Orthogonal Sub-band Image Transforms

by

Eero Peter Simoncelli

B.A. Physics, Harvard University (1984)

Submitted to the
Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of

Master of Science

May 1988

Author

Eero P. Simoncelli
April 29, 1988

Certified by

Edward H. Adelson
Associate Professor of Vision Science

Accepted by

Arthur C. Smith
Chairman, Department Committe on Graduate Students

Orthogonal Sub-band Image Transforms

by

Eero Peter Simoncelli

B.A. Physics, Harvard University (1984)

Submitted to the
Massachusetts Institute of Technology
Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degree of

Master of Science

May 1988

Thesis supervisor: Dr. Edward H. Adelson
Title: Associate Professor of Vision Science

Abstract

This paper proposes a class of linear transformations which are particularly well suited for image processing tasks such as data compression, progressive transmission, and machine vision. The basis functions of these transformations form a complete orthogonal set and are localized in both the spatial and spatial frequency domains. In addition, they may be implemented efficiently using cascaded convolutions with relatively small filters. Formulation of the problem is discussed in both the spatial frequency and spatial domains. Frequency domain formulation allows the isolation of aliasing errors and simple analysis of cascaded systems. Spatial domain formulation simplifies the problem of transform inversion, and provides a more obvious interpretation of the issues involved in filter design. Two simple design methods are proposed: a general spatial domain technique which is easily extended to multiple dimensions, and a frequency domain technique for the design of one-dimensional transforms. Examples of data compression and progressive transmission are given, and the extension of the results to two and three dimensions with arbitrary sampling geometries is discussed.

Acknowledgements

I would like to thank the vision group at SRI (formerly RCA) Sarnoff Research Laboratories, where I began this work as a summer employee. I would also like to thank the Advanced Television Research Group at the MIT Media Laboratory for providing digitized images and a photographic output device. Thanks to all the members of the Vision Sciences Group for their patience and support, and most of all to my advisor and friend Ted Adelson, whose intuition never ceases to amaze me.

Contents

Introduction

The task of vision may be viewed as the transformation of an image from an array of intensity values into a more suitable representation [1,2]. The same statement may be made concerning many image processing applications, particularly data compression. According to this viewpoint, the central issue is the choice of a representation. Some representations are preferable to others due to the fact that "... any particular representation makes certain information explicit at the expense of information that is pushed into the background and may be quite hard to recover"[1]. A common approach is to narrow the set of choices to the class of *linear* transformations, which are easily characterized and well understood. The rapid growth of the field of digital image processing is due in part to the usefuleness of linear transformations in analyzing signals and converting them into more desirable representational forms. This paper is concerned with the choice of linear transformations for efficient image representation.

One property that is widely accepted as being a necessity for effective image processing is an explicit representation of *scale* [1,3,2,4,5,6]. The argument is usually presented that images contain objects and features of many different sizes which may be viewed over a large range of distances, and therefore, a transformation should analyze the image simultaneously at different scales. To further support this notion, Pentland [7] has shown that many natural textures may be modeled with fractals.

As an equivalent description of the notion of scale, we may refer to sub-bands in the frequency domain representation of the image. A localized region

in the frequency plane corresponds to information in the image which occurs at a particular scale and orientation. Thus, the basis functions of a useful transformation should partition the input signal into localized patches in the frequency domain. In addition to localization in frequency, it is advantageous for the basis functions to be spatially localized. The necessity of spatial localization is particularly apparent in machine vision systems, where information about the location of features in the image is critical.

These properties may also be motivated by analogy with biological systems. It is commonly known that an important characteristic of the simple cells found in the mammalian visual cortex is a response that is selective for frequency and orientation [8,9,10,11]. Furthermore, each of these cells processes the information contained in a limited region of the entire visual field, so that an array of such cells may be considered to perform an operation analogous to convolution. Thus, the early stages of the visual system use a representation which is localized in space as well as in spatial frequency.

This notion of joint localization may be contrasted with the two most common representations used for the analysis of linear systems: the sampled or *discrete* signal, and its Fourier transform. For each of these representations (and for any other linear transform representation), we may write the sampled signal as a weighted sum of *basis functions.* The standard basis functions for discrete signals consist of impulses located at each sample location and are thus maximally localized in space, but convey no information about scale. The Fourier basis set is composed of even and odd phase sinusoidal sequences, whose usefulness is primarily due to the fact that they are the eigenfunctions of the class

of linear shift-invariant systems. Although they are maximally localized in the frequency domain, they are not localized in space. It is clear that representation in the space or frequency domains is extremely useful for purposes of system analysis, but this does *not* imply that impulses or sinusoids are the best way to represent signal information. In a number of recent papers [8,12,13,14,15,16,17], the importance of this issue is addressed and related to a 1946 paper by Dennis Gabor [18], who showed that the class of linear transformations may be considered to span a range of joint localization with the impulse basis set and the Fourier basis set at the two extemes, and that one-dimensional signals can be represented in terms of basis functions which are localized both in time *and* frequency.

Traditionally, transformations used in signal processing are broken into two categories: orthogonal transforms and sub-band transforms. This distinction is especially obvious in the area of image coding [19,20], and is due in part to the nature of the computational methods used for the two types of transform.

Orthogonal Transforms

The phrase "orthogonal transform" is usually used to refer to transforms such as the discrete Fourier transform, the discrete cosine transform, and the Karhunen-Loeve transform. Two advantages of using an orthogonal transform are the simplicity of transform inversion and the isolation of different types of information in different coefficients, which provides a more obvious interpretation of the meaning of the transform coefficients [19,21]. In addition, many of these transforms have highly efficient implementations involving cascades of

"butterfly" computations [22,23]. Based on the previous discussion, an obvious disadvantage is the nearly complete loss of positional information.

Sub-band Transforms

Sub-band transforms are generally computed by convolving the input signal with bandpass filters and decimating the resulting images. Each sub-band image contains information occuring at a particular scale. The most obvious sub-band system incorporates ideal or "brick-wall" bandpass filters to avoid the occurance of aliasing when the filtered images are subsampled. But filters with sharp transitions produce ringing (Gibbs phenomenon) in the spatial domain which is perceptually undesirable.

The distinction between orthogonal and sub-band transforms is an arbitrary one. Consider, for example the block DCT, in which the signal (image) is divided into non-overlapping blocks and a discrete cosine transform is computed for each block. Several of the basis functions for this transform are depicted in figure 1. The basis functions are clearly orthogonal, since the DCT is orthogonal and the blocks are chosen so that they do not overlap. In addition, we can view this as a sub-band transform. Computing a DCT on non-overlapping blocks is equivalent to convolving the image with each of the block DCT basis functions and then subsampling by a factor equal to the block size. The Fourier transform of the basis functions (also shown in figure 1) indicates that this is a sub-band transform, although it is clear that the sub-band localization is rather poor: the sharp discontinuities at the block boundaries lead to ringing in the frequency domain.

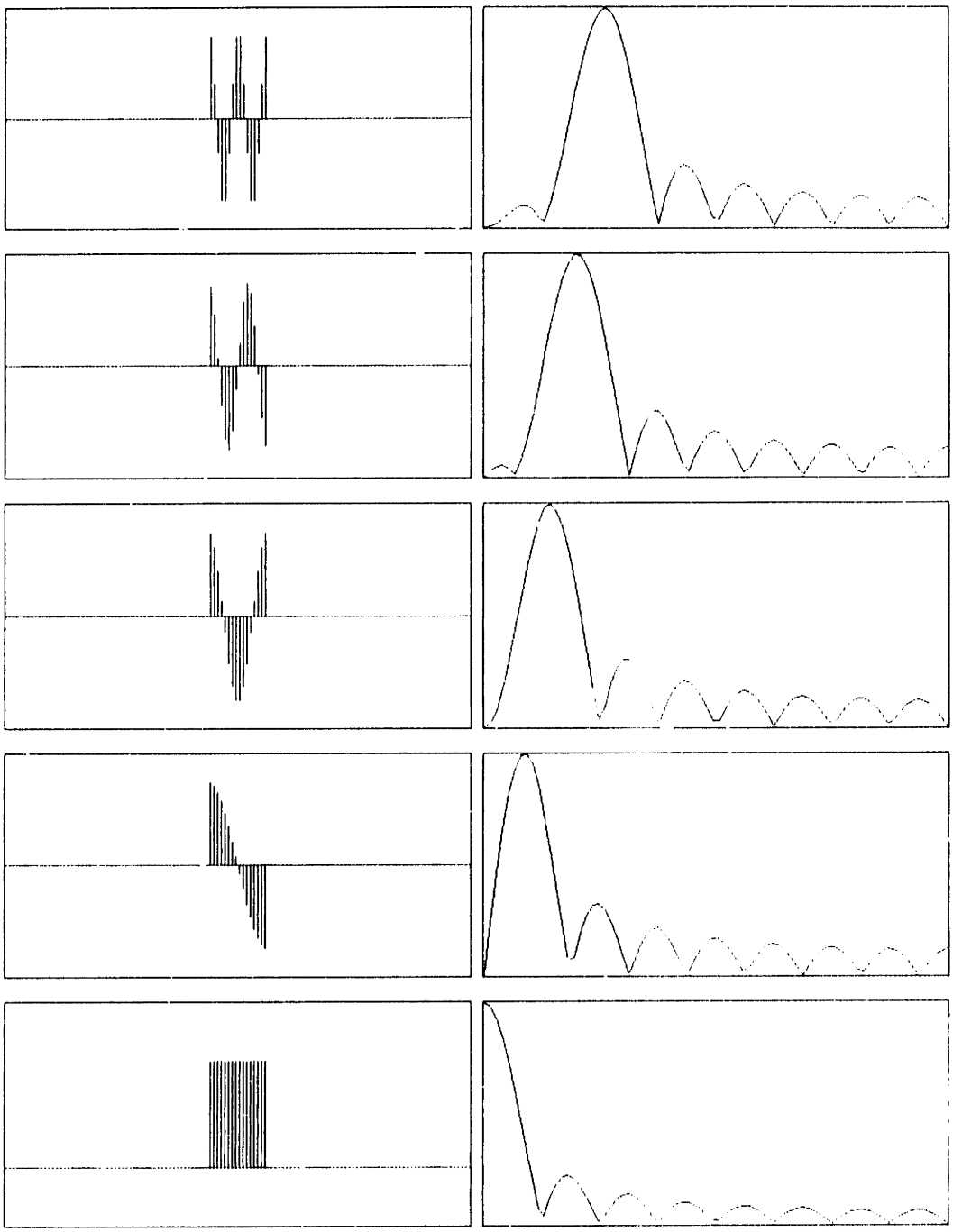


Figure 1: Several of the 16-point DCT basis functions and their DTFTs, plotted with a linear scale.

Smoother lowpass filters have been employed by several authors. Burt et al. [24,25,2] built a multi-scale representation from small gaussian-like filters which they termed a *pyramid*. The basis functions of this representation are jointly localized in space and spatial frequency, the computation of the representation is efficient, and the transformation is easily inverted, despite the non-orthogonality of the basis functions. The cascaded (pyramid) construction method is of particular interest and will play an important role in the transforms discussed in this paper. The primary disadvantages of this transform are that the basis functions are not oriented and the representation is overcomplete, requiring an increase by a factor of $\frac{4}{3}$ in the number of sample points over the original image.

Another recent development in the area of sub-band transforms is that of quadrature mirror filters (QMF). These were developed by Croiser, Esteban and Galand in [26,27] for sub-band coding of speech signals. They found that it was possible to choose non-ideal FIR bandpass filters and still avoid aliasing in the overall system output. These filters provide a central example for the type of transformations advocated in this paper, and will be discussed in section 2.

The purpose of the present paper is to describe orthogonal linear transformations which explicitly represent information at different scales and which satisfy the criterion of joint localization. It is divided into five major sections. Section 1 introduces the matrix and frequency domain notations which are used throughout the paper. Since the design constraints involve both the spatial and the spatial frequency domains, we will describe the problem and results in both domains. Section 2 discusses two-band one-dimensional systems, which

includes the class of QMF. Section 3 discusses filter design. We restrict the filter design problem to FIR filters, thus avoiding complicated multi-dimensional stability issues. A general spatial domain design method is described, along with a one-dimensional frequency domain method. In section 4, some examples of data compression and progressive transmission are presented. Finally, section 5 discusses the extension of subsampling and filter design issues to two- and three-dimensional systems.

1 Linear Transformations on Finite Images

The results presented in this paper are based on analysis in both the signal and the frequency domains, and thus rely on two separate notational frameworks: the standard matrix notation used in linear algebra, and the frequency domain representations commonly used in digital signal processing. In this section, we describe the two types of notation and make explicit the connection between them.

Analysis/Synthesis Filter Bank Formulation

As mentioned in the introduction, we are interested in linear transformations on images of a *finite* size which may be expressed in terms of convolutions with finite impulse response (FIR) filters. The schematic diagram in figure 2 depicts a convolution-based system known as an analysis/synthesis (A/S) filter bank [28]. The notation in the diagram is standard for digitial signal processing [22], except that for the purposes of this paper, the boxes $\boxed{F_i(\omega)}$ indicate *circular* convolution of a finite input image of size N with a filter with impulse response $f_i[n]$ and discrete time Fourier transform (DTFT)

$$F_i(\omega) = \sum_n f_i[n] e^{-j\omega n}$$

We do not place a causality constraint on the filter impulse responses, since they are meant for application to images. We do, however, assume that the length (region of support) of the filter is less than the image size. The boxes $\boxed{k_i \downarrow}$ indicate that the sequence is subsampled by a factor of k_i where k_i is an integer for all i. The boxes $\boxed{k_i \uparrow}$ indicate that the sequence should be upsampled by

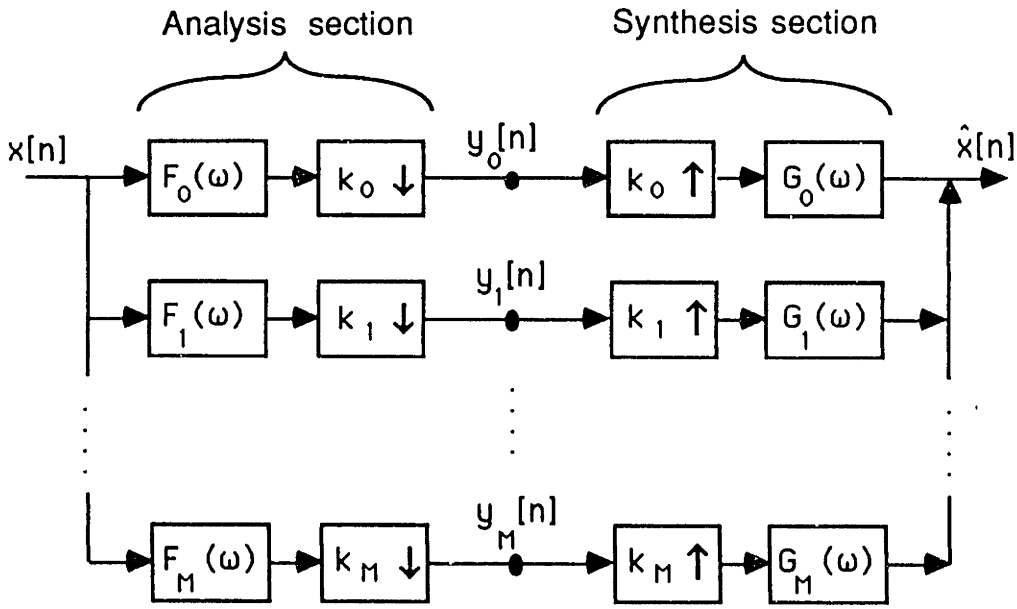


Figure 2: An analysis/synthesis filter bank.

inserting $k_i - 1$ zeros between each sample. For most of this paper, we will assume that $k_i = k$ for each i and that the integer k divides N evenly.

The analysis section of the A/S system takes an input sequence $x[n]$ of length N and performs a linear transformation to decompose it into M signals $y_i[n]$ of length N/k. As we will show shortly, such an analysis system is general enough to represent *any* linear transformation. The synthesis section corresponds to the inverse of the analysis transformation. Here the M signals $y_i[n]$ are upsampled and, after filtering with filters $g_i[n]$, are combined additively to give an approximation $\hat{x}[n]$ to the original signal. Note that although one-dimensional signals are indicated in the diagram, the system is equally valid for multi-dimensional signals if we replace occurances of scalars n, ω, k_i with vectors $\boldsymbol{n}$, $\boldsymbol{\omega}$, and matrix $\mathbf{K}_i$, respectively. The multi-dimensional problem will be discussed in section 5.

The use of the A/S formulation emphasizes the computation of the trans-

form coefficients through convolution. As we mentioned in the introduction, this is desirable since we would like to treat different regions of the image in the same manner. Furthermore, the expression of the problem in the frequency domain allows us to easily separate the error $e[n] = \hat{x}[n] - x[n]$ into two parts: an aliasing component and a shift-invariant component. To see this, we may write the contents of the intermediate signals $y_i[n]$ in the frequency domain as

$$Y_i(\omega) = \frac{1}{k} \sum_{j=0}^{k-1} F_i\left(\frac{\omega}{k} + \frac{2\pi j}{k}\right) X\left(\frac{\omega}{k} + \frac{2\pi j}{k}\right)$$

and the A/S system output is

$$\hat{X}(\omega) = \sum_{i=0}^{M-1} Y_i(k\omega) G_i(\omega)$$

where we have used well-known facts about the effects of upsampling and downsampling in the frequency domain [22]. Combining the two gives

$$\begin{aligned} \hat{X}(\omega) &= \frac{1}{k} \sum_{i=0}^{M-1} \left[\sum_{j=0}^{k-1} F_i\left(\omega + \frac{2\pi j}{k}\right) X\left(\omega + \frac{2\pi j}{k}\right) \right] G_i(\omega) \\ &= \frac{1}{k} \sum_{i=0}^{M-1} F_i(\omega) G_i(\omega) X(\omega) \\ &\quad + \frac{1}{k} \sum_{j=1}^{k-1} X\left(\omega + \frac{2\pi j}{k}\right) \sum_{i=0}^{M-1} F_i\left(\omega + \frac{2\pi j}{k}\right) G_i(\omega) \end{aligned} \tag{1}$$

The first term corresponds to a linear shift-invariant system response, and the second contains the system aliasing.

Cascaded Systems

A further advantage of the A/S system is that it allows explicit depiction and analysis of hierarchically constructed transformations. If we assume that we are dealing with A/S systems with perfect response (that is, $\hat{x}[n] = x[n]$), then

an intermediate signal $y_i[n]$ of one A/S system may be further decomposed by application of any other A/S system. To make this notion more precise, an example is given in the diagram of figure 3 in which an A/S system has been re-applied to its own intermediate signal $y_0[n]$. If the original A/S system (as shown in figure 2) had a perfect response then it is clear that the two-stage system shown in figure 3 will also have a perfect response. If the cascading is applied to each of the M intermediate signals $y_i[n]$, we will call the system a *uniform cascade* system. Otnerwise, it will be termed a *non-uniform* or *pyramid* cascade.

Matrix Formulation

An alternative to the frequency domain notation associated with the A/S filter bank is the matrix notation of linear algebra. An image of finite extent which has been sampled on a discrete lattice may be written as a finite length column vector $\boldsymbol{x}$ which corresponds to a point in $\mathbf{R}^N$, the set of all real N-tuples. The value of each component of $\boldsymbol{x}$ is simply the corresponding sample value in the image. Multi-dimensional images are converted to this vector format by imposing an arbitrary but fixed order on the lattice positions. If we let N be the length of the vector $\boldsymbol{x}$, a linear transformation on the image corresponds to multiplication by some matrix $\mathbf{M}$ with N columns.

