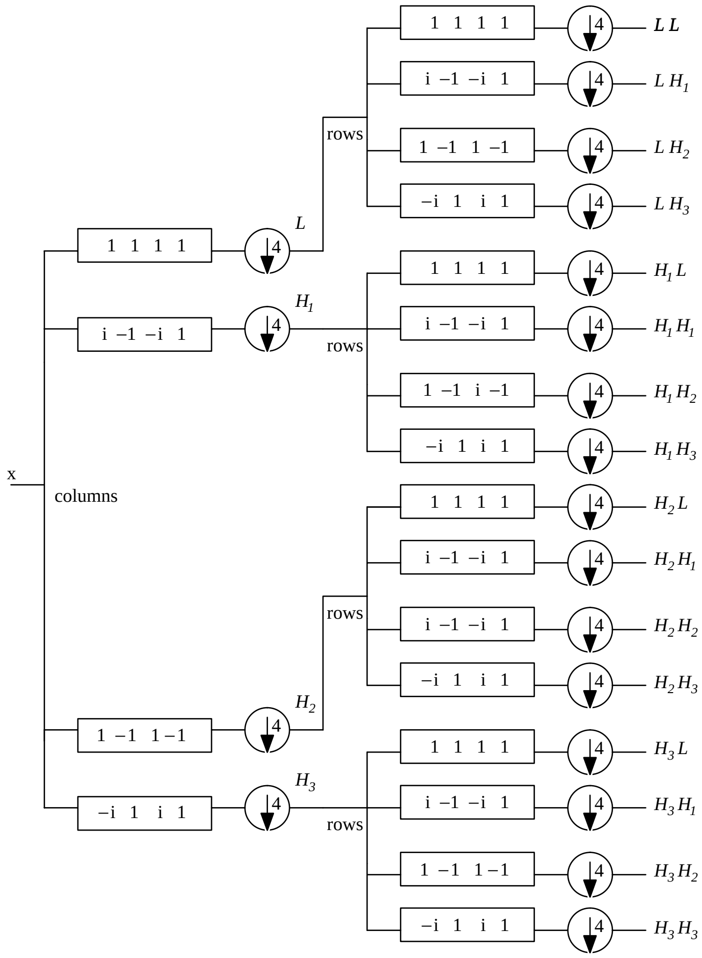

Fig. 3. 2D 4-channel analysis filter bank.

so that the spectrum Y is determined by first applying the 1D WPT to each column of H and then to each row of the result to obtain the rows of Y. The column-row operations are portrayed in the direct form as a 2D 4-channel decimated analysis filter bank as shown in Fig. 3. Consider first the 2D WPT at the first level of decomposition and let the input image H be of dimension $N \times N$ where $N = 4^M$. Write the first level of decomposition as $Y^1 = C^1 H(C^1)^t$, where C^1 represents the one-level 1D WPT. We have then

$$Y^1 = \begin{bmatrix} L \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} [H][L^t B_1^t B_2^t B_3^t] \tag{5.1}$$

where matrices $L = B_0$, B_1, B_2, and B_3 are of size $(N/4) \times N$. If each B_r is viewed as an $(N/4) \times (N/4)$ block matrix whose blocks consist of 1×4 vectors, then B_r is the block diagonal matrix with block $[\,1 \quad i^{-r} \quad i^{-2r} \quad i^{-3r}\,]$ down the diagonal. In other words, multiplication by B_r represents convolution by the downsampled-by-four filters:

$$\begin{aligned} &h_1 = [\,1 \quad 1 \quad 1 \quad 1\,] \qquad h_2 = [\,i \quad -1 \quad -i \quad 1\,] \qquad h_3 = [\,-1 \quad 1 \quad -1 \quad 1\,] \\ &h_4 = [\,-i \quad -1 \quad i \quad 1\,] \end{aligned} \tag{5.2}$$

where the filter coefficients represent causal sequences $[\,h_i(0) \quad h_i(1) \quad h_i(2) \quad h_i(3)\,]$, for $i = 1, \ldots, 4$. Multiplying through, we get

$$Y^1 = \begin{bmatrix} LHL^t & LHB_1^t & LHB_2^t & LHB_3^t \\ B_1HL^t & B_1HB_1^t & B_1HB_2^t & B_1HB_3^t \\ B_2HL^t & B_2HB_1^t & B_2HB_2^t & B_2HB_3^t \\ B_3HL^t & B_3HB_1^t & B_3HB_2^t & B_3HB_3^t \end{bmatrix} \tag{5.3}$$

where each submatrix of Y^1 represents 16 size $(N/4) \times (N/4)$ scaled and directionally filtered replicas of the image. Thus, the submatrices are of the form $Q^a H Q^{b^t}$, where Q^a, Q^b are L, B_1, B_2, or B_3.