Since the analysis and synthesis stages of the system in figure 2 each represent a linear transformation, it is clear that we may represent the same transformations using matrix notation. From the definition of convolution, and assuming a one-dimensional system for simplicity, we may write

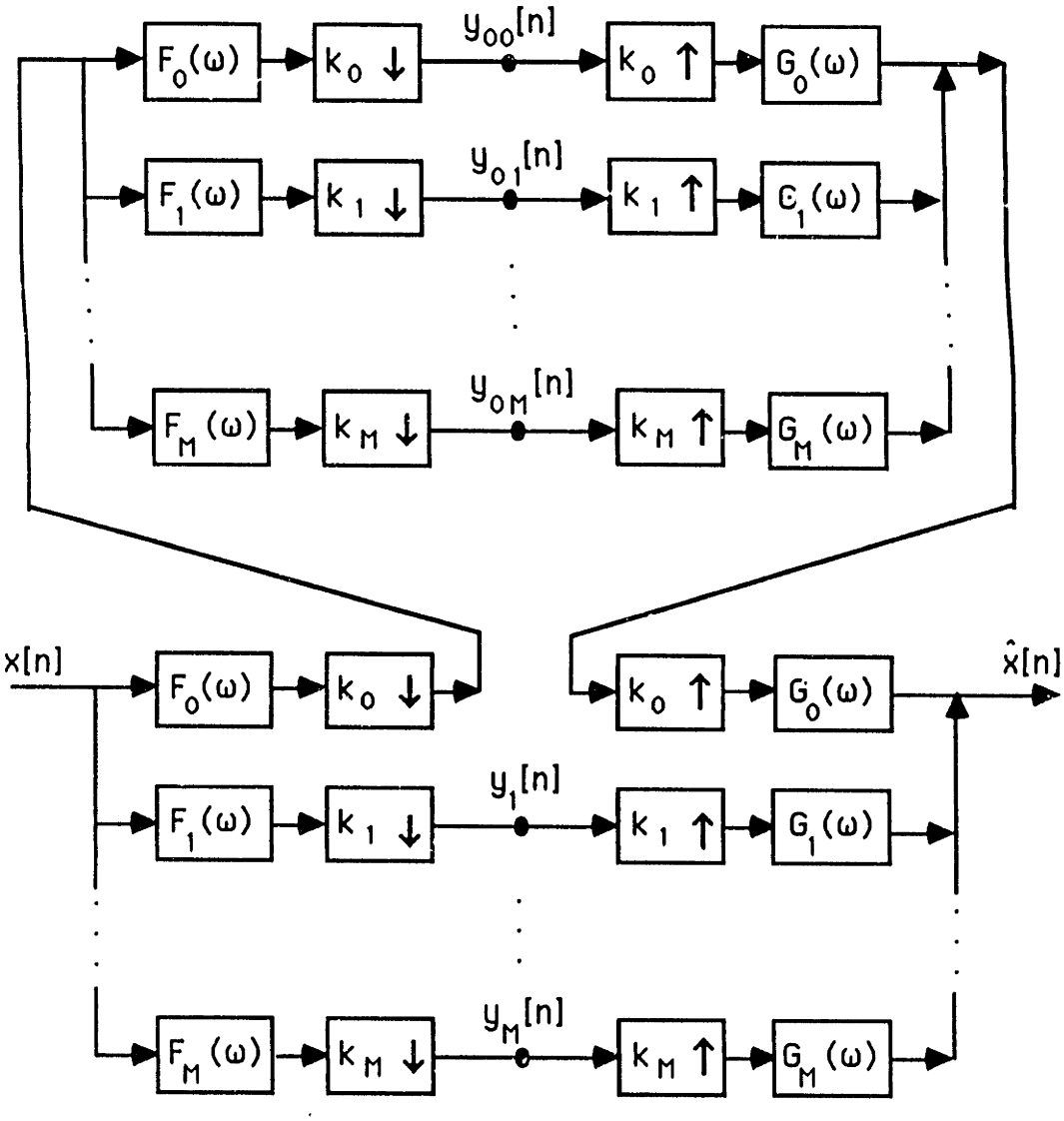


Figure 3: A non-uniformly cascaded analysis/synthesis filter bank.

$$y_i[m] = \sum_{l=0}^{N-1} x[l] f_i[k_i m - l]$$

and

$$\hat{x}[n] = \sum_{i=0}^{M-1} \sum_{m=0}^{\frac{N}{k_i}-1} y_i[m] g_i[n - k_i m]$$

where the filter sample locations $(k_i m - l)$ and $(n - k_i m)$ are computed modulo N. These expressions may be formulated as matrix-vector products

$$\boldsymbol{y} = \mathbf{F}^t \boldsymbol{x}$$

and

$$\hat{\boldsymbol{x}} = \mathbf{G}\boldsymbol{y}$$

or combining these two equations

$$\hat{\boldsymbol{x}} = \mathbf{G}\mathbf{F}^t \boldsymbol{x} \tag{2}$$

where $\boldsymbol{y}$ and $\hat{\boldsymbol{x}}$ are N-vectors, the superscript t indicates matrix transposition, and

$$\mathbf{F} = \begin{bmatrix} f_0[0] & f_0[k_0] & & f_1[0] & f_1[k_1] & \\ f_0[-1] & f_0[k_0-1] & & f_1[-1] & f_1[k_1-1] & \\ f_0[-2] & f_0[k_0-2] & & f_1[-2] & f_1[k_1-2] & \\ \vdots & f_0[k_0-3] & \cdots & \vdots & f_1[k_1-3] & \cdots \\ & f_0[k_0-4] & & & f_1[k_1-4] & \\ f_0[2] & \vdots & & f_1[2] & \vdots & \\ f_0[1] & & & f_1[1] & & \end{bmatrix} \tag{3}$$

and

$$\mathbf{G} = \begin{bmatrix} g_0[0] & g_0[k_0] & & g_1[0] & g_1[k_1] & \\ g_0[1] & g_0[k_0+1] & & g_1[1] & g_1[k_1+1] & \\ g_0[2] & g_0[k_0+2] & & g_1[2] & g_1[k_1+2] & \\ \vdots & g_0[k_0+3] & \cdots & \vdots & g_1[k_1+3] & \cdots \\ & g_0[k_0+4] & & & g_1[k_1+4] & \\ g_0[-2] & \vdots & & g_1[-2] & \vdots & \\ g_0[-1] & & & g_1[-1] & & \end{bmatrix} \tag{4}$$

It is worth noting that the columns of **G** are composed of copies of the filters $g_i[n]$ shifted by increments of k_i, but those of **F** are composed of copies of the *time-inverted* filters $f_i[-n]$ shifted by increments of k_i. Also, it is apparent that the order of the columns may be changed without changing the characteristics of the transformation, as long as we rearrange both **F** and **G** in the same way. The ordering indicated above, in which the amount of shift varies faster than the filter number as we cycle through the columns will be referred to as the *standard* form of the transformation matrix. For systems in which $k_i = k$ for each i, an alternative ordering of the columns will be useful. In this ordering, the filter number is varied faster than the amount of shift. For example, the matrix **F** in equation (3) would be written as

$$\mathbf{F} = \begin{bmatrix} f_0[0] & f_1[0] & f_2[0] & & f_0[k] & f_1[k] & \\ f_0[-1] & f_1[-1] & f_2[-1] & & f_0[k-1] & f_1[k-1] & \\ f_0[-2] & f_1[-2] & f_2[-2] & & f_0[k-2] & f_1[k-2] & \\ \vdots & \vdots & \vdots & \cdots & f_0[k-3] & f_1[k-3] & \cdots \\ f_0[2] & f_1[2] & f_2[2] & & f_0[k-4] & f_1[k-4] & \\ f_0[1] & f_1[1] & f_2[1] & & \vdots & \vdots & \end{bmatrix}$$

We will refer to this ordering as the *interdigitated* form of the transformation matrix.

From the above discussion, it is clear that we can express any A/S system in matrix form. The converse of this result is also true: there is an A/S system corresponding to the linear transformation and inverse transformation defined by any invertible matrix **M**. Given a transformation matrix **M** with l rows, we simply create an analysis filter bank with $k_i = N$ for each i, containing l different filters, each defined by a row of the matrix **M**.

Inverse Transforms

A primary advantage of the matrix notation is the ease with which a given transformation may be inverted. From equation (2), we see that in order for the A/S system to perfectly reconstruct the original signal $x[n]$, the corresponding matrices must obey

$$\mathbf{G}\mathbf{F}^t = \mathbf{I} \tag{5}$$

where **I** is the identity matrix. If **F** has rank N and is square, we may choose a synthesis matrix

$$\mathbf{G} = (\mathbf{F}^{-1})^t \tag{6}$$

which will also be square with rank N. Thus, transform inversion in the spatial domain is a conceptually simple procedure and we will find it useful in the analysis and design of A/S systems.

If the matrix **F** is of rank N but is *not* square (that is, the representation is overcomplete), we may always build a perfect reconstruction system by choosing

$\mathbf{G}$ to be the generalized inverse or pseudo-inverse [29] of $\mathbf{F}$:

$$\mathbf{G} = (\mathbf{F}\mathbf{F}^t)^{-1}\mathbf{F}$$

If $\mathbf{F}$ is square, this reduces to the solution given in equation (6). Similarly, if we start with a (possibly non-square) matrix $\mathbf{G}$ of rank N, we may choose $\mathbf{F} = (\mathbf{G}\mathbf{G}^t)^{-1}\mathbf{G}$.

In the square matrix case, we may use the commutative property of matrix inversion to switch the order of $\mathbf{G}$ and $\mathbf{F}^t$, which gives

$$\mathbf{F}^t\mathbf{G} = \mathbf{I}$$

This corresponds to switching the analysis and synthesis sections of the system, thus converting the A/S system to a *multiplexing* system. Such a system combines M input signals $x_i[n]$ to form a single signal $y[n]$. This signal may later be decomposed into approximations $\hat{x}_i[n]$ of the original signals. This is illustrated in figure 4. In this way, *all* of the results of this paper and all of the filter designs may be used for the purpose of multiplexing.

Orthogonal Transforms

As discussed in the introduction, this paper will be primarily concerned with *orthogonal* transformations. A matrix $\mathbf{A}$ corresponding to an orthogonal transformation is a square matrix with the property that

$$\mathbf{A}\mathbf{A}^t = \mathbf{A}^t\mathbf{A} = \mathbf{I} \tag{7}$$

This condition places a number of restrictions on the corresponding A/S system. Since the transformation matrix must be square, the number of samples in the

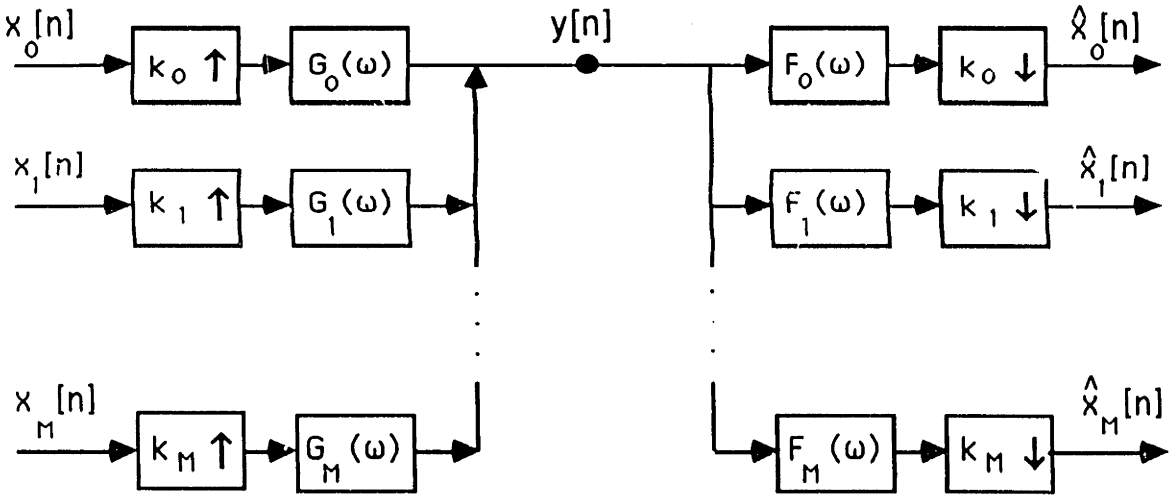


Figure 4: A multiplexing system.

transformed signal must be equal to N, the number of samples in the original image. For the A/S system, this means that

$$\sum_{i=0}^{M-1} \frac{1}{k_i} = 1$$

where we have assumed that N is divisible by all of the k_i. Such a system has been termed a *maximally decimated* filter bank [28].

A second, more important constraint is placed on the A/S system by orthogonality. Combining the perfect reconstruction requirement in (5) with the orthogonality constraint in (7) gives

$$\mathbf{G} = \mathbf{F}$$

If we consider the relationships between the A/S and matrix formulations given in (3) and (4), this means that the filters must obey

$$g_i[n] = f_i[-n], \qquad \text{for all } i \tag{8}$$

The synthesis filters should be *time-reversed* versions of the analysis filters.

Another property of orthogonal matrices will prove to be helpful in understanding A/S systems: the inner product of distinct columns of an orthogonal matrix must be zero. In terms of the filtering matrix $\mathbf{F}$, it is useful to express this condition in the form of two separate equations:

$$\sum_{n=0}^{N-1} f_i[n] f_i[n - mk_i] = 0, \qquad \text{for all } m \neq 0, \quad \text{for all } i \tag{9}$$

and

$$\sum_{n=0}^{N-1} f_i[n] f_j[n - mk_j] = 0, \qquad \text{for all } m, \quad \text{for all } i \neq j \tag{10}$$

where, as before, the filter sample locations are computed modulo N. Equation (9) states that the inner product of the impulse response of a filter with a copy of itself delayed by a multiple of k_i is zero. We will refer to this as the *self-orthogonality* condition. Equation (10) states that the dot product of the impulse responses of any two shifted distinct filters f_i and f_j is zero. Consider the columns of the matrix $\mathbf{F}$ containing shifted copies of the impulse response of one of the filters, say $f_i[n]$. We will refer to the subspace of $\mathbf{R}^N$ spanned by these columns as the *associated subspace* of the filter $f_i[n]$. Then equation (10) states that for an orthogonal transform, the M associated subspaces of the filters $f_i[n]$ are orthogonal subspaces of $\mathbf{R}^N$. In other words, the M subspaces, each spanned by the columns of $\mathbf{F}$ containing shifted copies of one of the filters f_i, are orthogonal. We will refer to this as the *mutual orthogonality* condition. These two equations provide necessary and sufficient conditions for the orthogonality of the matrix $\mathbf{F}$, and we will refer to them often in the next sections.

2 One-Dimensional, Two-band Systems

In order to elucidate the concept of FIR orthogonal filters we will first consider the splitting of a finite-length one-dimensional signal into two orthogonal sub-signals. For such a system, the subsampling factors are $k_0 = k_1 = 2$, and we will assume that the size of the image N is a power of 2. In addition to being the simplest to understand, this type of system lends itself easily to the construction of octave-band representations. An octave bandwidth partition of the frequency domain has ben advocated by several authors [1,6,4] for use in image analysis and machine vision. For illustrative purposes, consider the use of lowpass and highpass filters in the two-band A/S system. If we form a pyramid by applying the transform recursively to the lowpass sub-signal, the resultant sub-signals will correspond to a segmentation of the original into octave width frequency sub-bands, as illustrated in the idealized diagram in figure 5. Such one-dimensional systems yield surprisingly useful transforms which may be applied separably to multi-dimensional signals.

There is a substantial literature concerned with the construction of band-pass filter sets which allow signal reconstruction without aliasing. The history of these efforts has been reviewed in [28] and more recently in [30] and [31]. The initial concepts for two-band systems were established by Croiser, Esteban and Galand in [26,27] and the resulting set of filters were termed *quadrature mirror filters* (QMF). These authors were interested in sub-band coding of speech signals using filters of reasonable length. The authors also discuss the extension of these results to multiple band systems using uniform cascading techniques. Clearly, this only includes cases where the number of bands is a power of two.

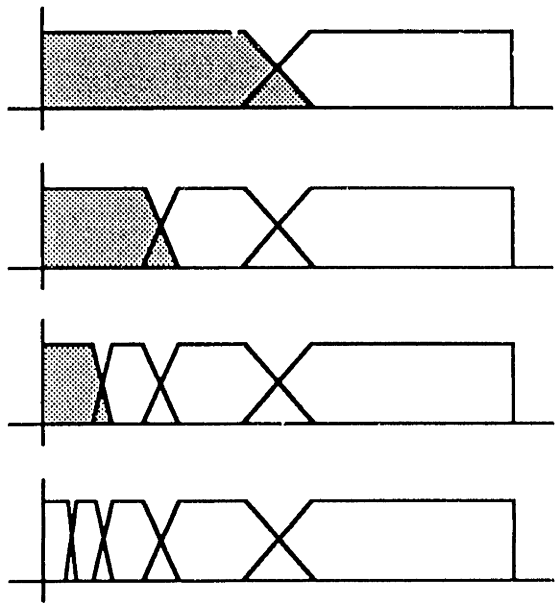


Figure 5: Octave band splitting produced by a four-level pyramid cascade of a two-band A/S system. The top picture represents the splitting of the two-band A/S system. Each succesive picture shows the effect of reapplying the system to the lowpass signal (indicated in grey) of the previous picture. The bottom picture gives the final four-level partition of the frequency domain. All frequency axes cover the range from 0 to π.

In addition, they suggested an efficient polyphase filter implementation. Several suggestions for more efficient realizations have been made by the same authors [32,33].

The issue of the general multiple-band problem was first discussed by Rothweiler [34], who gave a complex solution. This was made more precise by Galand and Nussbaumer in [35], where the case of odd-length symmentric filters was also presented using suitable delays in the A/S system. The complex solution was also discussed by Chu in [36]. Nussbaumer and Vetterli [37,38] discovered a real multiple-band solution which cancelled aliasing only in adjacent bands, thus relying on negligible aliasing between non-adjacent bands.

Smith and Barnwell [39] were the first to publish perfect reconstruction systems. A similar result was published by Wackersreuther in [40], which in-

cluded the multiple band case. Smith and Barnwell clarified their orignal work in [41] and an efficient implementation was found by Galand and Nussbaumer [42]. Analysis of the problem in terms of a matrix of Z-transform polynomials known as the *aliasing cancellation* or A/C matrix was independently proposed by Ramstad [43], Smith and Barnwell [44], and Vetterli [45].

Vetterli [45,46,47,31] used the A/C matrix notation to prove that there is no solution for the general multiple band case using FIR modulated filters, derived some multiple-band solutions, and showed that the multiplexing problem is closely related. Vaidyanathan [48,49] further clarified the multiple-band problem. For reference, several authors [50,51,52,53] have explored IIR solutions for the two-band case.

Some Simple Examples

The simplest example of a two-band orthogonal system is a matrix containing non-overlapping planar rotation submatrices along the diagonal:

$$\mathbf{F}(\theta) = \begin{bmatrix} c_\theta & s_\theta & 0 & 0 & \cdots & & 0 \\ -s_\theta & c_\theta & 0 & 0 & & & \\ 0 & 0 & c_\theta & s_\theta & & & \\ 0 & 0 & -s_\theta & c_\theta & & & \\ \vdots & & & & \ddots & & \\ & & & & & c_\theta & s_\theta \\ 0 & & & & & -s_\theta & c_\theta \end{bmatrix} \tag{11}$$

where $c_\theta = \cos(\theta)$ and $s_\theta = \sin(\theta)$ for some arbitrary θ. This transformation may also be viewed as a two-point DFT. The matrix is clearly orthogonal since

the individual rotation matrices are orthogonal and they have been placed in the matrix so as to act upon distinct pairs of points in the image. The arrangement of the columns in the matrix corresponds to the interdigitated ordering, and the impulse responses of the corresponding A/S filters may be read directly from the matrix in equation (11) using the relationships defined in equations (3) or (4) and will each contain two samples. Clearly, the transformation matrix $\mathbf{F}(0)$ corresponds to the identity transformation. In addition, both the Haar and the Walsh-Hadamard transformation may be obtained through pyramid and uniform cascading, respectively, of the A/S bank corresponding to $\mathbf{F}(\pi/4)$. In this case, the impulse responses of the filters are $[1, 1]$ and $[1, -1]$. An illustration of the construction of the Haar basis set in the spatial domain is given in figure 6. The construction of the Hadamard basis set is given in figure 7. The transformation corresponding to $\mathbf{F}(\pi/4)$ will be referred to as the *primitive Haar* transformation. Although this transformation is simple to implement and computationally efficient, its poor frequency response limits its usefulness.

A more interesting transformation can be formed by concatenating two planar rotation matrices: one like the one given above, the other shifted along the diagonal by one sample

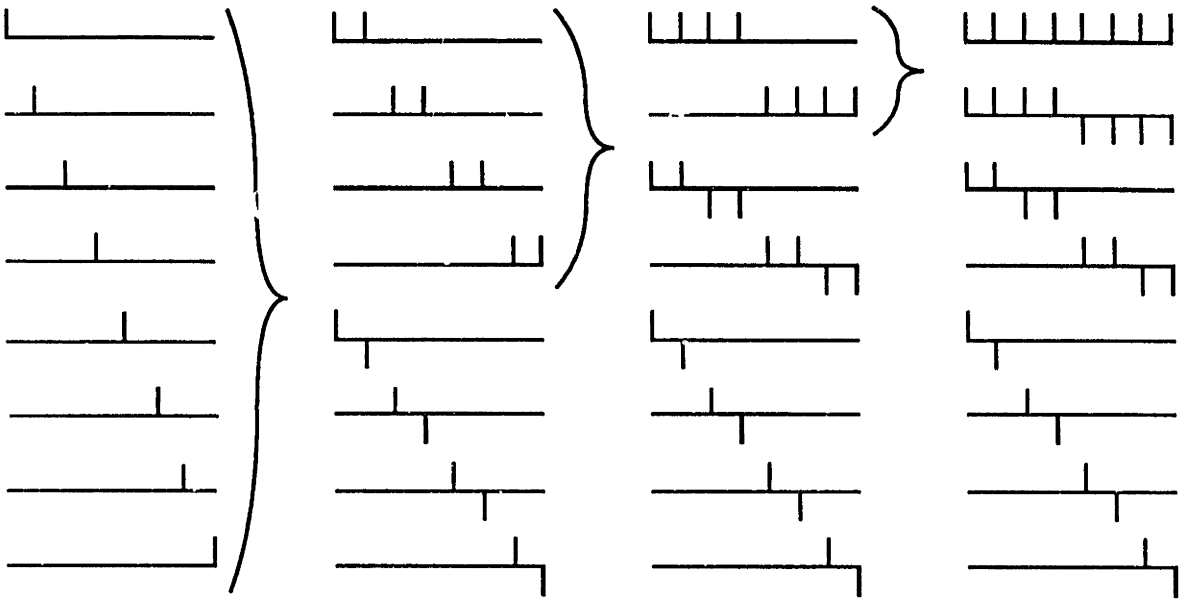


Figure 6: Pyramid construction of the eight-sample Haar basis set. On the left is the original impulse basis set. The next column indicates the primitive Haar set defined in equation (11) above. At each stage, the primitive transformation is applied to the lowpass basis functions.

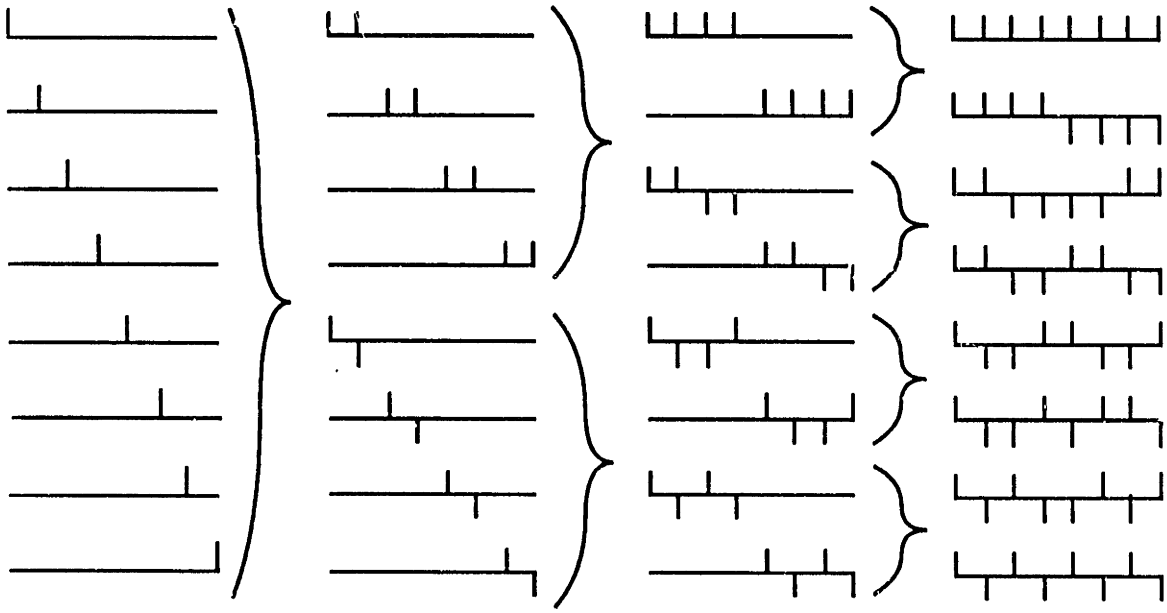


Figure 7: Uniform cascaded construction of the Hadamard basis set. See figure 6.

$$\mathbf{F}(\theta_0,\theta_1) = \begin{bmatrix} c_0 & s_0 & 0 & 0 & \cdots & & 0 \\ -s_0 & c_0 & 0 & 0 & & & \\ 0 & 0 & c_0 & s_0 & & & \\ 0 & 0 & -s_0 & c_0 & & & \\ \vdots & & & & \ddots & & \\ & & & & & c_0 & s_0 \\ 0 & & & & & -s_0 & c_0 \end{bmatrix} \cdot \begin{bmatrix} c_1 & 0 & 0 & 0 & 0 & \cdots & -s_1 \\ 0 & c_1 & s_1 & 0 & 0 & & 0 \\ 0 & -s_1 & c_1 & 0 & 0 & & \\ 0 & 0 & 0 & c_1 & s_1 & & \\ 0 & 0 & 0 & -s_1 & c_1 & & \\ \vdots & & & & & \ddots & \\ s_1 & & & & & & c_1 \end{bmatrix}$$

$$= \begin{bmatrix} c_0c_1 & s_0c_1 & s_0s_1 & 0 & 0 & \cdots \\ -s_0c_1 & c_0c_1 & c_0s_1 & 0 & 0 & \\ 0 & -c_0s_1 & c_0c_1 & s_0c_1 & s_0s_1 & \\ 0 & s_0s_1 & -s_0c_1 & c_0c_1 & c_0s_1 & \\ 0 & 0 & 0 & -c_0s_1 & c_0c_1 & \\ 0 & 0 & 0 & s_0s_1 & -s_0c_1 & \\ \vdots & & & & & \ddots \end{bmatrix} \tag{12}$$

where

$$c_0 = \cos(\theta_0), \quad s_0 = \sin(\theta_0), \quad c_1 = \cos(\theta_1), \quad \text{and} \quad s_1 = \sin(\theta_1).$$

The matrix $\mathbf{F}(\theta_0,\theta_1)$ is clearly orthogonal since the product of two orthogonal matrices is orthogonal. The construction of the transformation as a product of two planar rotation matrices indicates that the transformation could be implemented efficiently as a butterfly network.

As in the previous example, the matrix is in interdigitated form, and we may simply read off the two filters, each of length four, for use in the corresponding A/S system. This transformation class may be reduced to a one

parameter family by arbitrarily selecting one of the filters to be a highpass filter by constraining it to have a D.C. frequency response of zero. For example, one choice is

$$s_0c_1 + c_0c_1 - c_0s_1 + s_0s_1 = 0$$

which becomes, using standard trigonometric identities,

$$\sin(\theta_0 - \theta_1) + cos(\theta_0 - \theta_1) = 0$$

This equation has two solutions

$$\theta_0 - \theta_1 = \tfrac{3\pi}{4}, \qquad \text{or} \quad \theta_0 - \theta_1 = \tfrac{7\pi}{4}$$

The two solutions are the same up to an overall sign change. If we had chosen the other filter as a highpass, the roles of θ_0 and θ_1 would be reversed. An example with $\theta_0 = 0.167\pi$ and $\theta_1 = 1.917\pi$ is shown in figure 8. The values of θ_i were chosen to minimize aliasing energy in the lowpass A/S branch. As desired, the transform is both orthogonal and sub-band.

Although these filters are capable of better localization in the frequency domain than the those in (11), they are always asymmetric (except for degenerate cases such as $\theta_0 = 0$ or $\theta_1 = 0$ which are the same as the primitive Haar set defined in (11)) and thus do not have a linear phase response. One can continue to generate even-length filters by cascading additional rotation matrices, but for the purposes of this paper, we will take a more general approach in order to develop symmetric (linear phase) filters.

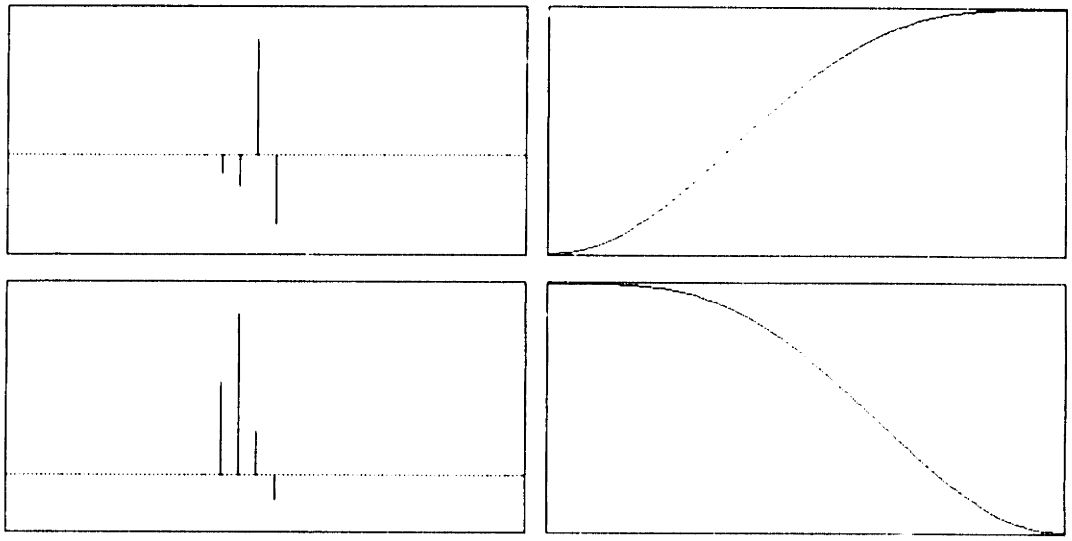


Figure 8: An example of a two-band A/S system corresponding to the matrix in equation (12). On the left are the low and highpass filter impulse response samples. On the right are the corresponding DTFT's, plotted from 0 to π on linear scale axes.

General Solutions

To investigate more general filter solutions for the two-band A/S system, we begin by writing an expression for $\hat{X}(\omega)$, the DTFT of the two-band A/S system output. The intermediate (transform) signals are

$$Y_0(\omega) = \frac{1}{2}\left[F_0(\tfrac{\omega}{2})X(\tfrac{\omega}{2}) + F_0(\tfrac{\omega}{2}+\pi)X(\tfrac{\omega}{2}+\pi)\right]$$

and

$$Y_1(\omega) = \frac{1}{2}\left[F_1(\tfrac{\omega}{2})X(\tfrac{\omega}{2}) + F_1(\tfrac{\omega}{2}+\pi)X(\tfrac{\omega}{2}+\pi)\right]$$

and the A/S system output is

$$\hat{X}(\omega) = Y_0(2\omega)G_0(\omega) + Y_1(2\omega)G_1(\omega)$$

Combining these equations gives

$$\hat{X}(\omega) = \frac{1}{2}\Big[F_0(\omega)G_0(\omega) + F_1(\omega)G_1(\omega)\Big]X(\omega) + \frac{1}{2}\Big[F_0(\omega+\pi)G_0(\omega) + F_1(\omega+\pi)G_1(\omega)\Big]X(\omega+\pi) \quad (13)$$

As mentioned in the previous section, expression of the problem in the frequency domain has allowed us to write the system response as a sum of a shift-invariant component and an aliasing component, thus isolating the aliasing errors in the second term.

Since we are seeking an orthogonal transform, we will require, as in (8), that

$$g_i[n] = f_i[-n], \qquad \text{for all } i$$

or, in the frequency domain

$$G_i(\omega) = F_i(-\omega)$$

By observing the relationship between the filters corresponding to the examples in (11) and in (12), we may impose a relationship between filters f_0 and f_1 which will eliminate the aliasing term in (13) and will ensure that the filters obey the mutual orthogonality condition given in equation (10). The choice is to define

$$f_1[n] = (-1)^{(n+1)} f_0[-1-n] \quad (14)$$

or equivalently,

$$\begin{aligned} F_1(\omega) &= \sum_n f_0[-1-n](-1)^{(n+1)} e^{-j\omega n} \\ &= e^{j\omega} \sum_m f_0[m] e^{j(\omega-\pi)m} \\ &= e^{j\omega} F_0(-\omega+\pi) \end{aligned} \quad (15)$$

This definition clearly includes the examples given in equations (11) and (12). In addition it includes as as sub-case the filter choices used by Esteban and Galand:

$$f_0[n] = h[n], \quad f_1[n] = (-1)^n h[n],$$
$$g_0[n] = h[n], \qquad \text{and} \quad g_1[n] = (-1)^{(n+1)} h[n]$$

for $h[n]$ an even-length symmetric filter.

It is straightforward to show that the aliasing term of equation (13) vanishes:

$$\begin{aligned}
&\tfrac{1}{2}\Big[F_0(\omega+\pi)G_0(\omega) + F_1(\omega+\pi)G_1(\omega)\Big]X(\omega+\pi) \\
&\qquad = \tfrac{1}{2}\Big[F_0(\omega+\pi)F_0(-\omega) + e^{j(\omega+\pi)}F_0(-\omega)e^{-j\omega}F_0(\omega+\pi)\Big]X(\omega+\pi) \\
&\qquad = \tfrac{1}{2}\Big[F_0(\omega+\pi)F_0(-\omega) - F_0(-\omega)F_0(\omega+\pi)\Big]X(\omega+\pi) \\
&\qquad = 0
\end{aligned}$$

Similarly, it is easy to see that the self-orthogonality condition of equation (10) holds for this choice of filters:

$$\begin{aligned}
\sum_n f_0[n] f_1[n-2m] &= \sum_n f_0[n](-1)^{(n-2m+1)} f_0[-1-n+2m] \\
&= \sum_{n \text{ odd}} f_0[n] f_0[-1-n+2m] - \sum_{n \text{ even}} f_0[n] f_0[-1-n+2m] \\
&= \sum_{n \text{ odd}} f_0[n] f_0[-1-n+2m] - \sum_{n' \text{ odd}} f_0[-1-n'+2m] f_0[n'] \\
&= 0, \qquad \text{for all } m
\end{aligned}$$

The factor $e^{j\omega}$ in equation (15) suggests a modification of the A/S system where the sampling on the two branches is staggered. That is, we keep the even samples

on one branch and the odd samples on the other. This is illustrated in figure 9. The box $\boxed{e^{-j\omega}}$ indicates a one sample delay.

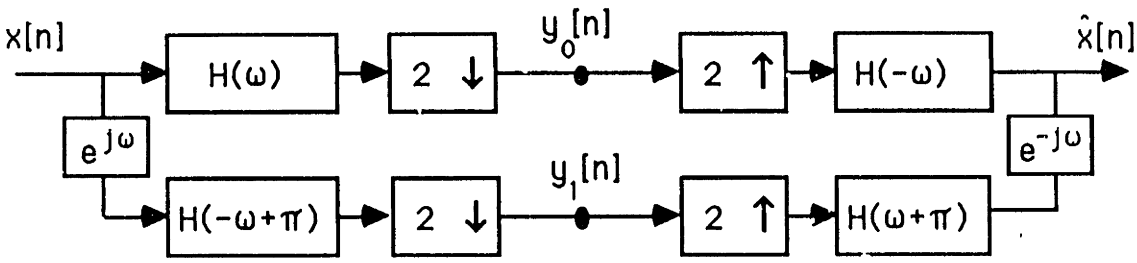


Figure 9: A staggered two-band analysis/synthesis filter bank.

If we let $F_0(\omega) = H(\omega)$, then from equations (14), (15), and (13), we see that the filter design problem is now reduced to design of a single filter $h[n]$ with DTFT obeying

$$H(\omega)H(-\omega) + H(-\omega+\pi)H(\omega+\pi) = 2$$

or

$$\left|H(\omega)\right|^2 + \left|H(\omega+\pi)\right|^2 = 2 \tag{16}$$

In the spatial domain, this requirement is

$$\sum_n h[n]h[n+2m] = \delta_{m0}, \qquad \text{for all } m \neq 0 \tag{17}$$

which is precisely the same as the self-orthogonality equation given in (9).

Symmetric Filters

We would like to design a *symmetric* filter satisfying equation (17). Symmetric filters have linear phase responses and in addition are more effiecient computationally, especially in systems where multiplication is a more expensive operation

than addition. A simple observation suffices to show that only trivial symmetric solutions (the identity transform and the primitive Haar transform) exist for filters of length $M < (\frac{N+4}{2})$ where N is the size of the image. Assume that the support of the filter $h[n]$ is the region $\{0 \leq n \leq M-1\}$, although the following argument holds for any region of support of length M. For $M > 1$ odd, consider the dot product of the first column of the standard form analysis matrix $\mathbf{F}$ with the $(\frac{M-1}{2})$th:

$$\begin{bmatrix} h[0] \\ 0 \\ 0 \\ \vdots \\ 0 \\ h[M-1] \\ h[M-2] \\ h[M-3] \\ \vdots \\ h[1] \end{bmatrix}^t \cdot \begin{bmatrix} h[M-1] \\ h[M-2] \\ h[M-3] \\ \vdots \\ h[1] \\ h[0] \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{18}$$

According to the self-orthogonality condition, this product must be zero. If $N \geq 2M-1$, then the product in (18) yields $h[0]h[M-1]$ which, by the symmetry of $h[n]$ is equal to $h[0]h[0]$. This is non-zero by our assumption concerning the region of support of the filter. If $N = 2M-2$ (remember we are assuming that N is a power of 2 for this section), then the product (18) yields twice this value, and is thus also non-zero. So if $M > 1$ and M is odd, we require that $N \leq 2M-4$ for the product (18) to be zero. For even-length symmetric

filters, the corresponding result is that $N \leq 2M - 1$. Vaidyanathan [54] has proven the corresponding result for one-dimensional time sequences which are infinite in length: there are *no* two-band symmetric FIR solutions except for the identity and the primitive Haar transform defined in (11). This fact is rather unfortunate. It means that non-trivial linear phase orthogonal filters must be approximately half the size of the image. In the next section, however, we will show that it is possible to design smaller filters which are very nearly orthogonal.

3 One-dimensional Filter Design

In this section, we will discuss the design of approximately orthogonal filters for use in A/S banks. The goal will be the determination of a simple design method that is readily extensible to A/S systems in multiple dimensions. The history of design methods for QMFs is somewhat sparse. The original papers of Croisier, Esteban and Galand [26,27] do not discuss design issues. Johnston [55] designed a set of widely used filters by minimizing an error function containing a shift-invariant error term and a weighted stopband ripple term for a fixed number of filter taps. Jain and Crochiere [56,57] used a similar error criterion in the time domain, and formulated an iterative design scheme which in which each iteration required the constrained minimization of a quadratic function. A technique for design of perfect reconstruction filter sets is given by Smith and Barnwell in [39]. They first design a *product filter* $F(\omega)$ which is factorable as

$$F(\omega) = \left|H(\omega)\right|^2$$

and which satisfies

$$f(n)\frac{\left(1+(-1)^n\right)}{2} = \delta(n)$$

The resulting $F(\omega)$ is factored to get h(n), the lowpass filter. Wackersreuther [40] independently arrived at an identical design method in the time domain. The problem with this design method is the somewhat arbitrary choice of the product filter.

In this paper, we present two design methods, each motivated by the fact that exactly orthogonal symmetric solutions exist only if the image size N satisfies $N \leq 2M-4$, where M is the filter size. The methods were not

developed as optimal design methods, but as useful research tools for exploring and understanding the class of orthogonal sub-band transforms. It has been found, however, that the resulting filters perform quite well in practice.

Spatial Domain Design

The first method is a spatial domain technique. Given an initial set of analysis filters for use in any two-band A/S system, we would like to orthogonalize the corresponding transformation matrix, retaining the form of the A/S system matrix. In other words, the columns of the resulting matrix should correspond to shifted versions of filters. Gram-Schmidt orthogonalization [29] is inappropriate here, since it orthogonalizes the matrix columns (or rows) in an arbitrary order, thus breaking the symmetry of the A/S system. We need instead a method which orthogonalizes the rows of the matrix in parallel. One possibility is that of iteratively averaging the analysis and synthesis matrices. Assuming that we start with a square full-rank analysis matrix $\mathbf{A}$, the appropriate synthesis matrix is given by equation (6), and one step of the iterative procedure is written

$$\mathbf{A}' \leftarrow \frac{1}{2}\left[\mathbf{A} + (\mathbf{A}^{-1})^t\right] \tag{19}$$

We emphasize again that we are assuming here that the initial matrix $\mathbf{A}$ is invertible. The result of this iterative procedure will clearly depend on the starting point. Several observations may be made concerning the properties of the iterative step.

Proposition 1 *If the initial matrix* $\mathbf{A}$ *corresponds to a particular A/S filter bank, the resulting matrix* $\mathbf{A}'$ *will be consistent with the same filter bank. In*

other words, if the matrix $\mathbf{A}$ *is formed from shifted versions of filter impulse responses, the matrix* $\mathbf{A}'$ *will have the same form.*

Proof: Assume the matrix $\mathbf{A}$ is written in interdigitated form. Then we can express the fact that $\mathbf{A}$ is composed of shifted versions of filters by the equation

$$\mathbf{A}_{i,j} = \mathbf{A}_{i\oplus k,j\oplus k} \tag{20}$$

where the subscripts indicate row and column of a matrix entry, k is the subsampling factor (assumed to be the same for each branch of the A/S system), and the $\oplus$ operator indicates addition modulo N. By definition, the inverse matrix $\mathbf{A}^{-1}$ must obey

$$\sum_{n=0}^{N-1} (\mathbf{A}^{-1})_{i,n}(\mathbf{A})_{n,j} = \delta_{i,j} \tag{21}$$

where $\delta_{i,j}$ is one when $i = j$ and zero otherwise. Substituting equation (20) into equation (21) gives

$$\sum_{n=0}^{N-1} (\mathbf{A}^{-1})_{i,n}\mathbf{A}_{n\oplus k,j\oplus k} = \delta_{i,j}$$

which becomes, with a change of variables

$$\sum_{n=0}^{N-1} (\mathbf{A}^{-1})_{i\ominus k,n\ominus k}\mathbf{A}_{n,j} = \delta_{i\ominus k,j\ominus k} = \delta_{i,j}$$

where the $\ominus$ operator indicates subtraction modulo N. This means that

$$(\mathbf{A}^{-1})_{i\ominus k,j\ominus k} = (\mathbf{A}^{-1})_{i,j}$$

or, with another change of variables

$$(\mathbf{A}^{-1})_{i,j} = (\mathbf{A}^{-1})_{i\oplus k,j\oplus k}$$

So $\mathbf{A}^{-1}$ retains the original interdigitated symmetry. It is clear that $(\mathbf{A}^{-1})^t$ and thus $\mathbf{A}'$ will have the same symmetry. ∎

Proposition 2 *If the original matrix* $\mathbf{A}$ *is full rank, then the procedure given in (19) will produce a full rank matrix* $\mathbf{A}'$.

Proof: Assume $\mathbf{A}$ is full rank but $\mathbf{A}'$ is not. Then there exists some non-trivial vector $\boldsymbol{v}$ such that

$$\mathbf{A}'\boldsymbol{v} = \left[\mathbf{A} + (\mathbf{A}^{-1})^t\right]\boldsymbol{v} = 0$$

This can be rewritten as

$$\mathbf{A}^t\mathbf{A}\boldsymbol{v} = -\boldsymbol{v}$$

This equation states that $\boldsymbol{v}$ is an eigenvector of the matrix $\mathbf{A}^t\mathbf{A}$ with eigenvalue -1. But since by assumption, $\mathbf{A}$ was full rank, $\mathbf{A}^t\mathbf{A}$ must be positive definite [29] and so cannot have negative eigenvalues. Thus $\mathbf{A}'$ must be full rank. ∎

Proposition 3 *An orthogonal matrix is a stable limit point of the iteration, and if the initial matrix* $\mathbf{A}$ *is such that the eigenvalues* λ_i *of* $\mathbf{A}^t\mathbf{A}$ *are all greater than 0.2, the iteration will converge to an orthogonal matrix.*

Proof: It is easy to see that orthogonal matrices represent limit points since $(\mathbf{A}^{-1})^t = \mathbf{A}$ if $\mathbf{A}$ is orthogonal. To demonstrate convergence, we want to show that $(\mathbf{A}')^t\mathbf{A}'$ is in some sense "closer" to the identity matrix $\mathbf{I}$ than is $\mathbf{A}^t\mathbf{A}$. We can do this by showing that the product of the matrix $(\mathbf{A}')^t\mathbf{A}'$ with any vector $\boldsymbol{v}$ produces a vector which is closer to $\boldsymbol{v}$ than the corresponding product using the matrix $\mathbf{A}$. Mathematically, we want to show that

$$\left\|\mathbf{A}^t\mathbf{A}\boldsymbol{v} - \mathbf{I}\boldsymbol{v}\right\| \geq \left\|(\mathbf{A}')^t\mathbf{A}'\boldsymbol{v} - \mathbf{I}\boldsymbol{v}\right\|$$

or equivalently,

$$\boldsymbol{v}^t\left[\left(\mathbf{A}^t\mathbf{A} - \mathbf{I}\right)\left(\mathbf{A}^t\mathbf{A} - \mathbf{I}\right)\right]\boldsymbol{v} - \boldsymbol{v}^t\left[\left((\mathbf{A}')^t\mathbf{A}' - \mathbf{I}\right)\left((\mathbf{A}')^t\mathbf{A}' - \mathbf{I}\right)\right]\boldsymbol{v} \geq 0$$

for *any* unit vector $\boldsymbol{v}$, with equality occuring when $\mathbf{A}$ is orthogonal. Since $\mathbf{A}$ is full rank, the product $\mathbf{A}^t\mathbf{A}$ is positive definite and will have N eigenvalues λ_i and N orthogonal eigenvectors $\boldsymbol{v}_i$. If we can show that the above expression holds for all of the eigenvectors, then by superposition it will hold for any vector $\boldsymbol{v}$.