Observe that each element $Y^1(k,l)$ of Y^1 can be written in bilinear form $c_k^{1^t} H c_l^1$ where $c_k^{1^t}$ and c_l^1 are the kth row and lth column of C^1 respectively, $k, l = 1, 2, \ldots, N$. Equivalently, in terms of submatrices Q^a, Q^b, any entry of Y^1 can be written as the product of the corresponding row k of Q^a and column l of Q^b, for all $k, l = 1, 2, \ldots, (N/4)$. Hence, $Y^1(k,l) = q_k^{a^t} H_{k,l}\, q_l^b$, where q_k^a and q_l^b are 4×1 nonzero vectors and $H_{k,l}$ the 4×4 submatrix of H with the top lefthand coordinates $(4(k-1)+1, 4(l-1)+1)$, for $k, l = 1, 2, \ldots, (N/4)$. If the bilinear forms are rewritten as the sum of the point-by-point product of $H_{k,l}$ and the outer product $q_k^a\, q_l^{b^t}$, then each entry $Y(k,l)$ may be thought of as the result of 2D filtering of 4×4 blocks of H with a 4×4 filter given by the outer product $h_k^t\, h_l$ of the 1D filters of Eq. (5.2). Hence the 2D WPT gives a 2D block transform where the 16 4×4 filters operate on 4×4 blocks of the $N \times N$ image to generate 16 $(N/4) \times (N/4)$ filtered images.

We now consider the second level of decomposition $Y^2 = A^2 H (A^2)^t$, where A^2 represents 1D WPT at level two. Equivalently, we can operate on the first level decomposition Y^1

$$Y^2 = \begin{bmatrix} L & \\ B_1 & \\ B_2 & O \\ B_3 & \\ & \\ O & I \end{bmatrix} \begin{bmatrix} L_1 & Q_1 & Q_2 & Q_3 \\ P_1 & & & \\ P_2 & & H_1 & \\ P_3 & & & \end{bmatrix} \begin{bmatrix} L^t & B_1^t & B_2^t & B_3^t & O \\ & & & & \\ & O & & & I \\ & & & & \end{bmatrix}$$

$$= \begin{bmatrix} LL_1L^t & LL_1B^t & .. & .. & .. & LQ_1 & LQ_2 & LQ_3 \\ . & & & & & B_1Q_1 & B_1Q_2 & B_1Q_3 \\ . & & & & & B_2Q_1 & B_2Q_2 & B_2Q_3 \\ . & & & & & B_3Q_1 & B_3Q_2 & B_3Q_3 \\ P_1L^t & P_1B_1^t & P_1B_2^t & P_1B_3^t & & & & \\ P_2L^t & P_2B_1^t & P_2B_2^t & P_2B_3^t & & H_1 & & \\ & & & & & & & \\ P_3L^t & P_3B_1^t & P_3B_2^t & P_3B_3^t & & & & \end{bmatrix}. \quad (5.4)$$

The matrix in the center of the first equation is a block matrix description of Y^1 in (5.3). In particular, L_1 and H_1 represent respectively the lowpass and all highpass outputs at the first level, while P_j, Q_j for $j = 1, 2, 3$ represent the bandpass outputs. Thus, the second level decomposition preserves the highpass output H_1, while operating on the lowpass image L_1 as before (i.e. as in the first level decomposition). In addition, subbands P_i and Q_i for $i = 1, 2, 3$ are operated upon in one dimension, each yielding four subbands in horizontal and vertical directions respectively. Further levels of decomposition, for a total of M levels, follow a similar pattern, where each lowpass approximation L_i is decomposed into 16 subbands, while subbands lying in its rows and columns are operated on in one dimension. This type of decomposition is a consequence of the 2D WPC group-based formulation CHC^t, where C is the 1D WPT matrix. Alternatively, the 2D WPT could be formulated as a filter bank following a logarithmic structure, iterating only on the lowpass approximation at each level. This would eliminate the 1D processing of the aforementioned subbands.

5.2. The 2D WPT spectrum

We now interpret the 2D WPT spectrum of an $N \times N$ image. (For comparison purposes, there is an extensive discussion of the 1D quadtree scanned WPT of an $N \times N$ image in Section V of Foote et al. (2000), including its capabilities for edge detection and comparisons with the 1D Haar transform—that discussion is not repeated herein, although we do follow its nomenclature and development.) We consider the 2D WPT spectrum in the context of the associated 2D filters. At the first level of decomposition the 16 subbands may be obtained by (block) convolution of the image $f(m, n)$ with filters $h_{k,l}(m, n)$. The 4×4 matrix $h_{k,l}$ is defined to be the outer product $h_k^t h_l$ of the filters in (5.2). Equivalently, these filters may be defined by 2D spatial exponentials of the form $\mathrm{e}^{i(2\pi/N)um}\mathrm{e}^{i(2\pi/N)vn}$, for $m, n = 0, 1, 2, 3$, where u, v are the spatial frequency parameters in cycles per pixel (cpp). Here u, v assume values 0, $(N/4)$, $(N/2)$, and $(3N/4)$.

As with the 1D WPT for an image with real entries, the "off-diagonal" spectral decomposition blocks at each level occur in complex conjugate pairs (located symmetrically across the main diagonal in the 1D quadtree spectrum); this is a reflection of the fact that the corresponding convolution filters are complex conjugate pairs. Likewise for the 2D separable WPT the nonreal convolution filters occur in conjugate pairs, and consequently for real images at level one of the 2D WPT, convolutions with only nine of the 16 filters $h_{k,l}$ need be computed.

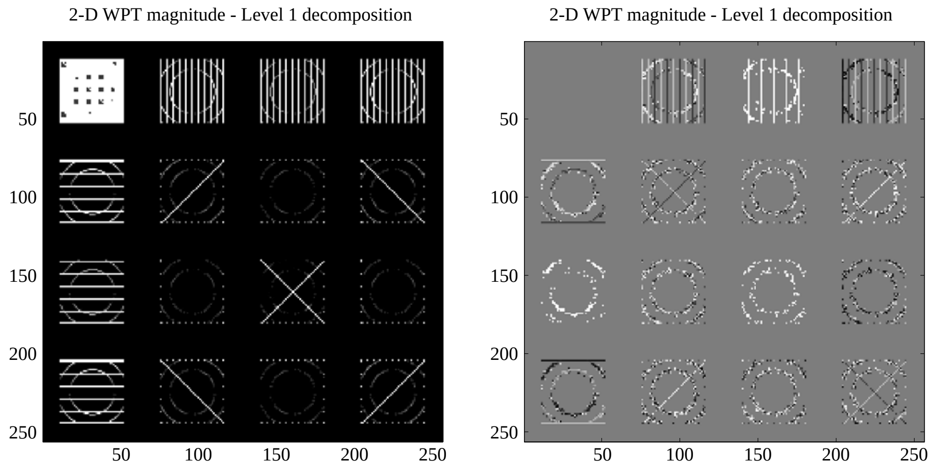

Fig. 4. 2D rectilinear WPT magnitude and phase.