So we will assume that $\boldsymbol{v}$ in the above expression is a normalized eigenvector of $\mathbf{A}^t\mathbf{A}$ with eigenvalue λ. Then

$$\boldsymbol{v}^t\left[\mathbf{A}^t\mathbf{A}\right]\boldsymbol{v} = \lambda$$

and

$$\begin{aligned}\boldsymbol{v}^t\left[(\mathbf{A}')^t\mathbf{A}'\right]\boldsymbol{v} &= \tfrac{1}{4}\boldsymbol{v}^t\left[\mathbf{A}^t\mathbf{A} + 2\mathbf{I} + \mathbf{A}^{-1}(\mathbf{A}^{-1})^t\right]\boldsymbol{v} \\ &= \tfrac{1}{4}\left[\lambda + 2 + \tfrac{1}{\lambda}\right]\end{aligned}$$

and so the condition that we are trying to prove becomes

$$\begin{aligned}\boldsymbol{v}^t\left[\left(\mathbf{A}^t\mathbf{A}\right)^2 - 2\mathbf{A}^t\mathbf{A} + \mathbf{I}\right]\boldsymbol{v} - \boldsymbol{v}^t\left[\left((\mathbf{A}')^t\mathbf{A}'\right)^2 - 2(\mathbf{A}')^t\mathbf{A}' + \mathbf{I}\right]\boldsymbol{v} \\ &= \left[\lambda^2 - 2\lambda + 1\right] - \tfrac{1}{16}\left[\lambda^2 + 4\lambda + 2 + 4 + \tfrac{4}{\lambda} + \tfrac{1}{\lambda^2}\right] \\ &= \tfrac{15}{16}\lambda^2 - \tfrac{7}{4}\lambda + \tfrac{5}{8} + \tfrac{1}{4}\left(\tfrac{1}{\lambda}\right) - \tfrac{1}{16}\left(\tfrac{1}{\lambda^2}\right) \\ &\geq 0\end{aligned}$$

A minimal amount of investigation reveals that this expression holds for all $\lambda > 0.2$, and thus if our starting matrix $\mathbf{A}$ is such that the eigenvalues of $\mathbf{A}'\mathbf{A}$ are all greater than 0.2, the iterative procedure will converge. Note that the eigenvalues of $\mathbf{A}^t\mathbf{A}$ are the squares of the singular values of the matrix $\mathbf{A}$. ∎

Proposition 4 *If the the associated subspaces of the original filters are orthogonal subspaces (and thus the original transformation matrix* $\mathbf{A}$ *partitions* $\mathbf{R}^N$

into orthogonal subspaces), then $(\mathbf{A}^{-1})^t$ *and thus* $\mathbf{A}'$ *will partition* $\mathbf{R}^N$ *into the same orthogonal subspaces.*

Proof: We will show this result for a two-band system. The extension to multiple subspaces is straightforward. Consider a standard form transformation matrix $\mathbf{A}$ which is composed of two filters obeying the mutual orthogonality condition given in equation (10):

$$\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2]$$

where $\mathbf{A}_1$ and $\mathbf{A}_2$ are $N \times \frac{N}{2}$ matrices, each containing shifted versions one of the two filters, and therefore spanning orthogonal subspaces. Then the inverse of the matrix may be written

$$\begin{aligned}
\mathbf{A}^{-1} &= \left[\mathbf{A}^t\mathbf{A}\right]^{-1}\mathbf{A}^t \\
&= \begin{bmatrix} [\mathbf{A}_1^t\mathbf{A}_1]^{-1} & 0 \\ 0 & [\mathbf{A}_2^t\mathbf{A}_2]^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{A}_1^t \\ \mathbf{A}_2^t \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{A}_1^\dagger \\ \mathbf{A}_2^\dagger \end{bmatrix}
\end{aligned}$$

where $\dagger$ indicates the Moore-Penrose pseudo-inverse [29]:

$$\mathbf{A}_i^\dagger = \left[\mathbf{A}_i^t\mathbf{A}_i\right]^{-1}\mathbf{A}_i^t$$

From the form of $\mathbf{A}^{-1}$, and since a matrix and its pseudo-inverse span the same space [29], it is clear that $\mathbf{A}' = \frac{\mathbf{A}+(\mathbf{A}^{-1})^t}{2}$ will partition $\mathbf{R}^N$ into the same orthogonal subspaces as $\mathbf{A}$. ∎

The disadvantages of this iterative method are that it depends on the initial conditions and that we have no control over the frequency response of the outcome. The advantages of this method are its simplicity and its flexibility, both of which make it useful as an exploratory tool. We shall see in section 5 that it is easily extended to multiple dimensions.

Proposition 4 allows us to reduce the computational load of the design method if the initial filters are chosen so that their associated subspaces are orthogonal (ie, if they are chosen to satisfy equation (10). For the two-band system, designing such filters involves selecting an initial lowpass filter and a matrix (image) dimension N', placing shifted copies of the filter into the columns of an $N' \times \frac{N'}{2}$ matrix $\mathbf{A}_1$, and iteratively averaging the matrix with the transpose of its pseudo-inverse:

$$\mathbf{A}'_1 \quad \leftarrow \quad \frac{1}{2}\left[\mathbf{A}_1 + (\mathbf{A}_1^{\dagger})'\right]$$

From the results of the previous section, the size of the resulting orthogonalized filter will be $M > (\frac{N+4}{2})$ and in practice, the iterative procedure produces filters of size N'. Thus, N' should be chosen to be the desired filter length. If the desir filter length is odd, the initial filter choice must be adjusted until the procedure produces filters with one impulse response sample equal to zero.

As an example of the flexibility of the method, consider a three-band design problem where we would like a highpass filter covering the top third of the frequency band and two lowpass filters with the same magnitude frequency response and (possibly) different phase covering the lower two thirds. For this example, we choose an image size $N' = 15$ (this must be divisible by the sub-

sampling factor 3) and initial filters

$$\begin{aligned} f_0 &= [x, \quad 1.0, \quad x] \\ f_1 &= [x, \quad 1.0, \quad x] \\ f_2 &= [-x, \quad 1.0, \quad -x] \end{aligned}$$

Since the subspaces associated with these filters are *not* orthogonal (ie, they do not satisfy equation (10)), choosing $x = 0.5$ will not guarantee that the filters produced by the iterative procedure will be highpass and lowpass. The parameter x can be adjusted until the resulting highpass filter has a zero D.C. response. This was found to occur at $x = 0.6735$, and the resulting set of filters is shown in figure 10. Note that the highpass filter is symmetric and the lowpass filters are inverted versions of each other. These filters are not optimal, but the iterative technique indicates that such filters exist, and indicates an appropriate sampling strategy and symmetry relationship between the two lowpass filters. An optimization routine could now be written to achieve filters with sharper transitions in the frequency domain. This will be demonstrated for the two-band case in the next section.

Frequency Domain Design

We now describe an alternative design procedure which is easily applied to the one-dimensional two-band case. It may be modified for application to any one-dimensional system. In the frequency domain, the condition for perfect reconstruction (assuming the filters have been chosen to eliminate the aliasing term as in equation (14)) is

$$\left|H(\omega)\right|^2 + \left|H(\omega+\pi)\right|^2 = 2 \tag{22}$$

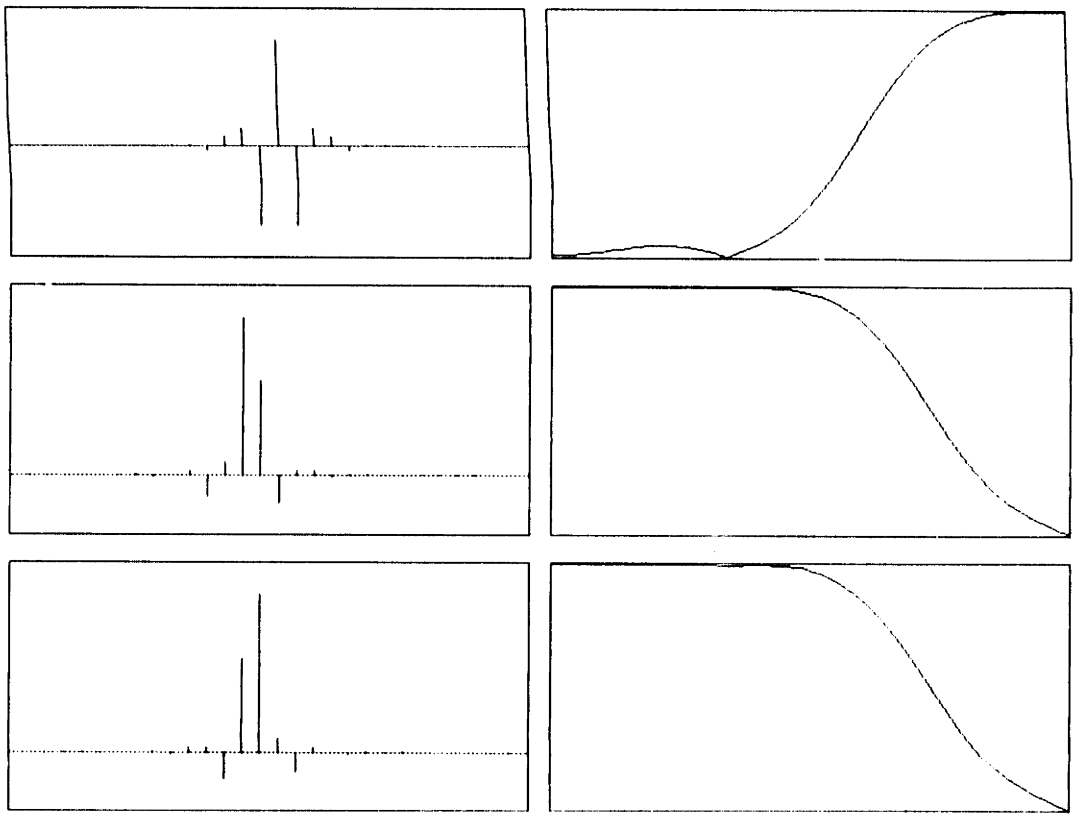


Figure 10: An asymmetric three-band system. On the left side are the filter impulse responses, and on the right are their DTFTs, plotted from 0 to π with a linear scale. The two lowpass filters have identical power spectra, but conjugate phase. The highpass filter is symmetric.

The matrix averaging method described above produces filters satisfying this constraint at the points $\omega = \frac{2\pi k}{N'}$ for $k \in \{0, 1, \ldots, N'-1\}$. This suggests that an alternative design method is that of frequency sampling, in which we force the filters to obey the constraint in (22) at the frequency samples of an N'-point DFT. It is known that digital filter designs based on frequency sampling are usually unsatisfactory due to poor behavior between the sample points [22]. This is because typical design specifications involve sharp transitions in the frequency domain. The present design constraints, however, do not necessitate sharp transitions and the frequency sampling designs perform remarkably well.

In designing the filter $h[n]$ by frequency sampling, there are $\left\lfloor \frac{N'-2}{4} \right\rfloor$ free parameters (where $\lfloor \cdot \rfloor$ indicates the floor function) corresponding to the frequency response $H(\frac{2\pi k}{N'})$ for $k \in \left\{0, 1, ..., \left\lfloor \frac{N'-2}{4} \right\rfloor \right\}$. The value of $H(\frac{2\pi k}{N'})$ for $k \in \left\{ \left\lfloor \frac{N'-2}{4} \right\rfloor + 1, \left\lfloor \frac{N'-2}{4} \right\rfloor + 1, ..., N'-1 \right\}$ are determined from the filter symmetry equation

$$H(\omega) = H(-\omega)$$

and, by rewriting equation (22):

$$H(\omega) = \sqrt{2 - H^2(\omega + \pi)}$$

where we have dropped the modulus symbol from equation (22) because $H(\omega)$ is real. Since we are interested in a lowpass filter, the typical choice is to let $H(\frac{2\pi k}{N'}) = \sqrt{2}$ for $k \in \left\{0, 1, ..., \left\lfloor \frac{N'-2}{4} \right\rfloor \right\}$. If the desired filter length is N', we need only invert the resulting transform $H(\omega)$ to get $h[n]$. If the desired filter length is odd, we must adjust one or more of the parameter choices until the resulting filter has one impulse response sample equal to zero. It was found empirically

that the best results were achieved by allowing only the last free parameter $H(\lfloor \frac{N'-2}{4} \rfloor)$ to vary.

We have used the frequency sampling technique discussed above to design filters of length 5, 7, 8, 9, 11, 12 and 13 for use in the two-band A/S system. For the 5-tap filter, the orthogonality equations may be solved analytically and this is done in Appendix A. In Appendix B, we provide the values of the impulse response samples for the filters. Table 1 gives several useful error measures for each of the filters given in appendix B, as well as for two of Johnston's filters [55]. The value of E_{aliasing} provides a measure of the aliasing energy in the lowpass filter. It is the RMS difference between the response of the lowpass branch of the A/S system to an impulse at $n = 0$ and one at $n = 1$. Let $\hat{x}^{(0)}[n]$ be the response of the lowpass branch of the A/S system to an impulse at $n = 0$. Let $\hat{x}^{(1)}[n]$ be the response of the lowpass branch of the A/S system to an impulse at $n = 1$. Then define

$$E_{\text{aliasing}} \equiv \left[\sum_n \left(\hat{x}^{(0)}[n] - \hat{x}^{(1)}[n+1] \right)^2 \right]^{\frac{1}{2}}$$

The value of E_{orth} provides a mean-squared measure of the overall system reconstruction error. Let $\hat{x}[n]$ be the overall A/S system impulse response. Then define

$$E_{\text{orth}} \equiv \left[\sum_n (\hat{x}[n] - \delta[n])^2 \right]^{\frac{1}{2}}$$

The value of E_{abs} indicates the absolute value of the largest error which occurs when the A/S system is applied to the 256×256 "Lena" image.

Several of Johnston's filters [55] are provided for comparison with the even-length filters. These are indicated by the letter "J". An A/S system based

filter	E_{aliasing}	E_{orth}	E_{abs}
5	0.617	8.19E-3	2.597
7	0.595	7.95E-5	0.023
8A	0.595	6.53E-4	0.265
8B	0.519	3.86E-3	1.567
8J	0.519	2.93E-3	1.572
9	0.478	1.35E-3	0.432
11	0.486	7.54E-4	0.283
12A	0.478	8.63E-4	0.271
12B	0.455	1.59E-3	0.415
12J	0.455	1.14E-3	0.899
13	0.404	1.59E-3	0.593

Table 1: Error measures of the filters given in Appendix B.

on his filters has a slightly lower MSE than the filters designed here, but the maximum absolute value error is in general higher.

Figures 11 through 13 give plots of the basis functions and their DTFTs for a four-level pyramid using the 5, 9 and 13-tap filters. These should be compared to the DTFTs of several of the 16-point DCT basis functions which are given in figure 1 of the introduction. Although the basis functions of the full-size DCT are maximally localized in spatial frequency, the block DCT is rather poorly localized. The block nature of the transform leads to exaggerated Gibbs ringing in the frequency domain and suggests that the orthogonal subband representation may be preferable for many tasks.

Non-orthogonal Systems

For some applications, it is desirable to be able to perform either the analysis or the synthesis computations with inexpensive hardware. In such cases, a non-orthogonal system may be useful. For example, suppose we wish to reconstruct the image with the bilinear interpolation filters $g_0 = [1, 2, 1]$ and $g_1 = [-1, 2, -1]$ which could be implemented efficiently with arithmetic shifts and additions. Note that the associated subspaces of these two filters are orthogonal, even though the filters do not correspond to an orthogonal basis set. From the discussion of section 1, we can find suitable analysis filters by inverting the matrix G formed from shifted versions of the filters $g_i[n]$. The true inverse would require inverting an $N \times N$ matrix where N is the number of pixels in the image — clearly an impractical solution. In keeping with the concept of designing the filters by placing them in a reduced size image, we can find approximate analysis

filters by inverting an $N' \times N'$ matrix formed from shifted versions of the filters $g_i[n]$, where $N' << N$. As an example, the inverse filters f_i have been computed with matrix sizes $N' = 16$ and $N' = 22$. The resulting impulse response samples are given in Appendix B. Use of these filters will be demonstrated in the next section.

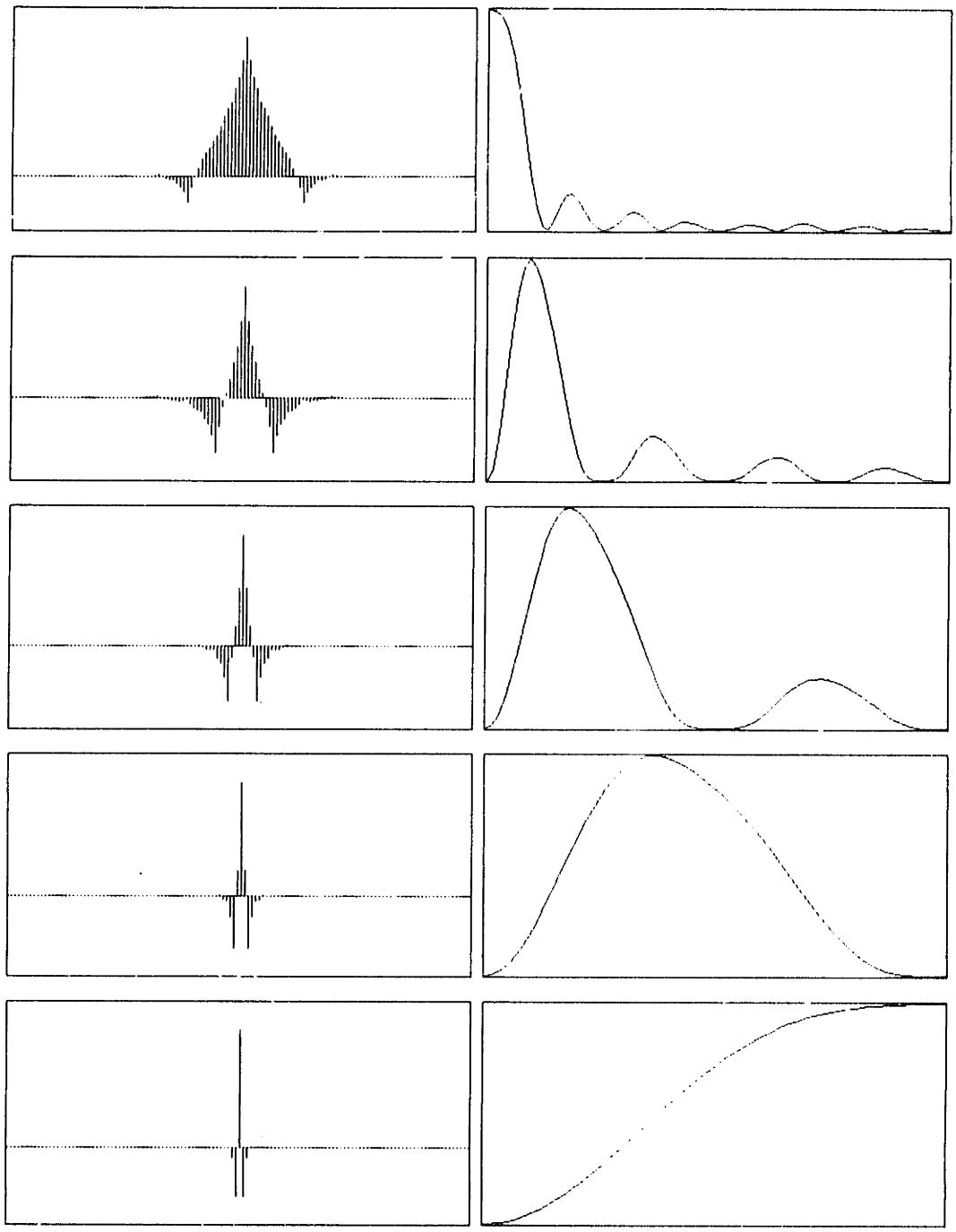


Figure 11: Basis functions and DTFTs for a four-level pyramid constructed from the 5-tap filter given in Appendix A. DTFTs are plotted on linear scale axes.

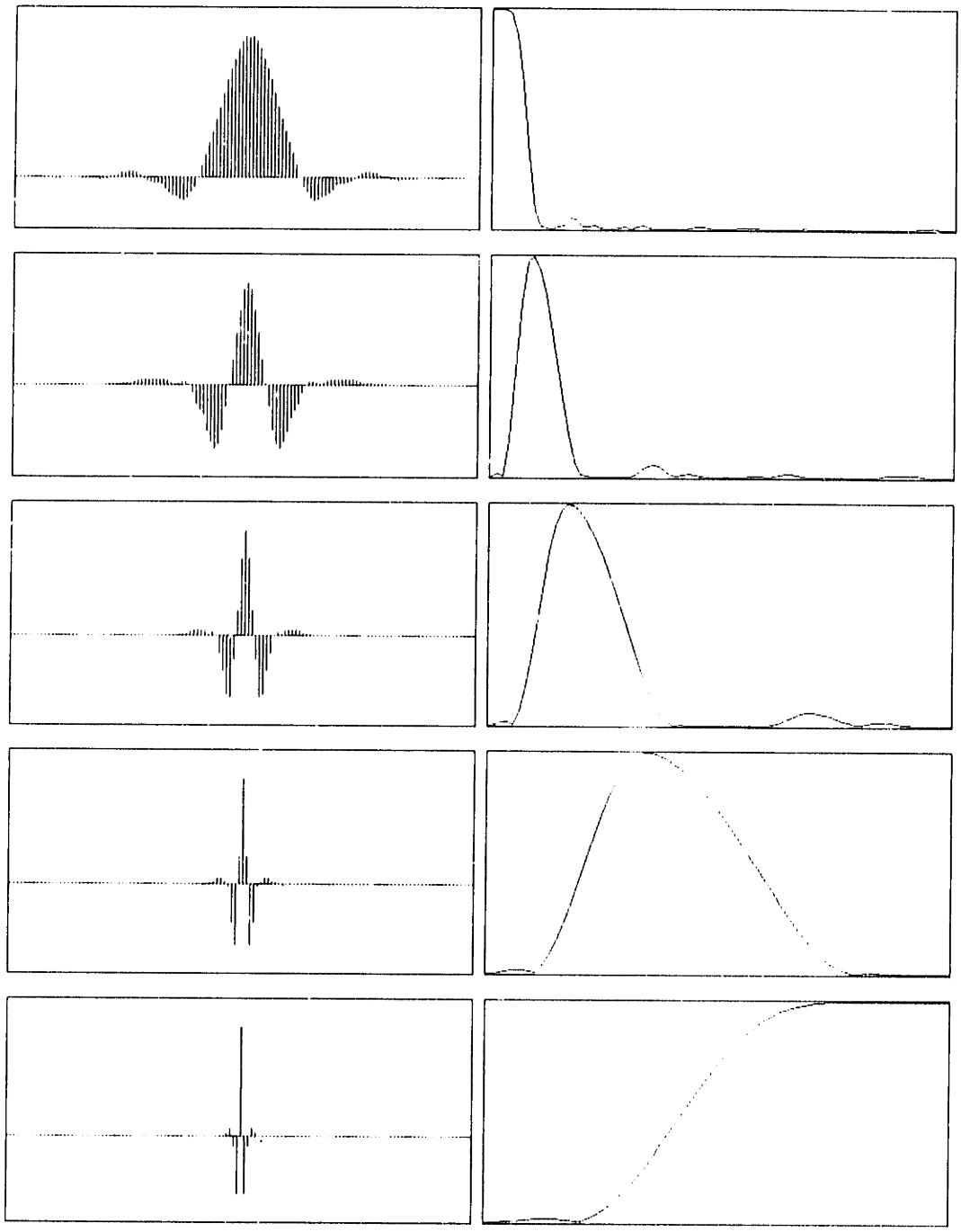


Figure 12: Basis functions and DTFTs for a four-level pyramid constructed from the 9-tap filter given in Appendix B. DTFTs are plotted on linear scale axes.