With this in mind we consider the effect of the filters on the spectrum, for particular input signals. Filters $h_{1,2}(m,n)$, $h_{1,3}(m,n)$ and $h_{1,4}(m,n)$ are sensitive to frequencies $e^{i(2\pi/N)um}e^{i(2\pi/N)vn}$, where $u = 0$ and $v = (N/4)$, $(N/2)$, $(3N/4)$ respectively; that is, frequencies with only a horizontal variation. Hence these correspond to edges that run vertically or have a vertical component. In an analogous fashion, filters $h_{2,1}(m,n)$, $h_{3,1}(m,n)$ and $h_{4,1}(m,n)$ are sensitive to signals with a vertical variations: $v = 0$ and $u = (N/4)$, $(N/2)$, $(3N/4)$, respectively, and therefore to edges that run horizontally or those that have horizontal components. For such signals which constitute part of the basis set for the 2D 4×4 DFT, and which have a zero average value, the 2D WPT spectrum consists of a constant value 16 in the corresponding spectral block and zero elsewhere. (Equivalently, for such signals defined only on a 4×4 grid, the spectrum would be one point of value 16 in the corresponding spectral block.) In the top row of Fig. 4 it is possible to identify image edges that have components along vertical edges with frequencies defined by $e^{i(2\pi/N)0m}e^{j(2\pi/N)vn}$, where $v = (N/4)$, $(N/2)$, $(3N/4)$. Edges with components along horizontal edges with frequencies defined by $e^{i(2\pi/N)um}e^{i(2\pi/N)0n}$, where $u = (N/4)$, $(N/2)$, $(3N/4)$ appear in the lefthand column. Clearly, horizontal and vertical spatial exponential signals defined above, with a specific phase factor result in a spectrum of value 16 with the same phase factor. The interpretation in the context of edges is that a nonzero phase angle reflects a shifted edge, and the sign of the phase angle reflects whether the edge is increasing or decreasing. For the output of real filters $h_{1,3}(m,n)$ and $h_{3,1}(m,n)$, the phase spectrum is either 0 or π. For horizontal and vertical spatial exponentials with frequencies other than $(\pi/2)$, π, $(3\pi/4)$, the spectrum consists of components only with horizontal or vertical frequencies $(\pi/2)$, π, $(3\pi/4)$ respectively. This is due to the orthogonality of the 2D exponential signals.

Filters $h_{k,l}(m,n)$, for $k, l = 2, 3, 4$ are responsive to exponentials with frequencies in both horizontal and vertical directions, that is $u, v = 0$, $(N/4)$, $(N/2)$ and $(3N/4)$, and hence to diagonal edges. For such signals the spectrum has a value 16 in the corresponding signal block. For signals (diagonal edges) with other horizontal and vertical non-zero frequencies, the spectrum will typically have components in all 16 subbands.

Finally, for an image, the phase of the 2D WPT spectrum has a different interpretation than that for the 1D WPT. In the latter transform the phase angle of the second harmonic component—obtained by convolution with filter h_2—represents the angle of a gradient vector calculated with respect to an axis rotated $-\pi/4$ with respect to the x-axis. Hence it measures the angle of the corresponding edge on a 2×2 grid. In the 2D WPT, a zero phase angle represents an edge orthogonal to the 2D spatial exponentials $e^{i(2\pi/N)um}e^{i(2\pi/N)vn}$ where $u, v = 0, (N/4), (N/2), (3N/4)$. A nonzero phase angle still represents an edge orthogonal to the same exponential, except that the exponential now has the added phase. That is, for a given edge in a 4×4 block, the 15 associated spectral points give the decomposition of the edge in 15 directions (zero phase). The phase portion of the spectra defines the shift of the edges from the 15 directions.

6. Future directions

We conclude with some remarks on promising avenues for further investigation and applications. First of all, the theory developed herein is predicated on the action of an underlying group G on the spatial domain $L(X)$ of signals. By the basic harmonic analysis of this group action there is a group-based convolution given by the formula

$$(f \star h)(x) = \frac{|X|}{|G|} \sum_{\beta \in G} f(\beta a) h(\beta^{-1} x), \qquad \text{for all } x \in X,$$

where $f, h \in L(X)$ and a is a fixed but arbitrary base point in X (see Section II of Mirchandani et al., 2000). Convolution in turn gives rise to families of group-invariant filters $f \mapsto f \star h$. Filters are elemental to most signal processing applications, such as pattern recognition, data compression, and noise reduction.

In the particular case of the 1D WPC group action on $L(X)$, group-based convolution is explored in some detail in (ibid.). In particular, efficient algorithms are given for computing 1D WPC convolution, examples are computed, and applications are explored. Similar formulas for the 2D WPT may easily be derived from the 1D formulas. Indeed, since the 1D representations afforded by the action of WPC groups on trees are multiplicity-free (no repeated irreducible components in the spectral decomposition), it follows easily that the tensor product of two such representations is also multiplicity-free. Hence the 2D decomposable WPC representations on $L(X_1 \times X_2)$ are multiplicity free. The 2D WPC convolution of two 2D signals f and h may be efficiently computed as follows: first compute their 2D group-based Fourier transforms, $\widehat{f}$ and $\widehat{g}$, "multiply" these spectra in the spectral domain, and then take the inverse group-based Fourier transform to render the convolution product in the spatial domain (delta function) basis. Because the 2D WPC representation is multiplicity-free, this "multiplication" in the spectral domain reduces to a set of scalar-matrix multiplications, one for each irreducible component.