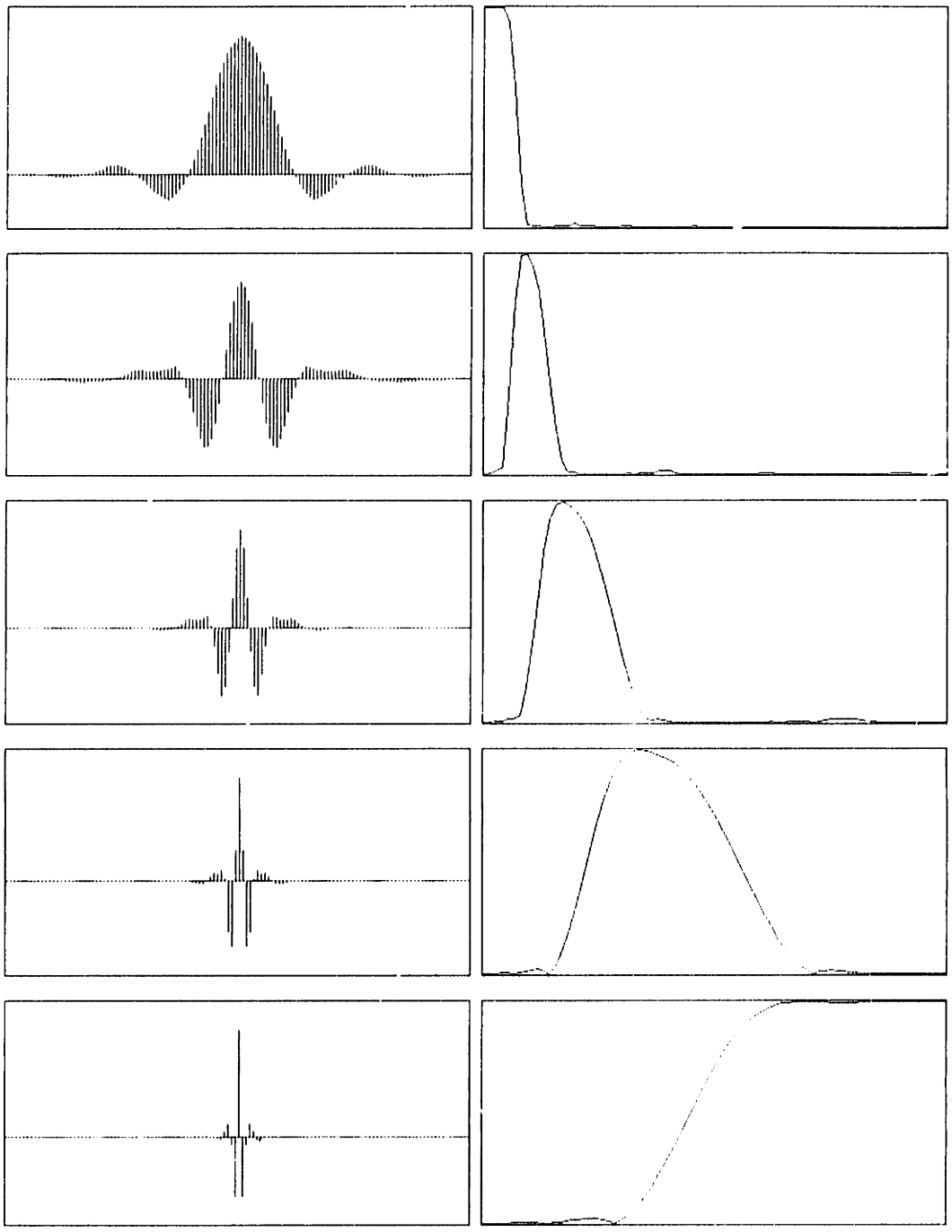


Figure 13: Basis functions and DTFTs for a four-level pyramid constructed from the 13-tap filter given in Appendix B. DTFTs are plotted on linear scale axes.

4 Applications

In this section, we will give examples of two image processing applications using the filters designed in the previous section in a separable manner.

Data Compression

An optimal transform for data compression should minimize the bit rate for a given allowable error in the reconstructed image. If the basis functions of the transform are orthonormal, and if expected mean square differences are used as an error measure, this is equivalent to maximizing the following expression for the gain in coding over PCM [20]:

$$G = \frac{\frac{1}{N}\sum_{j=0}^{N-1}\sigma_j^2}{\left[\prod_{j=0}^{N-1}\sigma_j^2\right]^{1/N}}$$

where σ_j^2 is the variance of the jth transform coefficient. This measure was computed for some of the filters given in Appendix B, and is displayed in table 2. Values were computed assuming Markov second order signal statistics, where the autocorrelation matrix $\mathbf{R}_{xx}$ is a symmetric Toeplitz matrix of the form

$$\mathbf{R}_{xx} = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^N \\ \rho & 1 & \rho & & \rho^{(N-1)} \\ \rho^2 & \rho & 1 & & \rho^{(N-2)} \\ \vdots & & & \ddots & \\ \rho^N & \rho^{(N-1)} & \rho^{(N-2)} & & 1 \end{bmatrix}$$

A value of $\rho = 0.95$ was used to compute the numbers given in table 2, Values are given for circular convolution on an image of size $N = 256$ and for linear convolution, where the image edges were handled by reflecting the filters in a manner that preserves the orthogonality of the basis set. This reflection technique, which is described in Appendix C, is preferable because in circular convolution, the transform coefficients near the image edges will contain information from opposite sides of the image which are highly uncorrelated. Comparable values for a 16-point block DCT and a 32-point block DCT are also given. A 9-tap sub-band filter gives slightly better value than the 16-point DCT, and the 13-tap sub-band filter is substantially better. This does not necessarily correspond to measurements of subjective quality. We have found that images compressed with a 9-tap subband transform are perceptually superior to the 16-point DCT, primarily because of the absence of block artifacts.

Several authors have used separable QMF filters for the purposes of image coding. Vetterli [58] was the first to apply QMF filters separably to two-dimensional images. Woods and O'Neill [59] used DPCM to code the sixteen bands formed by a uniform cascade of separable filters. They used 32 and 80-tap filters designed by Johnston [55]. Gharavi and Tabatabai [60] used a pyramid of separable filters and in [61], applied it to color images. Tran et. al. [62] used an extension of Chen and Pratt's [63] combined Huffman and run-length coding scheme to code QMF pyramids.

The separable staggered two-band A/S system produces a partition of the frequency domain as illustrated in the diagram in figure 14. Also shown are the four (staggered) subsampling lattices. Several examples of data compression

filter	$G_{circular}$	$G_{reflected}$
5-tap	5.83	5.93
7-tap	6.99	7.10
8-tap	7.98	8.17
9-tap	8.93	9.05
13-tap	12.043	12.20
DCT-16	8.82	
DCT-32	9.49	

Table 2: Theoretical Coding gains over PCM for a four-level pyramid based on several different one-dimensional symmetric orthogonal sub-band filters. Second order signal statistics were assumed to be Markov with $\rho = 0.95$. Gain factors are given for circular convolution and reflected convolution, as defined in Appendix C.

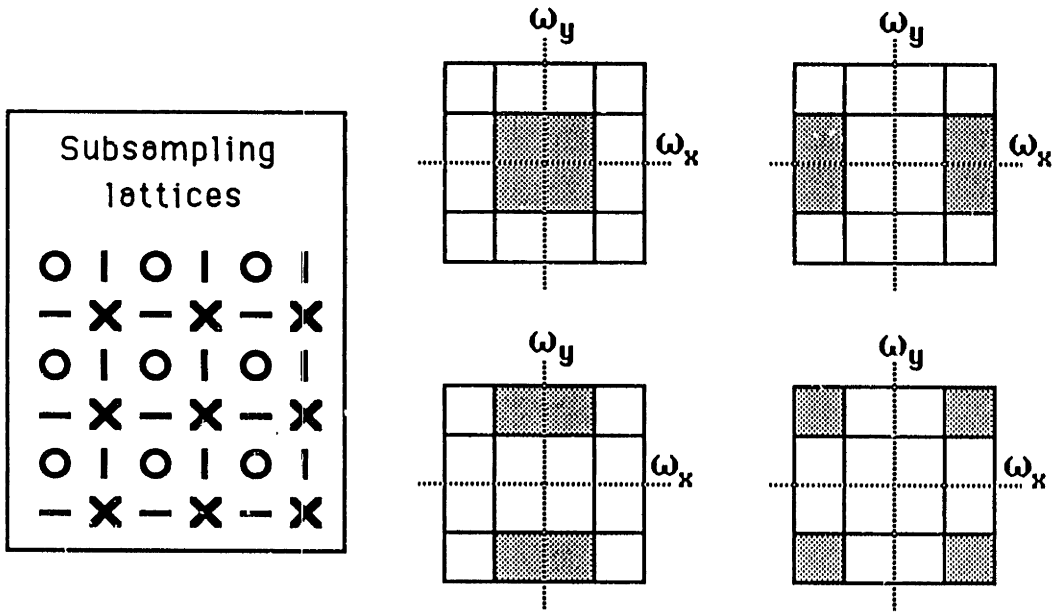


Figure 14: Illustration of the sampling grid and idealized frequency domain partition for four-band separable system. An "o" indicates that a sample from the lowpass-filtered image is retained at this location. A "–" indicates that a highpass vertical sample is retained. A "-" indicates that a highpass horizontal sample is retained. An "x" indicates that a highpass diagonal sample is retained.

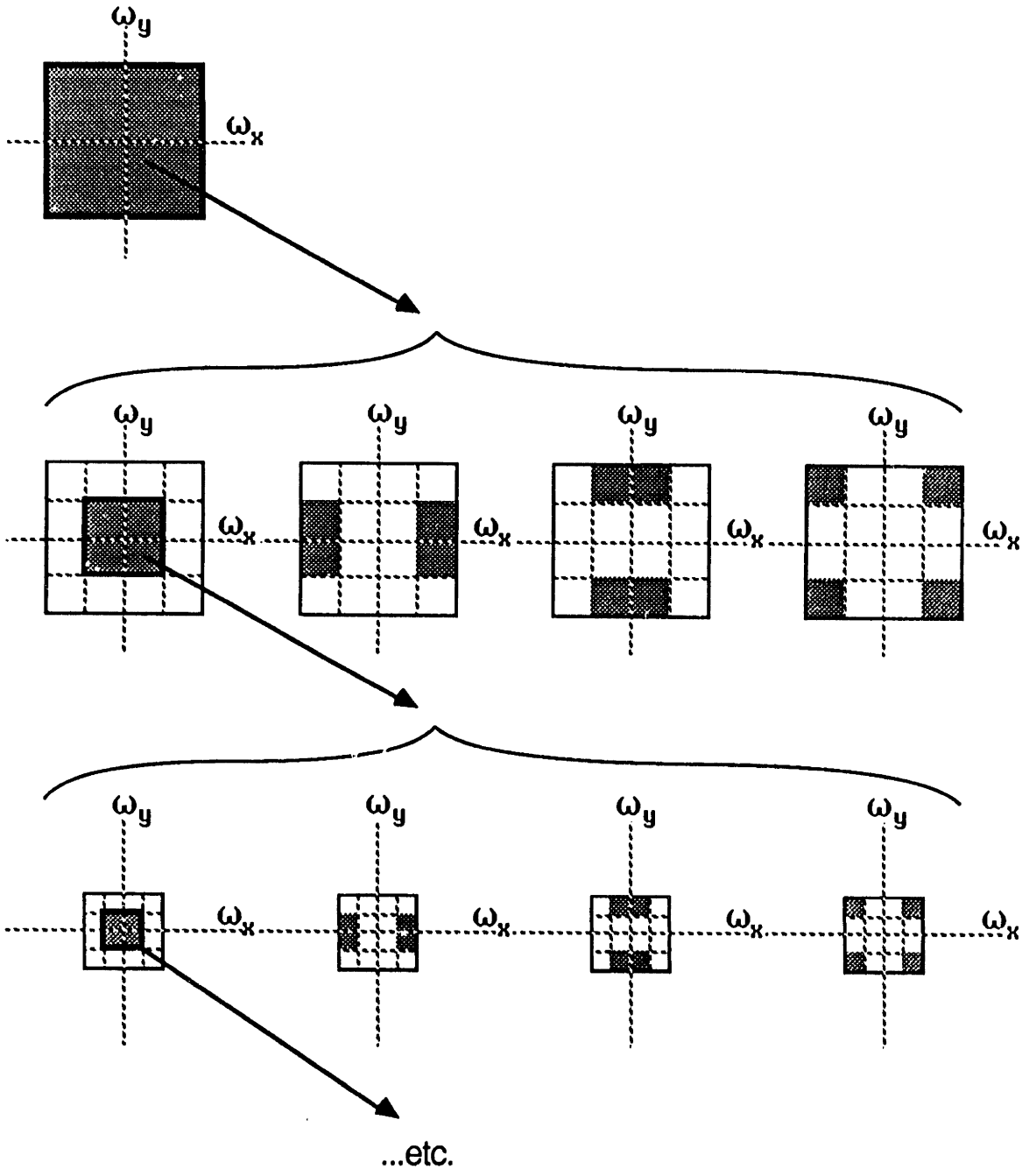


Figure 15: Idealized diagram of the partition of the frequency plane resulting from a 4-level pyramid cascade of separable 2-band filters. The top plot represents the frequency spectrum of the original image, with axes ranging from $-\pi$ to π. This is divided into four sub-bands at the next level. On each subsequent level, the lowpass sub-band (outlined in bold) is sub-divided further.

of the 256 × 256 "Lena" image were computed using this separable system and are given in figures 16 through 21. A four-level separable pyramid transform was applied to the original image. This transformation partitions the frequency domain into octave-spaced oriented sub-bands, as illustrated in the idealized frequency diagram of figure 15.

An overall bit rate (entropy) was fixed and the bit rates assigned to the coefficients of the transform were determined using the standard optimal allocation formula [20]:

$$R_k = R + \frac{1}{2}\log_2 \frac{\sigma_k^2}{\left[\prod_{j=0}^{N-1} \sigma_j^2\right]^{1/N}}$$

where, as before, σ_k^2 is the variance of the kth coefficient in the transform. Negative values of R_k were set to zero and the other bit rates raised to maintain the correct overall bit rate R. Note that if we assume stationary image statistics, the σ_k^2 are the same for all coefficients belonging to the same sub-image of the transform. It has been shown [20] that the optimal quantizer for entropy coding is nearly uniform for bit rates which are high enough that the image probability distribution is approximately constant over each bin. Even though the examples shown were compressed to relatively low bit rates, uniform quantization was used due to its simplicity. Each sub-image was quantized with the bin size chosen to give a first order entropy equal to the optimal bit rate R_k for that subimage. An interesting comparison may be made between compression with the 9-tap and 13-tap filters, as shown in figures 20 and 21, respectively. The 9-tap filter produces more aliasing, but the Gibbs ringing is more noticeable in the case of the 13-tap filter. Schreiber et al [64] noted that some amount of aliasing may

be perceptually preferable to the Gibbs ringing that results from using sharp filters.

Progressive Transmission

Another application of the pyramid sub-band representation is that of progressive transmission. This involves sending an image through a low-capacity channel in such a way that a low resolution version of the image becomes available quickly, and higher resolution information is added in a gradual manner. In the case of a sub-band pyramid, this is easily accomplished by sending the information in order from lowest to highest resolution. Figure 22 shows a sequence of intermediate progressive transmission images for a four-level pyramid transform using 9-tap separable filters.

Since in progressive transmission, it is likely that one would desire a simple inexpensive receiver, a non-orthogonal system may be preferable. A sequence of progressive transmission images for a four-level pyramid using the linear interpolation filter and its inverse (given in Appendix B) is shown in figure 23. The initial images in this sequence exhibit more obvious aliasing effects than those of figure 22, but the significant computational advantages could outweigh the modest increase in distortion.

Figure 16: Original 8-bit 256 × 256 "Lena" image.

Figure 17: Data compression example using a four-level pyramid with 9-tap separable filter. First order entropy is 1.0 bit/pixel. RMS error is 3.86.

Figure 18: Data compression example using a four-level pyramid with 9-tap separable filter. First order entropy is 0.35 bit/pixel. RMS error is 8.13.

Figure 19: Data compression example using a four-level pyramid with 5-tap separable filter. First order entropy is 0.5 bit/pixel. RMS error is 7.05.

Figure 20: Data compression example using a four-level pyramid with 9-tap separable filter. First order entropy is 0.5 bit/pixel. RMS error is 6.02.

Figure 21: Data compression example using a four-level pyramid with 13-tap separable filter. First order entropy is 0.5 bit/pixel. RMS error is 5.86.

Figure 22: Progressive transmission images for a four-level pyramid transform using 9-tap separable filters. The first image is produced by reconstructing only the lowpass signal at the fourth pyramid level. The next image results from reconstructing the entire fourth level, or equivalently, the lowpass image on the third level. Bit rates (entropies) are: 0.03, 0.09, 0.24, and 0.49 bits/pixel and the image size is 256×256. The final image at 0.65 bits/pixel is not shown.

Figure 23: Progressive transmission images for a four-level pyramid transform using 15-tap separable analysis filters and 3-tap separable synthesis filters. Bit rates are essentially the same as for the 9-tap example above.

5 Extensions to Two and Three Dimensions

The application of QMF concepts to two-dimensional images was first suggested by Vetterli in [58], where he discussed a four-band separable and a two-band case. Pei and Jaw [65] suggested a separable application of the perfect reconstruction filters developed by Smith and Barnwell. In this section, we discuss a more general extension of orthogonal sub-band filters to two and three dimensions.

Sampling of Multi-dimensional Signals

In order to understand the extension of orthogonal sub-band filter concepts to multiple dimensions, we must first review the effects of sampling in multiple dimensions. It will be shown that the subsampling scheme places constraints on the filter spectra. Sampling a one-dimensional signal corresponds to a replication of its frequency spectrum in the frequency domain. The sampling "lattice" is parameterized by a single number—the sample spacing T—and replication of the spectrum in the frequency domain occurs at intervals of $\frac{2\pi}{T}$. In d dimensions, a sampling lattice may be parameterized by a set of d linearly independent vectors $\{\boldsymbol{v}_i : 0 \leq i \leq d-1\}$. The locations of the lattice sampling points consist of all *integral* linear combinations of the vectors $\boldsymbol{v}_i$:

$$\sum_{i=0}^{d-1} n_i \boldsymbol{v}_i, \qquad n_i \in \mathbf{N} \text{ for all } i$$

More compactly, we may incorporate the lattice vectors into a square non-singular sampling matrix $\mathbf{V}$[23]. This allows us to describe a lattice point location as a matrix-vector product $\mathbf{V}\boldsymbol{n}$, where $\boldsymbol{n} \in \mathbf{N}^d$, the space of all integer d-tuples. An example will be given shortly, but first we will consider the issue

of uniqueness.

Although there is a unique lattice associated with every sampling matrix $\mathbf{V}$, the converse statement is not true. We may alter the order of the columns of $\mathbf{V}$ without affecting the lattice. We may also multiply any column by -1. More generally, we may add an integer multiple of any column of $\mathbf{V}$ to any other column, as we show in the following proposition.

Proposition 5 *The lattice corresponding to a matrix* $\mathbf{V}$ *is unaltered if an integer multiple of one column of* $\mathbf{V}$ *is added to another column of* $\mathbf{V}$.