A promising line of research would therefore be to investigate the properties of 2D convolution for application to various classical image processing problems, particularly for edge detection and multiresolution similarity detection. Comparisons could be made between the existing 1D WPT results (which use the quadtree scanning method for $2^n \times 2^n$ images), and both the rectilinear and quadtree 2D separable WPT.

Since trees are 1D (thick) affine buildings, another promising area for exploring algebraic approaches to multi-dimensional signal processing may be to seek ways of extending the WPC group analysis to higher rank buildings. A first step would be to attach a simplicial complex geometry to selected subsets of the sample nodes X, and use this to provide a multiresolution filtration of $L(X)$. One might begin by considering truncated affine buildings or buildings of spherical type acted on by various finite groups. Since the family of finite groups acting on buildings of spherical type includes all the Chevalley groups and their twisted versions, these considerations are aimed toward developing a rich pool of possibilities from which to draw. Computational effectiveness will play an important role in dictating the geometries and groups that are appropriate for signal processing.

Acknowledgements

R Foote's research was partially supported by AFOSR, G Mirchandani's research was partially supported by AFOSR and NASA\JOVE, and D Rockmore's research was partially supported by AFOSR and NSF.

References

An, M., Cooley, J.W., Tolimieri, R., 1990. Factorization method for crystallographic Fourier transforms. Adv. Appl. Math. 11 (3), 358–371.

Auslander, L., Johnson, J.R., Johnson, R.W., 1996. Multidimensional Cooley–Tukey algorithms revisited. Adv. Appl. Math. 17 (4), 477–519.

Auslander, L., Johnson, R.W., Vulis, M., 1988. Evaluating finite Fourier transforms that respect group symmetries. Acta Cryst. A 44 (4), 467–478.

Conway, J., Curtis, R., Norton, S., Parker, R., Wilson, R., 1985. Atlas of Finite Groups. Clarendon Press, Oxford.

Curtis, C., Reiner, I., 1981. Methods of Representation Theory with Applications to Finite Groups and Orders, I. John Wiley & Sons, New York.

Dummit, D., Foote, R., 1999. Abstract Algebra, second ed. John Wiley & Sons, New York.

Foote, R., Mirchandani, G., Rockmore, D., Healy, D., Olson, T., 2000. A wreath product group approach to signal and image processing: part I—multiresolution analysis. IEEE Trans. Signal Process. 48 (1), 102–132.

Gonzalez, R.E., Woods, R.C., 1993. Digital Image Processing. Addison-Wesley, Reading, MA.

Jain, A., 1989. Fundamentals of Digital Image Processing. Prentice-Hall, New Jersey.

Mirchandani, G., Foote, R., Rockmore, D., Healy, D., Olson, T., 2000. A wreath product group approach to signal and image processing: part II—convolution, correlation, and applications. IEEE Trans. Signal Process. 48 (3), 749–767.

Moura, J., Johnson, J.R., Püschel, M., Haentjens, G., Sepiashvili, D., 2001. In search of optimal implementations for signal processing transforms. In: O'Sullivan, J.A. (Ed.), Statistical Methods in Imaging: In Medicine, Optics, and Communication. Springer-Verlag, NY. See also the SPIRAL (Signal Processing algorithms Implementation Research for Adaptable Libraries). Available from http://www.ece.cmu.edu/~spiral.

Pitsianis, N.P., 1997. The Kronecker product in optimization and fast transform generation. Ph.D. Thesis. Department of Computer Science, Cornell University.

Pratt, W., 1991. Digital Image Processing. John Wiley & Sons, New York.

Scheffé, H., 1999. The Analysis of Variance. A Wiley Publication in Mathematical Statistics, John Wiley & Sons, New York (reprint of the 1959 original. Wiley Classics Library).

Strang, G., Nguyen, T., 1996. Wavelets and Filter Banks. Wellesley-Cambridge Press, Mass.

Vetterli, M., Kovacevic, J., 1995. Wavelets and Subband Coding. Prentice-Hall, New Jersey.