Proof: We first write the original sampling matrix as

$$\mathbf{V} = \begin{bmatrix} | & | & & | \\ \boldsymbol{v}_0 & \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_d \\ | & | & & | \end{bmatrix}$$

where the $\boldsymbol{v}_i$ are column vectors of height d (d is the dimensionality of the image space being sampled). Let

$$\mathbf{V}' = \begin{bmatrix} | & | & & | & | & | & & | \\ \boldsymbol{v}_0 & \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_{k-1} & \boldsymbol{v}_k + \boldsymbol{v}_j & \boldsymbol{v}_{k+1} & \cdots & \boldsymbol{v}_{d-1} \\ | & | & & | & | & | & & | \end{bmatrix}$$

for some j and k such that $0 \leq j < k \leq d-1$. Next we choose some arbitrary vector $\boldsymbol{n} \in \mathbf{N}^d$, thus defining a particular point in the lattice associated with $\mathbf{V}$:

$$\mathbf{V}\boldsymbol{n} = \sum_{i=0}^{d-1} n_i \boldsymbol{v}_i$$

We need to show that this point is also contained in the lattice defined by the matrix $\mathbf{V}'$. We define vector $\boldsymbol{n}' \in \mathbf{N}^d$ by

$$\boldsymbol{n}' = \left(n_0, n_1, \ldots, n_{j-1}, (n_j - n_k), n_{j+1}, \ldots, n_{d-1}\right)^t$$

Then

$$\begin{aligned}
\mathbf{V}'\boldsymbol{n}' &= \sum_{i=0}^{j-1} n_i \boldsymbol{v}_i + (n_j - n_k)\,\boldsymbol{v}_j + \sum_{i=j+1}^{k-1} n_i \boldsymbol{v}_i + n_k\,(\boldsymbol{v}_j + \boldsymbol{v}_k) + \sum_{i=k+1}^{d-1} n_i \boldsymbol{v}_i \\
&= \sum_{i=0}^{d-1} n_i \boldsymbol{v}_i \\
&= \mathbf{V}\boldsymbol{n}
\end{aligned}$$

So the point $\mathbf{V}\boldsymbol{n}$ is contained in the lattice defined by $\mathbf{V}'$. The converse result may be demonstrated in the same manner, thus completing the proof. ∎

In one dimension, ideal sampling is equivalent to multiplication by a field of dirac delta functions. This is also true in multiple dimensions. We will denote the lattice associated with the identity sampling matrix as the *unit square* or *discrete* lattice:

$$L(\boldsymbol{r}) = \sum_{\boldsymbol{n} \in \mathbf{N}^d} \delta(\boldsymbol{r} - \boldsymbol{n})$$

Using this notation, we can define the lattice associated with sampling matrix $\mathbf{V}$ in terms of $L(\boldsymbol{r})$:

$$\begin{aligned}
L_{\mathbf{V}}(\boldsymbol{r}) &= \sum_{\boldsymbol{n} \in \mathbf{N}^d} \delta(\boldsymbol{r} - \mathbf{V}\boldsymbol{n}) \\
&= \sum_{\boldsymbol{n} \in \mathbf{N}^d} \delta(\mathbf{V}^{-1}\boldsymbol{r} - \boldsymbol{n}) \\
&= L(V^{-1}\boldsymbol{r}) \qquad (23)
\end{aligned}$$

where we have employed the useful fact that the argument of the delta function may be multiplied by any square invertible matrix. This equation says that any lattice may be viewed as resulting from the action of a matrix $\mathbf{V}$ on the points of the discrete lattice. This is illustrated in figure 24 for the hexagonal sampling

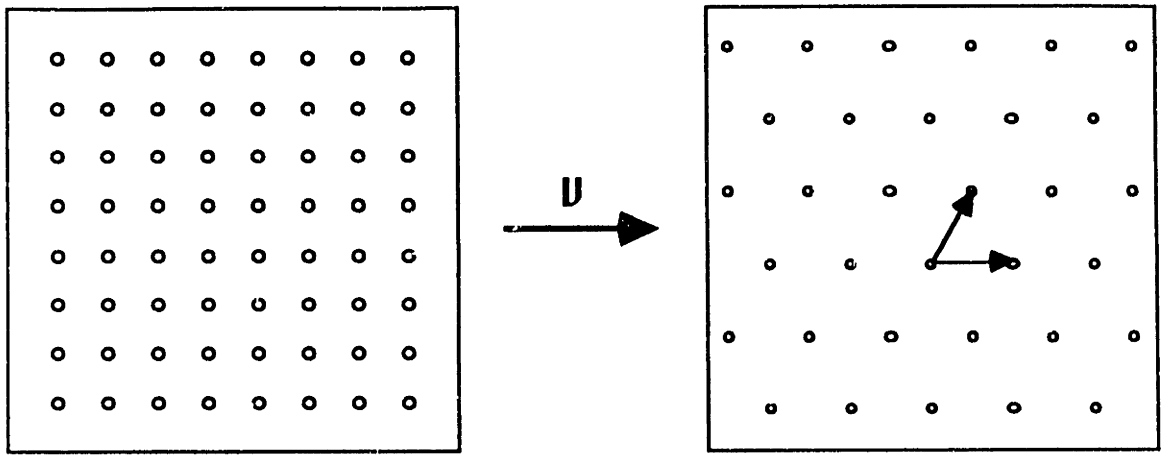


Figure 24: An illustration of the relationship between the unit square lattice and a hexagonal lattice.

matrix

$$\mathbf{V} = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix}$$

As in one dimension, the Fourier transform of lattice $L_{\mathbf{V}}(\boldsymbol{r})$ is also a lattice:

$$\begin{aligned}
\tilde{L}_{\mathbf{V}}(\boldsymbol{\Omega}) &= \int_{\boldsymbol{r}\in R^d} d\boldsymbol{r}\; L_{\mathbf{V}}(\boldsymbol{r}) e^{-j\boldsymbol{\Omega}^t\cdot\boldsymbol{r}} \\
&= \int_{\boldsymbol{r}\in R^d} d\boldsymbol{r} \sum_{\boldsymbol{n}\in \mathbf{N}^d} \delta(\boldsymbol{r}-\mathbf{V}\boldsymbol{n}) e^{-j\boldsymbol{\Omega}^t\cdot\boldsymbol{r}} \\
&= \sum_{\boldsymbol{n}\in \mathbf{N}^d} e^{-j\boldsymbol{\Omega}^t\cdot\mathbf{V}\boldsymbol{n}} \\
&= \sum_{\boldsymbol{n}\in \mathbf{N}^d} e^{-j(\mathbf{V}^t\boldsymbol{\Omega})^t\cdot\boldsymbol{n}} \\
&= \frac{(2\pi)^d}{|\mathbf{V}|} \sum_{\boldsymbol{n}\in \mathbf{N}^d} \delta(\boldsymbol{\Omega}-2\pi(\mathbf{V}^{-1})^t\boldsymbol{n}) \\
&= \frac{(2\pi)^d}{|\mathbf{V}|} L_{2\pi(\mathbf{V}^{-1})^t}(\boldsymbol{\Omega})
\end{aligned}$$

where $|\mathbf{V}|$ indicates the determinant of the matrix $\mathbf{V}$, and we have assumed convergence in order to switch the order of the integral and the sum. if we define a related matrix

$$\tilde{\mathbf{V}} \equiv 2\pi(\mathbf{V}^{-1})^t$$

then the Fourier transform of lattice $L_{\mathbf{V}}(\boldsymbol{r})$ may be written

$$\tilde{L}_{\mathbf{V}}(\boldsymbol{\Omega}) = \frac{(2\pi)^d}{|\mathbf{V}|} L_{\tilde{\mathbf{V}}}(\boldsymbol{\Omega}) = |\tilde{\mathbf{V}}| \; L_{\tilde{\mathbf{V}}}(\boldsymbol{\Omega}) \tag{24}$$

Again, except for the scale factor, the lattice may be viewed as resulting from the action of a matrix $\tilde{\mathbf{V}}$ on the unit square lattice. The relationship between the lattices in the frequency domain and spatial domain is illustrated in figure 25.

To summarize, if $x_c(\boldsymbol{r})$ is a continuous image, we will use the notation

$$x_{\mathbf{V}}(\boldsymbol{r}) = x_c(\boldsymbol{r}) \cdot L_{\mathbf{V}}(\boldsymbol{r}) \tag{25}$$

to indicate the corresponding sampled image, where $L_{\mathbf{V}}(\boldsymbol{r})$ is defined in equation (23), and we will use

$$X_{\mathbf{V}}(\boldsymbol{\Omega}) = \frac{X_c(\boldsymbol{\Omega}) * \tilde{L}_{\mathbf{V}}(\boldsymbol{\Omega})}{|\mathbf{V}|} \tag{26}$$

to indicate its Fourier transform, where $*$ indicates convolution and $\tilde{L}_{\mathbf{V}}(\boldsymbol{\Omega})$ is defined in equation (24).

Discrete Representation of Sampled Signals

The results of this section can be derived completely in the sampled notation given above, but they may be more easily manipulated in the unit square lattice sampling domain. As stated previously, any lattice is a transformed version of

the unit lattice. Thus, it makes sense to work with the corresponding *discrete* image representation, defined by:

$$x[\boldsymbol{n}] = x_c(\mathbf{V}\boldsymbol{n}) \tag{27}$$

It is important to recognize that the matrix $\mathbf{V}$ induces a somewhat arbitrary mapping from the discrete domain to the continuous image domain. Although integer column operations do not alter the sampling lattice $L_{\mathbf{V}}(\boldsymbol{r})$, they *do* alter the mapping from the lattice points to the discrete sample locations $\boldsymbol{n}$ and thus will affect the geometry of the mapping of a filter back into the continuous image domain. In the frequency domain, the standard multi-dimensional DTFT is defined as [22,23]

$$F_i(\boldsymbol{\omega}) = \sum_{n \in \mathbf{N}} f_i[\boldsymbol{n}] e^{-j\boldsymbol{\omega}^t \cdot \boldsymbol{n}}$$

and so we can relate the sampled and discrete representations by

$$\begin{aligned} X_{\mathbf{V}}(\boldsymbol{\Omega}) &= \int_{\boldsymbol{r} \in R^d} d\boldsymbol{r} \left[x_c(\boldsymbol{r}) \sum_{\boldsymbol{n} \in \mathbf{N}^d} \delta(\boldsymbol{r} - \mathbf{V}\boldsymbol{n}) \right] e^{-j\boldsymbol{\Omega}^t \cdot \boldsymbol{r}} \\ &= \sum_{\boldsymbol{n} \in \mathbf{N}^d} x_c(\mathbf{V}\boldsymbol{n}) e^{-j\boldsymbol{\Omega}^t \cdot \mathbf{V}\boldsymbol{n}} \\ &= X(\boldsymbol{\omega})\big|_{\boldsymbol{\omega} = \mathbf{V}^t \boldsymbol{\Omega}} \end{aligned}$$

or

$$X(\boldsymbol{\omega}) = X_{\mathbf{V}}(\boldsymbol{\Omega})\big|_{\boldsymbol{\Omega} = (\mathbf{V}^{-1})^t \boldsymbol{\omega}}$$

This relation is illustrated in figure 25. We will use this relation to understand the results of this section in terms of the original frequency domain.

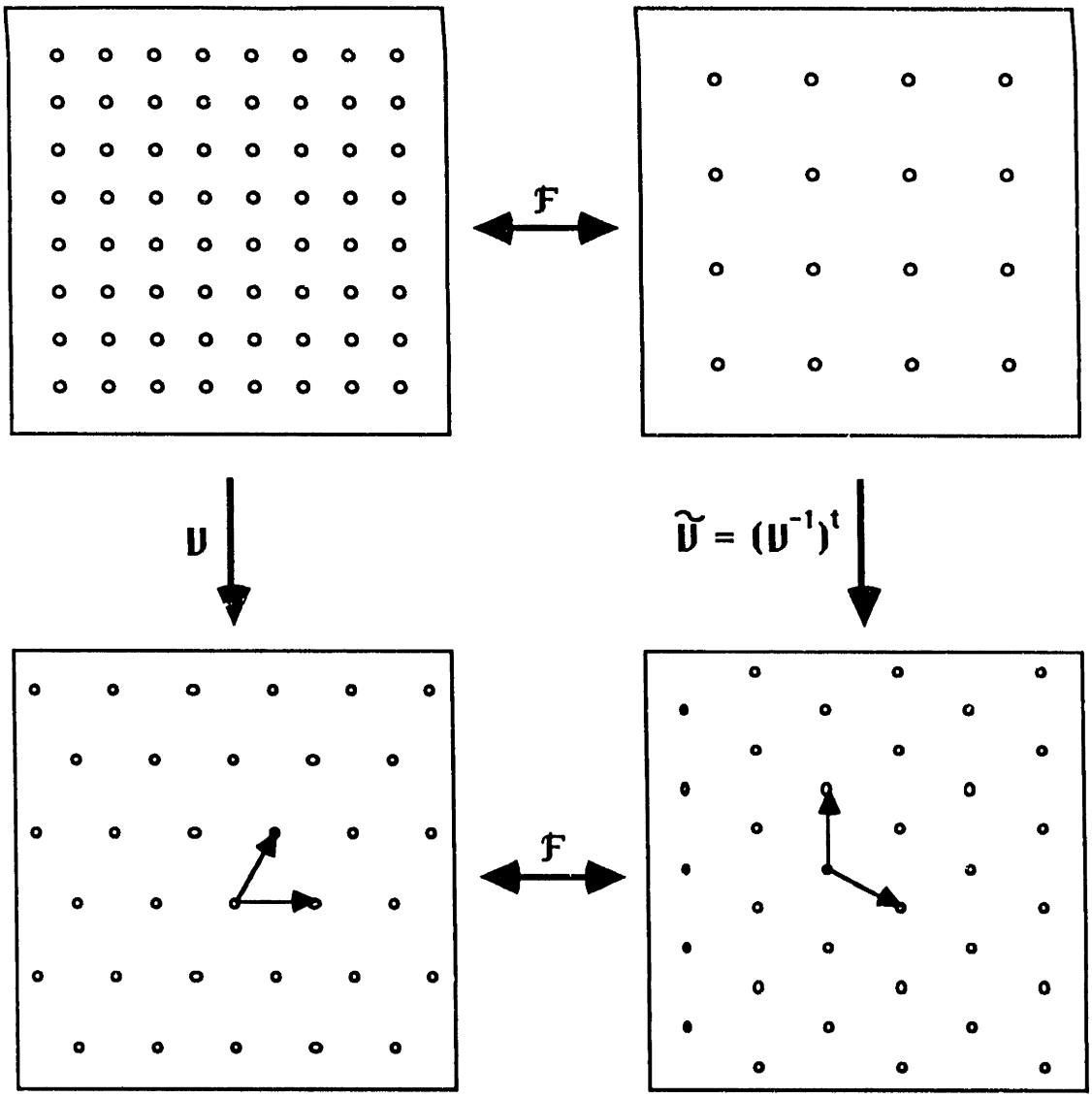


Figure 25: Relationship between lattices in different domains. Upper left is the discrete or unit lattice with points spaced one unit apart. Upper right is the DTFT of the unit lattice, which has points spaced 2π apart. Lower left is the lattice associated with sampling matrix V, which is defined in equation (23). The vectors correspond to the columns of **V**. Lower right is the DTFT of this lattice, defined in equation (24). The vectors correspond to the columns of $\tilde{\mathbf{V}}$.

Subsampling of Discrete Signals

Given a discrete image, we will be interested in the effects of subsampling and upsampling in the frequency domain. As with sampling, subsampling is parameterized by a square invertible matrix $\mathbf{K}$. The elements of the matrix, however, are constrained to be integers. The subsampled image is defined as

$$y[\boldsymbol{n}] = x[\mathbf{K}\boldsymbol{n}] = x_c(\mathbf{V}\mathbf{K} \cdot \boldsymbol{n}) \tag{28}$$

where we have made use of the definition in equation (27). The fraction of the original samples retained, or *subsampling density*, is $|\mathbf{K}|$, the determinant of the matrix $\mathbf{K}$.

As with the original sampling matrix $\mathbf{V}$, the subsampling matrix $\mathbf{K}$ is invariant to integral column operations. This fact, coupled with the requirement that the elements of $\mathbf{K}$ are integers allows us to write any subsampling matrix $\mathbf{K}$ in a *reduced* form as an upper-triangular matrix with the property that the off-diagonal elements are all less than the diagonal elements of the same row. The proof of this statement is given in appendix D. As an example, consider the following sequence of matrices which is generated by integer column operations:

$$\begin{bmatrix} 1 & 5 & 3 \\ 0 & 2 & 2 \\ 2 & 4 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 1 \\ 0 & 2 & 2 \\ 2 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 2 \\ 2 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

The final matrix is in reduced form.

As in one-dimensional signal processing, the effect of subsampling in the frequency domain corresponds to a sum of scaled and modulated copies of the

input spectrum. For a subsampled signal $y[\boldsymbol{n}]$ as in equation (28) above,

$$Y(\boldsymbol{\omega}) = \frac{1}{|\mathbf{K}|} \sum_{\boldsymbol{u} \in \mathbf{U}} X\left((\mathbf{K}^{-1})^t \boldsymbol{\omega} + 2\pi \boldsymbol{u}\right) \tag{29}$$

where

$$\mathbf{U} = \left\{(\mathbf{K}^{-1})^t \boldsymbol{n} \in [0,1)^d : \boldsymbol{n} \in \mathbf{N}^d\right\}$$

and where $[0,1)^d$ indicates the half-open unit interval in d dimensions. The equation indicates that every subsampling matrix $\mathbf{K}$ has associated with it a set of $|\mathbf{K}|$ *modulation vectors* $\mathbf{U} = \{\boldsymbol{u}_i\}$.

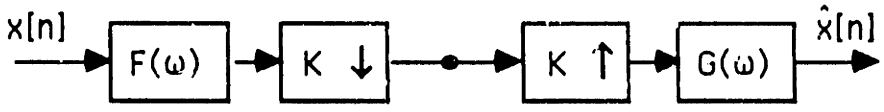
The effect of upsampling on the frequency domain is simpler to describe. If we define

$$y[\boldsymbol{m}] \equiv \begin{cases} x[\mathbf{K}^{-1}\boldsymbol{m}] & \boldsymbol{m} = \mathbf{K}\boldsymbol{n}, \quad \text{some } \boldsymbol{n} \in \mathbf{N} \\ 0 & \text{otherwise} \end{cases} \tag{30}$$

then the DTFT of $y[\boldsymbol{n}]$ will be

$$\begin{aligned} Y(\boldsymbol{\omega}) &= \sum_{\boldsymbol{n} \in \mathbf{N}} x[\boldsymbol{n}] e^{-j\boldsymbol{\omega}^t \mathbf{K}\boldsymbol{n}} \\ &= X(\mathbf{K}^t \boldsymbol{\omega}) \end{aligned} \tag{31}$$

The A/S system involves adjacent down and upsampling boxes, and so it is useful to combine the results above to get a response for a single A/S branch, as defined by the following diagram:



Combining equations (29) and (31) gives

$$\hat{X}(\boldsymbol{\omega}) = G(\boldsymbol{\omega}) \sum_{\boldsymbol{u} \in \mathbf{U}} F(\boldsymbol{\omega} + 2\pi \boldsymbol{u}) X(\boldsymbol{\omega} + 2\pi \boldsymbol{u}) \tag{32}$$

Thus, the spectrum of the subsampled signal is a sum of the input spectrum modulated by the vectors in the set **U**. The output of the entire A/S system will be a sum of such terms, one for each of the M branches. Clearly, for $d = 1$ this reduces to the one-dimensional equation given in (1).

In the sampled image domain, we may use equation (25) to write down an expression for a sampled image corresponding to the subsampled discrete image in equation (28):

$$x_c(\mathbf{r}) \cdot L_{\mathbf{VK}}(\mathbf{r}) = x_{\mathbf{VK}}(\mathbf{r})$$

and using equation (26), we may write the Fourier transform:

$$\frac{X_c(\Omega) * \tilde{L}_{\mathbf{VK}}(\Omega)}{|\mathbf{V}|\ |\mathbf{K}|} \tag{33}$$

Thus, the discrete subsampled image corresponds to sampling the original image using a lattice with a matrix which is the concatenation of the original sampling lattice and the subsampling lattice.

Two-Dimensional Systems

Having discussed the concepts of sampling and subsampling in multiple dimensions, we will now limit our attention to two-dimensional systems. When considering two-dimensional sampling matrices, we will assume without loss of generality that one of the vectors defining the matrix is the unit vector in the horizontal direction. For the purposes of this paper, we will only be interested in lattices which have reflection symmetry with respect to the x and y axes. Thus the form

of the sampling matrix must be

$$\mathbf{V} = \begin{bmatrix} 1 & a \\ 0 & b \end{bmatrix}$$

where $a \in \{0, \frac{1}{2}\}$ and b is any real number. We will denote the sampling patterns associated with these two choices of a as the *rectangular* and *staggered* sampling patterns, respectively. Some special cases which have additional symmetries are worth mentioning. The unit lattice is simply a rectangular lattice with $b = 1$. A *hexagonal* lattice is a staggered lattice with $b = \frac{\sqrt{3}}{2}$. This lattice corresponds to the centers of disks which are packed with maximal density in the plane. A *quincunx* lattice is a staggered lattice with $b = \frac{1}{2}$. This lattice is actually a rotated square lattice. Figure 25 illustrates the relationship between the discrete and sampling domains for a hexagonal sampling system.

Two-Band Systems

We consider first a two-band A/S system. We can easily list all reduced-form subsampling matrices with determinant two. Since they must be in upper triangular form, the determinant is the product of the two diagonal elements. Since the elements are integers, there are only two choices: $(1,2)$ and $(2,1)$ (negating either of these does not alter the lattice). Finally, if the upper left element is 2, the upper right element must be 0 or 1 according to proposition D. If the upper left element is 1, then the upper right element must be 0. So there are only three distinct reduced-form subsampling matrices with determinant two:

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}.$$

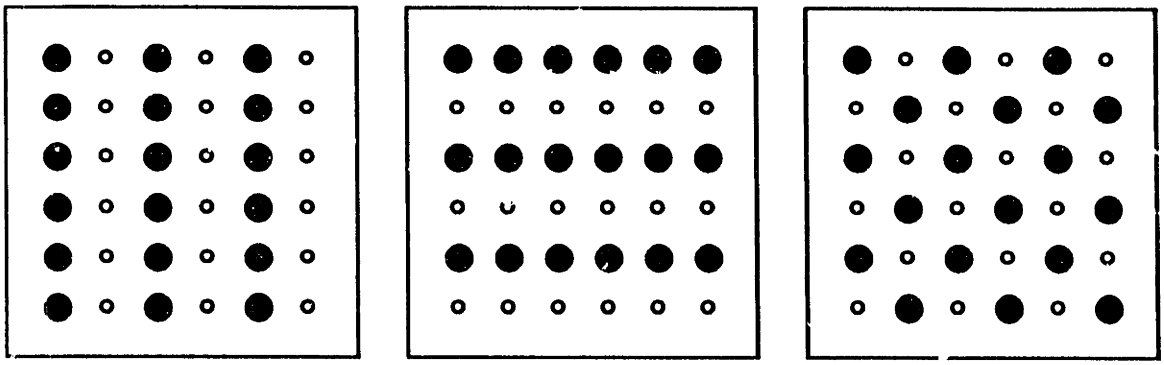


Figure 26: The three two-band subsampling schemes, shown on a square lattice. These are given in the same order as the matrices in the text.

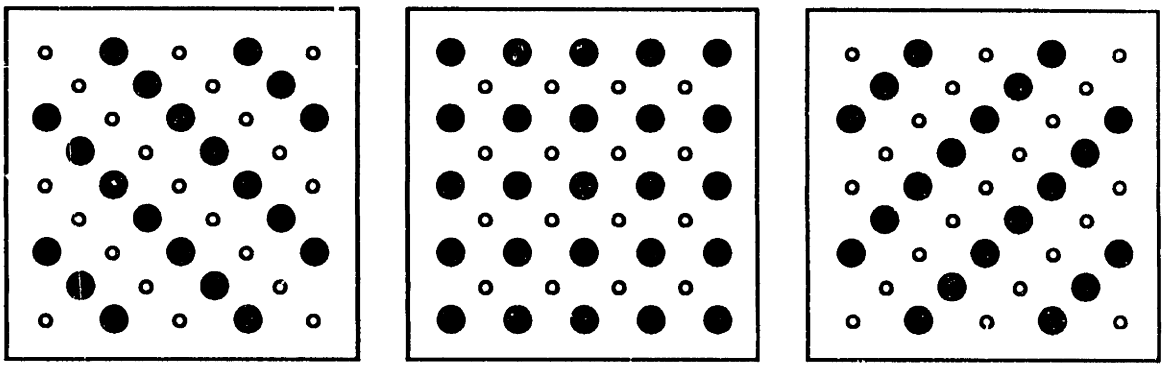


Figure 27: The three two-band subsampling schemes, shown on a quincunx lattice.

These are illustrated in figure 26 on the rectangular sampling lattice and in figure 27 on the staggered sampling lattice. The third sampling strategy has the nicest symmetry properties on the rectangular lattice, and so we will use this as an example. Note that when the initial sampling lattice is square, this subsampling produces a quincunx-sampled image. Similarly, the second subsampling strategy produces a unit-sampled image when the initial sampling lattice is quincunx.

Assume that we start with an image sampled on a unit lattice. Then using equation (33), the spectrum resulting from the subsampling given in the third matrix above is composed of the original spectrum convolved with a lattice

corresponding to the matrix

$$\tilde{\mathbf{V}}\tilde{\mathbf{K}} = 2\pi \begin{bmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{bmatrix}$$

Figure 28 illustrates the lattice in comparison to the original sampling lattice. It should be clear from the figure that the subsampled lattice may be written as a sum of the original spectrum with a copy of itself modulated by the vector $\begin{pmatrix} \pi \\ \pi \end{pmatrix}$. Thus, the set of modulation vectors defined in equation (29) is

$$\mathbf{U} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \right\}$$

Letting $\boldsymbol{u} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$, and assuming that $G_i(\omega) = F_i(-\omega)$, the overall A/S system response is

$$\hat{X}(\omega) = \frac{1}{2}\left[F_0(\omega)F_0(-\omega) + F_1(\omega)F_1(-\omega)\right]X(\omega)$$
$$+ \frac{1}{2}\left[F_0(\omega + 2\pi\boldsymbol{u})F_0(-\omega) + F_1(\omega + 2\pi\boldsymbol{u})F_1(-\omega)\right]X(\omega + 2\pi\boldsymbol{u})$$

As we did in the one-dimensional case, we may impose a relationship between the filters which eliminates aliasing and ensures that the basis functions are orthogonal to each other:

$$\begin{aligned} f_0[\boldsymbol{n}] &= h[\boldsymbol{n}] \\ f_1[\boldsymbol{n}] &= (-1)^{2\boldsymbol{u}^t\cdot\boldsymbol{n}+1}\left[-\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \boldsymbol{n}\right] \end{aligned}$$

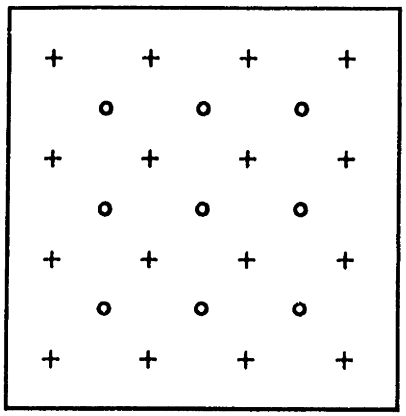


Figure 28: The frequency domain lattice for a two-band system. The locations of the original frequency domain lattice are indicated by a + sign. This illustrates the assertion that the lattice for the subsampled image is a sum of the original lattice and a shifted copy.

This choice cancels the aliasing term in equation and the remaining linear shift-invariant system response is

$$\hat{X}(\omega) = \frac{1}{2}\left[|F_0(\omega)|^2 + |F_0(\omega + 2\pi u)|^2\right.$$

This is illustrated in figure 30, along with the diamond-shaped partition of the frequency domain which is induced when $h[n]$ chosen to be a lowpass filter. Vetterli [58] derived a similar set of filters, but he assumed at the outset that the filters were symmetric with respect to inversion through the origin. The result above holds for symmetric or asymmetric filters. A small example filter was designed using the matrix averaging method of the previous section. A frequency surface plot of the lowpass filter is given in figure 29. The filter, and a more detailed description of the design method, are given in Appendix E.

Despite the fact that the subsampled images are in quincunx format, it is possible to use this A/S system in a cascade. We must spatially rotate the filters by $\pi/4$ in order to apply them to the quincunx image. Subsampling on the quincunx image is done according to the *second* subsampling matrix given

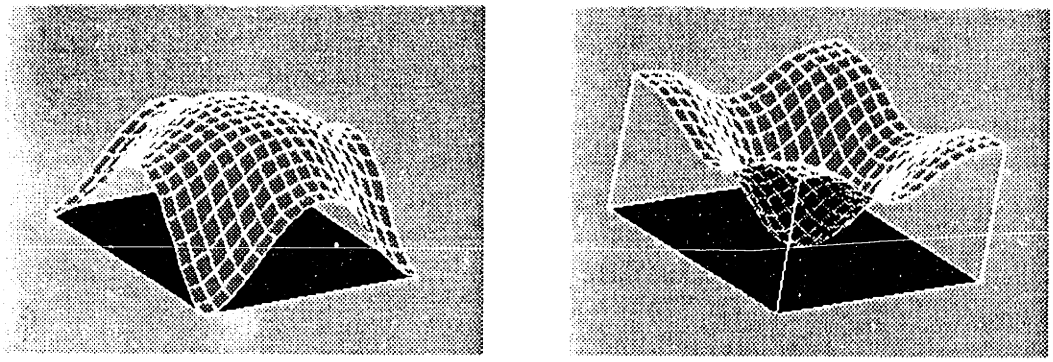


Figure 29: Frequency response for 2-band quincunx-subsampled system. The graphs extend from $-\pi$ to π in both x and y directions.

above, which will produce a square lattice as a result (see figure 27). A pyramid cascade of this sort of A/S system will produce a frequency domain partition in the form of nested diamonds and squares. This is illustrated in figure 31.

Four-band Systems

There are seven reduced-form subsampling matrices with determinant four:

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \quad \begin{bmatrix} 4 & 2 \\ 0 & 1 \end{bmatrix},$$

$$\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 4 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 4 & 3 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}.$$

These are illustrated in figure 32 on the rectangular sampling lattice and in figure 33 on the staggered sampling lattice. As an example, we will restrict our attention to the first of these, since this one maintains the symmetries of the original sampling scheme.

In the frequency domain, this subsampling lattice produces a sum of the

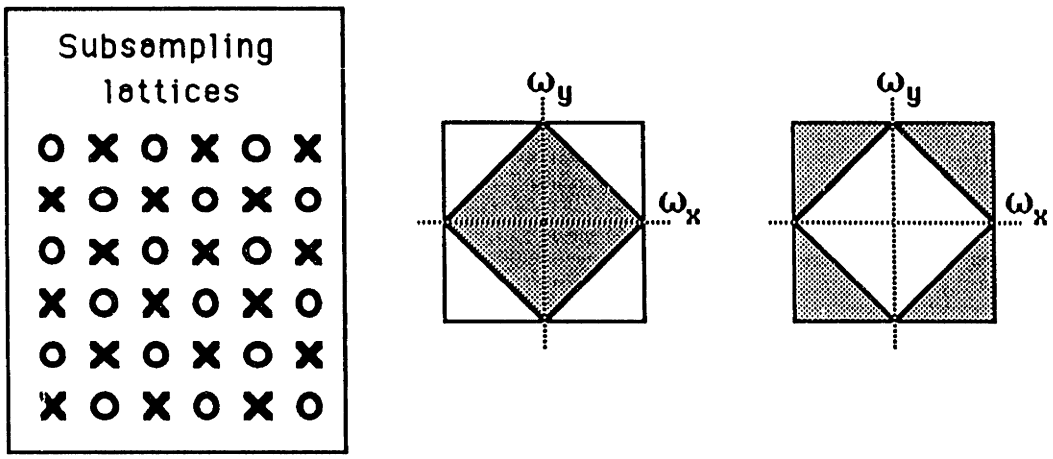


Figure 30: Illustration of the staggered sampling lattice and frequency domain partition for two-band system described above, applied to a unit-sampled image. The o's indicate that a lowpass sample is retained and the x's indicate that a highpass sample is retained.

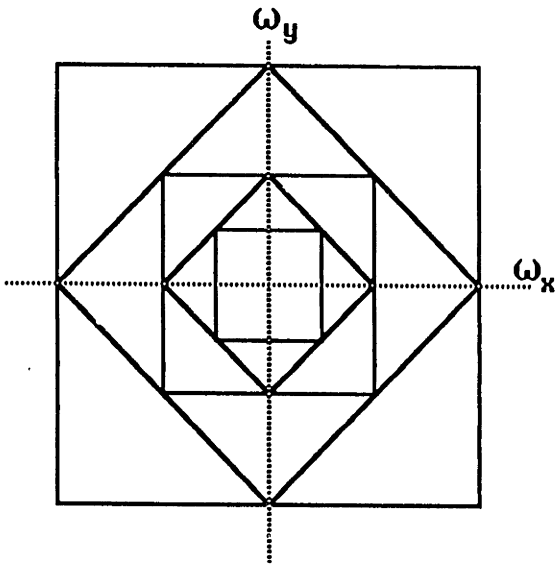


Figure 31: Nesting pattern in the frequency domain for a pyramid built on a two-band "quincunx" system.

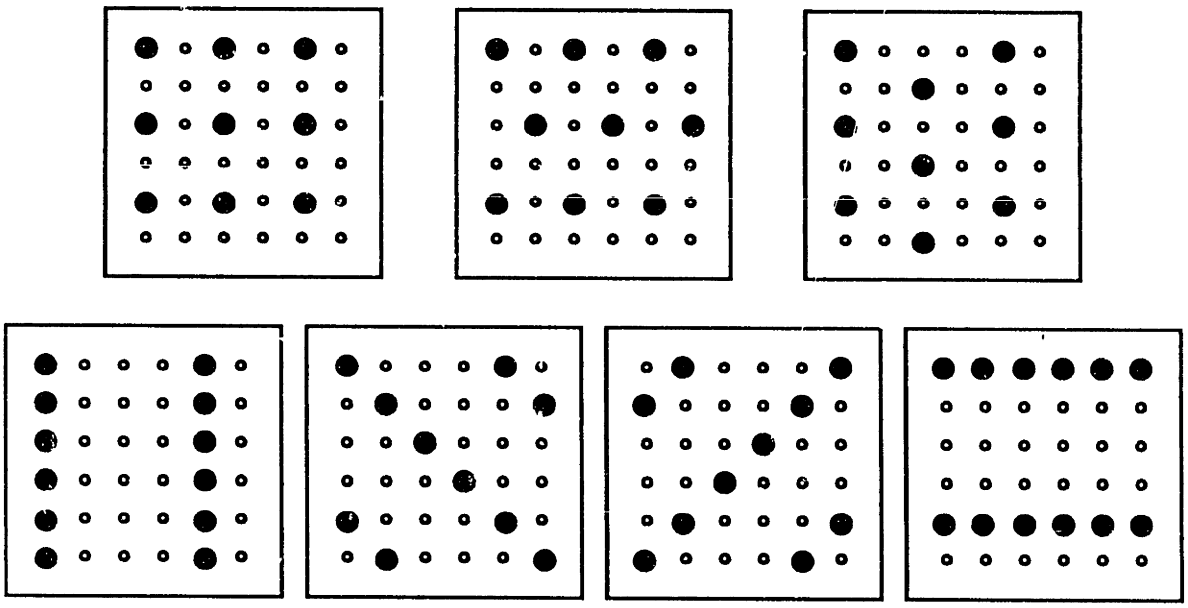


Figure 32: The seven four-band subsampling schemes, on a square lattice.

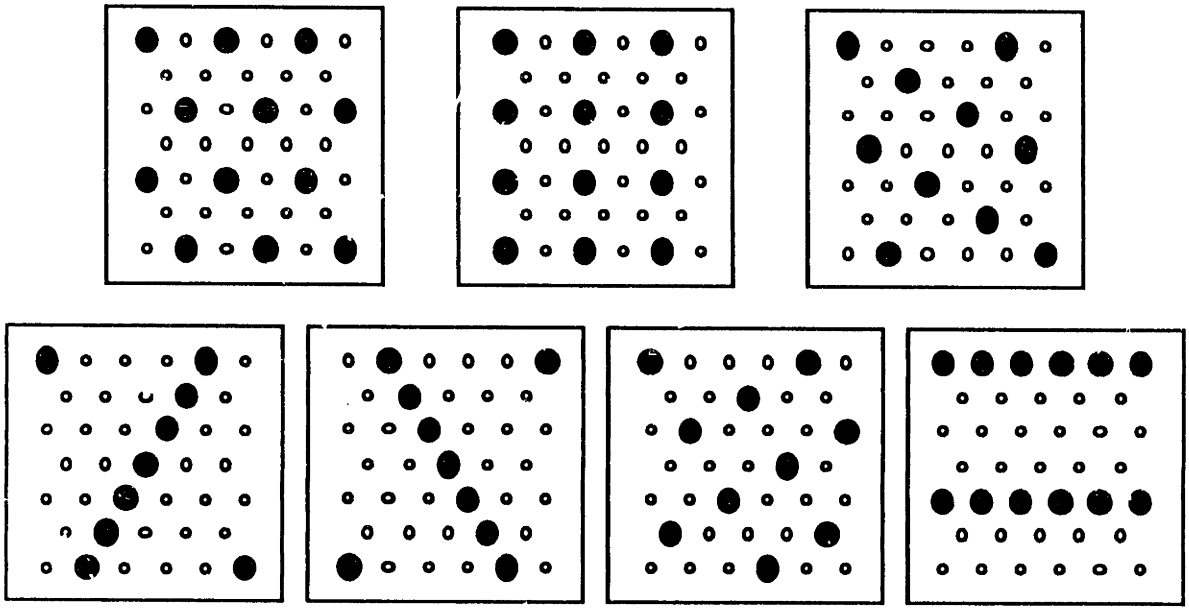


Figure 33: The seven four-band subsampling schemes, on a hex lattice.

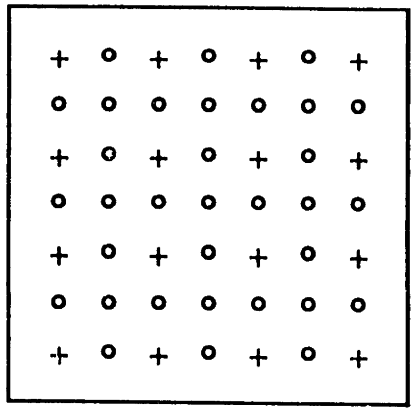


Figure 34: The frequency domain lattice for the first of the four-band systems. The locations of the original frequency domain lattice are indicated by a + sign. This illustrates the assertion that the lattice for the subsampled image is a sum of the original lattice and shifted copies of itself.

original spectrum modulated by the vectors

$$\mathbf{U} = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \right\} \equiv \{\mathbf{0}, \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$$

This is illustrated in figure 34.

It appears that simple modulated filter solutions do not exist for this case unless we constrain $F_0(\omega)$ to be symmetric. With this constraint we may make the following filter choices:

$$\begin{aligned} F_0(\omega) &= H(\omega) = H(-\omega) \\ F_1(\omega) &= e^{j\omega^t \cdot \boldsymbol{u}_1} H(-\omega + 2\pi \boldsymbol{u}_1) \\ F_2(\omega) &= e^{j\omega^t \cdot \boldsymbol{u}_3} H(-\omega + 2\pi \boldsymbol{u}_2) \\ F_3(\omega) &= e^{j\omega^t \cdot \boldsymbol{u}_2} H(\omega + 2\pi \boldsymbol{u}_3) \end{aligned} \tag{34}$$

As before, this choice of filters eliminates the overall system aliasing and ensures that the basis functions of the transformation will be orthogonal to each other.

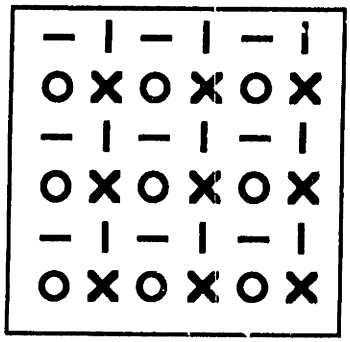


Figure 35: Illustration of the sampling grid for four-band system described in equation (34), applied to a square-sampled image. This should be compared to the separable example given in figure 14.

The resulting linear shift-invariant system response is:

$$\hat{X}(\omega) = \frac{1}{2}\Big[\sum_{i=0}^{3} |H(\omega + 2\pi u_i)|^2\Big] X(\omega)$$

In the rectangular sampling domain, this does *not* correspond to the separable solution because the vectors in the exponents of $F_2(\omega)$ and $F_3(\omega)$ have been switched. This is illustrated in figure 35, which should be compared to figure 14. This solution is not of practical interest for rectangular sampled images, since the separable solution is so much more efficient computationally. On a hexagonally sampled image, however, this solution has very nice symmetry properties. If the filter $H(\omega)$ is a lowpass filter, the other three filters will be highpass filters, each oriented with in one of the three hexagonal directions. This may be contrasted with the separable case, where one of the highpass filters contains both diagonal orientations. The sampling lattices and frequency domain partition for the hexagonal system is depicted in figure 36. An example filter, and an elaboration of the design method are given in Appendix E.

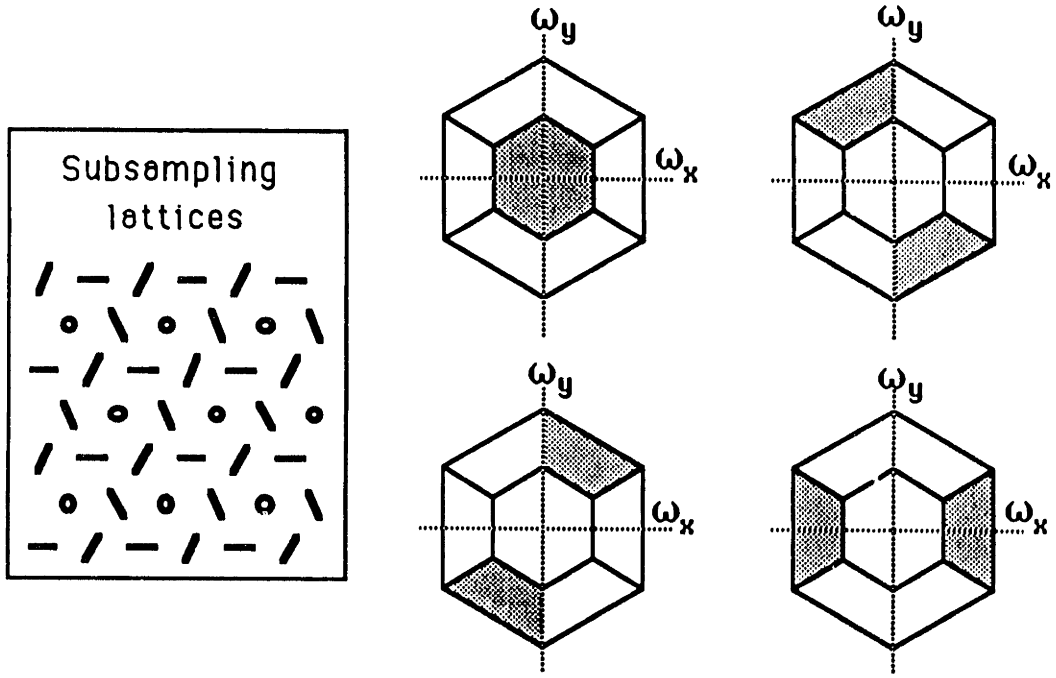


Figure 36: Illustration of the sampling grid and frequency domain partition for four-band system described in equation (34), applied to a hexagonally sampled image.

Three-dimensional Systems

Three dimensional systems are of interest for application to image sequences. We give a brief introduction to the problem. The sampling matrix in three dimensions may be written with the help of proposition D as

$$\mathbf{V} = \begin{bmatrix} 1 & a & c \\ 0 & b & d \\ 0 & 0 & e \end{bmatrix}$$

If we again assert the constraint of x and y reflection symmetry, then the parameter a must again satisfy $a \in \{0, \frac{1}{2}\}$. Adding the constraint of z-reflection symmetry makes things complicated, so for the purposes of this example, we will assume that $a = 0$. Then z reflection symmetry requires that $c, d \in \{0, \frac{1}{2}\}$. The case $(c, d) = (0, 0)$ is three-dimensional unit lattice. The pattern formed in

the case $(c,d) = \left(\frac{1}{2}, \frac{1}{2}\right)$ corresponds to the densest packing of spherical objects (such as marbles) and will be referred to as the *marble* lattice. This is analogous to the hexagonal case discussed above, which corresponds to the densest packing of disks in the plane. It is composed of square-sampled planes which are interleaved. Alternatively, this sampling structure may be rotated in space to give a pattern composed of interleaved hexagonally sampled planes. In the frequency domain, the spectrum of such a sampling scheme is replicated on another marble lattice. If one wishes to lowpass filter a continuous image to avoid the aliasing resulting from sampling with such a lattice, the frequency spectrum of the filter should be unity in a region in the shape of a *rhombic dodecahedron* [66], and zero elsewhere. In analogy with the four-band case described above, an interesting 8-band example occurs with the "separable" subsampling matrix:

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

This subsampling matrix preserves the sampling scheme (ie, the subsampled marble lattice image is again a marble lattice image). Preliminary studies indicate that using such a subsampling scheme, it is possible to construct an eight-band system with full aliasing cancellation. The resulting basis functions will be *spatio-temporally* oriented and will thus extract information about motion. Determination of the actual filter samples remains a topic for investigation.

Conclusions

This paper has investigated the design and application of linear transformations which are jointly localized in space and spatial frequency, and which explicitly represent information at multiple scales. In doing so, the advantages of a spatial domain matrix formulation have been emphasized, especially with regard to transform inversion and filter design. The filter design method of iterative inverse averaging was found to be a useful research tool for investigating the class of orthogonal linear transformations. In one dimension, the examples of orthogonal systems given in this paper are members of the class of QMF. We also discussed non-orthogonal systems. It was shown that these concepts may be extended to multi-dimensional systems with arbitrary regular sampling geometries. Of particular interest are the two-dimensional hexagonal and the three-dimensional "spherical packing" systems, which have appealing symmetry properties. For such systems, better filter designs are needed.

A 5-tap Filter Design

According to the expression in (18), a symmetric filter of length 5 requires an image size $N' \geq 2 \cdot 5 - 4 = 6$. Thus we design a 5-tap filter within an image of length 6. We write the filter samples as

$$h[n] = [\ a \quad b \quad c \quad b \quad a\ ]$$

Then the self-orthogonality equation (17) yields two constraint equations:

$$2a^2 + 2b^2 + c^2 \ = \ 1$$

and

$$2ac + b^2 + a^2 \ = \ 0$$

The first is a normalization equation, and the second is the "circular" inner product of the filter with itself shifted by two samples. We have three parameters and two equations, so we may we impose the additional constraint that the filter response is zero at $\omega = \pi$ (ie, the filter is lowpass):

$$c - 2b + 2a = 0$$

Combining these equations gives two possible solutions:

$$\begin{aligned} a &= (2\sqrt{2} \mp \sqrt{14})/12 \\ b &= \pm \sqrt{2}/4 \\ c &= 2b - 2a \end{aligned}$$

Analysis of the DTFT of these two filters indicates that the lower choice of signs leads produces a filter with very poor frequency localization. The filter described in the next appendix is thus the first solution.

B One-dimensional Filters

The filter impulse responses given in this appendix were designed using the frequency sampling method described in section 3. The odd-length filters were designed using an image size N' which was one larger than the desired filter size, and the frequency response at the point $\omega = \left\lfloor \frac{N'-2}{4} \right\rfloor$ was adjusted until the appropriate filter tap was zero. The frequency response at the DFT locations $w \in \{0, 1, \ldots, \left\lfloor \frac{N'-2}{4} \right\rfloor - 1\}$ was fixed at $H(\omega) = \sqrt{2}$, and if N' was divisible by 4, then the response at $\omega = \frac{N'}{4}$ was fixed at $H(\omega) = 1.0$. The even-length filters were designed using an image size which was equal to the desired filter size. The frequency response was fixed in the same manner as for the odd-length filters. The response at the point $\omega = \left\lfloor \frac{N'-2}{4} \right\rfloor$ was adjusted to produce the same measure of aliasing error as one of the odd-length filters (labelled with the letter "A") or as one of Johnston's filters (labelled with the letter "B"). Two of Johnston's filters are provided for reference and are labelled with a "J". Also given are a 15-tap and a 21-tap inverse filter for the linear interpolation filter that was described at the end of section 3.

5	7	9	11	13
-0.0761025	-0.0074972	0.0282204	0.0005612	-0.0145152
0.3535534	-0.0731952	-0.0603941	0.0244078	0.0211069
0.8593118	0.3610506	-0.0738819	-0.0558173	0.0406707
	0.8534972	0.4139475	-0.0732233	-0.0990339
		0.7984298	0.4088095	-0.0587709
			0.8047379	0.4314804
				0 7723375

Table 3: Odd-length filter tap values. Half of the impulse response sample values are shown for each of the normalised lowpass QMF filters (All filters are symmetric about $n = 0$). The appropriate highpass filters are obtained by delaying by one sample and multiplying with the sequence $(-1)^n$.

8A	8B	8J	12A	12B	12J
0.0042330	0.0138932	0.0132759	-0.0024175	-0.0056647	-0.0053876
-0.0545462	-0.0981376	-0.0999205	0.0165117	0.0266007	0.0266667
0.0545462	0.0981376	0.0981901	0.0019685	-0.0048733	-0.0038329
0.7028738	0.6932135	0.6929634	-0.1117252	-0.1185671	-0.1197755
			0.1141427	0.1242317	0.1251126
			0.6886266	0.6853794	0.6850152

Table 4: Even-length filter tap values. Half of the impulse response sample values are shown for each of the normalised lowpass QMF filters (All filters are symmetric about $n = -0.5$). The appropriate highpass filters are obtained by multiplying with the sequence $(-1)^n$.

3	15	21
0.40824829	-0.00150183	-0.00010665
0.81649658	-0.00450547	0.00031996
	0.01051268	0.07465681
	0.02553208	-0.00181309
	-0.06158052	-0.00437275
	-0.14865021	0.01055859
	0.35875590	0.02549003
	0.86761963	-0.06153892
		-0.14856482
		0.35865968
		0.86598770

Table 5: 3-tap linear interpolation filter, and two approximate inverse filters. Half of the impulse response sample values are shown for each of the normalized lowpass filters (All filters are symmetric about $n = 0$). The appropriate highpass filters are obtained by multiplying with the sequence $(-1)^n$.

C Orthogonal Edge-handling

In this appendix, we describe several methods of applying a filter near the boundaries of the image. Circular convolution or "wraparound" is the simplest solution to the problem. In this case, each edge of the image is associated with the opposite edge. This is often undesirable because it produces artificial discontinuities in image intensity at the "seams" between opposite edges. A method which is often preferable is to *reflect* the image through each of its boundaries, thus eliminating the discontinuity problem. The images presented in this paper were all processed using reflection at the boundaries.

For the separable orthogonal 4-band case, it is also possible to modify the reflection method in such a way that the resulting basis set retains its orthogonality. Consider a pair of one-dimensional filters $f_0[n]$ and $f_1[n]$ obeying the orthogonality equations given in (9) and (10). We then define the kth basis function $u_k[n]$ as follows. If k is even, then

$$u_k[n] = \begin{cases} f_0[k-n] + f_0[k+n], & n \neq 0 \\ \sqrt{2} f_0[k], & n = 0 \end{cases}$$

If k is odd, then

$$u_k[n] = \begin{cases} f_1[k-n] + f_1[k+n], & n \neq 0 \\ \sqrt{2} f_1[k], & n = 0 \end{cases}$$

Now consider the dot product of two of these basis functions. As an example, let $l \neq k$ both be even. Then the dot product is

$$\sum_{n=0}^{N-1} u_k[n] u_l[n] = 2 f_0[k] f_0[l] + \sum_{n=1}^{N-1} (f_0[k-n] + f_0[k+n])(f_0[l-n] + f_0[l+n])$$

$$
\begin{aligned}
&= \sum_{n=-N+1}^{N-1} f_0[k-n] f_0[l-n] + \sum_{n=-N+1}^{N-1} f_0[k+n] f_0[l-n] \\
&= \sum_{m} f_0[k-l+m] f_0[m] + \sum_{m} f_0[k-l+m] f_0[m] \\
&= 0 + 0
\end{aligned}
$$

by equation (9). The cases for k and/or l odd may be demonstrated in exactly the same manner. Since equations (9) and (10) provide sufficient conditions for the orthogonality of the A/S system matrix, the basis set formed with this sort of filter reflection at the image boundaries is orthogonal. Unfortunately, the factor of $\sqrt{2}$ at the edge filter sample means that the highpass or lowpass characteristics of the basis functions are not preserved. This reflection method was used only to compute the coding gains over PCM (see section 4), where an orthogonal basis set is required.

D Reduced-form Subsampling Matrices

In this appendix, we prove a proposition which allows us to write any subsampling matrix in a *reduced* form.

Proposition 6 *Any subsampling matrix* $\mathbf{K}$ *may be rewritten in upper-triangular form such that the elements in the* i*th row obey*

$$0 \leq \mathbf{K}_{ij} \leq \mathbf{K}_{ii} - 1, \qquad \textit{for } i > j.$$

Proof: Given an upper triangular matrix, it is not difficult to see that we can use column operations to make the expression above hold. For each row i from the $d-1$th to the first, we add multiples of that row to all rows to its right until the ith row of each of those columns is between 0 and $\mathbf{K}_{ii}$. Now we must show that the matrix may be reduced to upper triangular form. We prove this statement by induction.

- Let $d = 2$. Then we write

$$\mathbf{K} = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

  If b divides d or vice versa, we are done, since we may then eliminate the larger of the two by adding an integer multiple of the smaller. Otherwise, assume without loss of generality that $b < d$. Then subtract the first column multiplied by $\left\lfloor \frac{d}{b} \right\rfloor$ from the second column. This will create a new subsampling matrix with the lower right element strictly less than b. We repeat this procedure, each time reducing the minimum of the bottom two matrix elements, until one of them divides the other. The iteration

is limited since the minimum is a strictly decreasing integer and so will eventully reach zero.

- Now assume the statement holds for a $d \times d$ subsampling matrix $\mathbf{K}$. Let $\mathbf{K}'$ be any $(d+1) \times (d+1)$ subsampling matrix. Consider the $d \times (d+1)$ matrix, $\mathbf{A}$, which contains all but the top row of $\mathbf{K}'$. We will show that one column of $\mathbf{A}$ may be reduced to the zero vector using integer column operations. This will complete the proof, since the remaining d columns of $\mathbf{A}$ form a $d \times d$ matrix which can be reduced to upper diagonal form by the inductive assumption.

  The $d+1$ columns $\boldsymbol{v}_i$ of $\mathbf{A}$ must be linearly dependent and so there exist $d+1$ integers n_i such that

$$\sum_{i=0}^{d} n_i \boldsymbol{v}_i = 0 \tag{35}$$

  If one of the n_i, say n_j divides all of the others, we are done. We may assume that $n_j = 1$ since the equation above will still hold when all of the n_i are divided by n_j. Then we simply add $-n_i$ copies of each column vector $\boldsymbol{v}_i$ to column vector $\boldsymbol{v}_j$, thus eliminating it. If there is no n_i dividing all of the others, then write n_s for the smallest of the n_i and n_r for one of the n_i which is *not* divisible by n_s. Now we add $n_s \left\lfloor \frac{n_r}{n_s} \right\rfloor \boldsymbol{v}_r$ to vector $\boldsymbol{v}_s$. Equation (35) now becomes

$$\sum_{i \neq r,s} n_i \boldsymbol{v}_i + \left(n_r - n_s \left\lfloor \frac{n_r}{n_s} \right\rfloor \right) \boldsymbol{v}_r + n_s \boldsymbol{v}_s + n_s \left\lfloor \frac{n_r}{n_s} \right\rfloor \boldsymbol{v}_r$$

  The new coefficient of the vector $\boldsymbol{v}_r$ is $n_r' = n_r - n_s \left\lfloor \frac{n_r}{n_s} \right\rfloor < n_r$. Thus we have reduced the smallest n_i. We may repeat this process until the smallest n_i divides the others.

This completes the inductive proof. Incidentally, this theorem also applies to the original sampling matrix $\mathbf{V}$. ∎

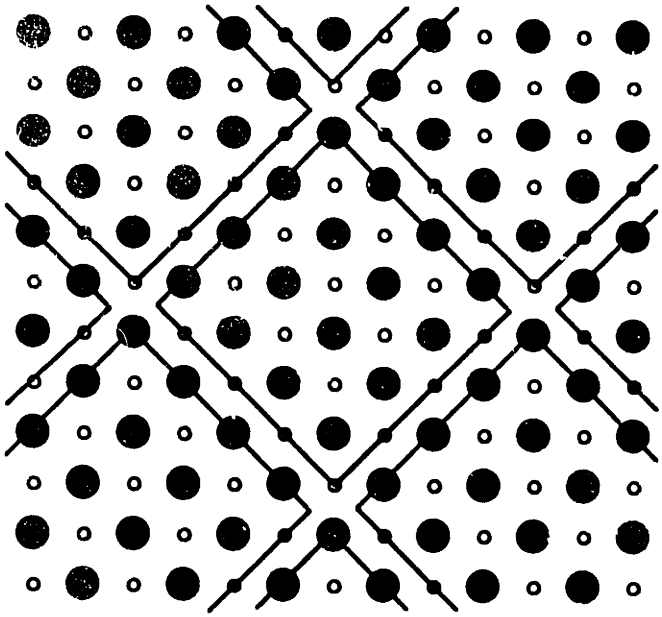


Figure 37: The design "image" for the two-band quincunx system. See text for explanation.

E Non-separable Two-dimensional Filters

The matrix averaging design method is easily applied to the multi-dimensional design problem. The center region indicated in each of the diagrams given below indicates the "image" in which the quincunx and hexagonal filters described in section 5 were designed. Each of these images is surrounded by neighboring copies of itself to illustrate how the circular convolution is to be done.

Consider the quincunx case. The image region is chosen to produce a diamond-shaped filter seven samples wide (from corner to corner). Such a filter fits into the image leaving a border of samples analogous to the leftover sample in the one-dimensional odd-length design case. Since the desired filter is smaller than the image, the initial filter choice must be adjusted to minimize

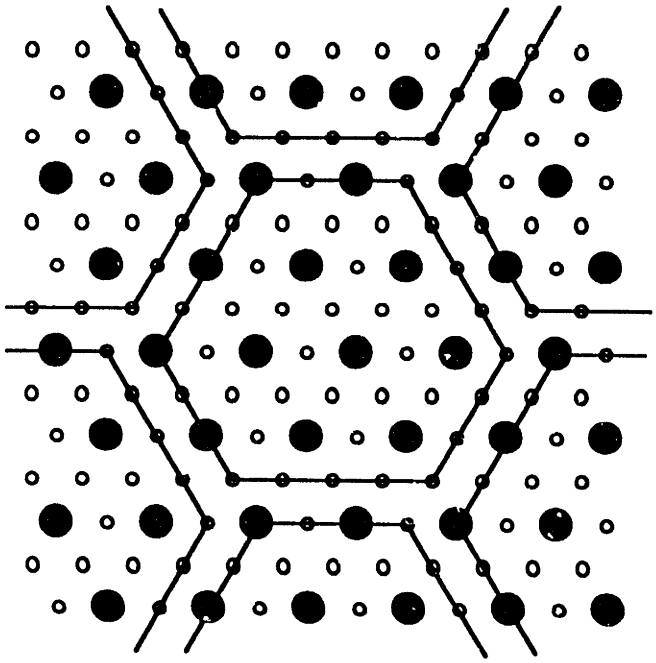


Figure 38: The design "image" for the four-band hexagonal system.

the resulting values contained in the unwanted samples. The sixteen blackened circles within the image indicate the samples which are retained after subsampling. The the samples indicated by small circles are discarded. A column of the analysis matrix $\mathbf{F}$ is obtained by placing the impulse response samples of an initial lowpass filter into the the image, centered on one of the blackened circles. All other samples of the image are set to zero. The image samples are then read off in an arbitrarily fixed order to form a column vector of height 32 which may be placed into the matrix $\mathbf{F}$. This procedure is repeated for each of the sixteen blackened disks, producing a 32×16 matrix. The complete A/S system matrix also includes the highpass basis functions, but since the system has been designed so that the initial high and lowpass basis functions span orthogonal spaces, proposition 4 of section 3 allows us to use the iterative procedure on the matrix containing only lowpass filters. The hexagonal system matrix is formed

in the same manner.

			e							-e				
		f	c	f					-f	c	-f			
	f	d	b	d	f			-f	d	-b	d	-f		
e	c	b	a	b	c	e	-e	c	-b	a	-b	c	-e	
	f	d	b	d	f			-f	d	-b	d	-f		
		f	c	f					-f	c	-f			
			e							-e				

Table 6: An example of two-band quincunx filters.

The following are some example filter sample values for the quincunx case:

$$
\begin{aligned}
a &= 0.9271061 \\
b &= 0.1828738 \\
c &= -0.0182986 \\
d &= -0.0366835 \\
e &= -0.0016742 \\
f &= -0.0022018
\end{aligned}
$$

			e		f		f		e			
		f		c		d		c		f		
	f		d		b		b		d		f	
e		c		b		a		b		c		e
	f		d		b		b		d		f	
		f		c		d		c		f		
			e		f		f		e			

			-e		-f		-f		-e			
		f		c		d		c		f		
	-f		-d		-b		-b		-d		-f	
e		c		b		a		b		c		e
	-f		-d		-b		-b		-d		-f	
		f		c		d		c		f		
			-e		-f		-f		-e			

Table 7: An example of four-band hexagonal filters. Only the lowpass and the horizontal highpass filters are shown. The other two highpass filters are identical to the horizontal, rotated by $2\pi/3$ and $4\pi/3$.

The following are some example filter sample values for the hexagonal case:

$$
\begin{aligned}
a &= 0.6935061 \\
b &= 0.2905438 \\
c &= -0.0323747 \\
e &= -0.0027388 \\
d &= -0.0319443 \\
f &= -0.0028679
\end{aligned}
$$

References

[1] David Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* W. H. Freeman and Company, San Fransisco, 1982.

[2] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden. Pyramid methods in image processing. *RCA Engineer*, 29(6):33–41, November/December 1984.

[3] David Marr, Tomaso Poggio, and Shimon Ullman. Bandpass channels, zero-crossings, and early visual information processing. *J. Opt. Soc. Am.*, 69:914–916, 1977.

[4] S. G. Mallat. *A theory for multiresolution signal decomposition: the wavelet representation.* GRASP Lab Technical Memo MS-CIS-87-22, University of Pennsylvania, Department of Computer and Information Science, 1987.

[5] Andrew P. Witkin. Scale-space filtering. In *Proc. Int. Joint Conf. Artificial Intelligence*, pages 1019–1021, 1985.

[6] Jan J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[7] Alex Pentland. Fractal based description of natural scenes. *IEEE Trans. PAMI*, 6(6):661–674, 1986.

[8] S. Marcelja. Mathematical description of the response of simple cortical cells. *J. Opt. Soc. Am.*, 70:1297–1300, 1980.

[9] F. W. Campbell and J. G. Robson. Application of fourier analysis to the visibility of gratings. *J. Physiology*, 197:551–566, 1968.

[10] J. G. Robson. *Visual Coding and Adaptablility*, chapter Neural images: The physiological basis of spatial vision. Lawrence Erlbaum, Hillsdale, NJ, 1980.

[11] L. Maffei. The visual cortex as a spatial frequency analyzer. *Vision Research*, 13:1255–1267, 1973.

[12] J. J. Kulikowski, S. Marcelja, and P. O. Bishop. Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex. *Biological Cybernetics*, 43:187–198, 1982.

[13] Martin J. Bastiaans. Gabor's signal expansion and degrees of freedom of a signal. *Optica Acta*, 29(9):1223–1229, 1982.

[14] John G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A*, 2(7):1160–1169, July 1985.

[15] John G. Daugman and Daniel M. Kammen. Pure orientation filtering: a scale-invariant image-processing tool for perception research and data compression. *Behavior Research Methods, Instruments, & Computers*, 18(6):559–564, 1986.

[16] M. R. Turner. Texture discrimination by gabor functions. *Biological Cybernetics*, 55:71–83, 1986.

[17] David J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, 4(12):2379–2394, December 1987.

[18] Dennis Gabor. Theory of communication. *J IEE*, 93:492–457, 1946.

[19] R. J. Clarke. *Transform Coding of Images. Microelectronics and Signal Processing Series*, Academic Press, London, 1985.

[20] N.S. Jayant and Peter Noll. *Digital Coding of Waveforms. Signal Processing Series*, Prentice-Hall, Englewood Cliffs, NJ, 1984.

[21] Robert M. Lerner. *Lectures on Communication System Theory*, chapter Representation of signals. McGraw-Hill, New York, 1961.

[22] Alan V. Oppenheim and Ronald W. Schafer. *Digital Signal Processing.* Prentice-Hall, Inc., Englewood Cliffs, 1975.

[23] Dan E. Dudgeon and Russell M. Mersereau. *Multidimensional Digital Signal Processing. Signal Processing Series*, Prentice-Hall, Englewood Cliffs, NJ, 1984.

[24] Peter J. Burt. Fast filter transforms for image processing. *Computer Graphics and Image Processing*, 16:20–51, 1981.

[25] P. J. Burt and Edward H. Adelson. The laplacian pyramid as a compact image code. *IEEE Trans. Communications*, COM-31(4):532–540, April 1983.

[26] A. Croisier, D. Esteban, and C. Galand. Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques. In *International Conference on Information Sciences and Systems*, pages 443–446, Patras, August 1976.

[27] D. Esteban and C. Galand. Application of quadrature mirror filters to split band voice coding schemes. In *Proceedings ICASSP*, pages 191–195, 1977.

[28] R. E. Crochiere and L. R. Rabiner. *Multirate Digital Signal Processing. Signal Processing Series*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[29] Gilbert Strang. *Linear Algebra and Its Applications.* Academic Press, Orlando, 1980.

[30] P. P. Vaidyanathan. Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques. *IEEE ASSP Magazine*, 4–20, July 1987.

[31] Martin Vetterli. A theory of multirate filter banks. *IEEE Trans. ASSP*, ASSP-35(3):356–372, March 1987.

[32] D. Esteban and C. Galand. HQMF: halfband quadrature mirror filters. In *Proceedings ICASSP*, pages 220–223, 1981.

[33] C. R. Galand and D. J. Esteban. Design and evaluation of parallel quadrature mirror filters (PQMF). In *Proceedings ICASSP*, pages 224–227, 1983.

[34] Joseph H. Rothweiler. Polyphase quadrature mirror filters – a new subband coding technique. In *Proceedings ICASSP*, pages 1280–1283, 1983.

[35] Claude R. Galand and Henri J. Nussbaumer. New quadrature mirror filter structures. *IEEE Trans. ASSP*, ASSP-32(3):522–531, June 1984.

[36] Peter L. Chu. Quadrature mirror filter design for an arbitrary number of equal bandwidth channels. *IEEE Trans. ASSP*, ASSP-33(1):203–218, February 1985.

[37] H. J. Nussbaumer and M. Vetterli. Pseusdo quadrature mirror filters. In *Proceedings Int. Conf. on DSP*, pages 8–12, Elsevier Science Publishers B.V. (North-Holland), 1984.

[38] H. J. Nussbaumer and M. Vetterli. Computationally efficient QMF filter banks. In *Proceedings ICASSP*, pages 11.3.1–11.3.4, 1984.

[39] M. J. T. Smith and T. P. Barnwell, III. A procedure for designing exact reconstruction filter banks for tree-structured subband coders. In *Proceedings ICASSP*, pages 27.1.1–27.1.4, 1984.

[40] Gunter Wackersreuther. On the design of filters for ideal QMF and polyphase filter banks. *Arch. Elekt. Ubertrag*, 39:123–130, 1985.

[41] Mark J. T. Smith and Thomas P. Barnwell, III. Exact reconstruction techniques for tree-structured subband coders. *IEEE Trans. ASSP*, ASSP-34(3):434–441, June 1986.

[42] Claude R. Galand and Henri J. Nussbaumer. Quadrature mirror filters with perfect reconstruction and reduced computational complexity. In *Proceedings ICASSP*, pages 14.3.1–14.3.3, 1985.

[43] Tor A. Ramstad. Analysis/synthesis filterbanks with critical sampling. In *Proceedings Int. Conf. on DSP*, pages 130–134, Elsevier Science Publishers B.V. (North-Holland), 1984.

[44] Mark J. T. Smith and Thomas P. Barnwell, III. A unifying framework for analysis/synthesis systems based on maximally decimated filter banks. In *Proceedings ICASSP*, pages 14.2.1–14.2.4, March 1985.

[45] M. Vetterli. Splitting a signal into subsampled channels allowing perfect reconstruction. In *Proc. IASTED Conf. Appl. Signal Processing and Digital Filtering*, Paris, June 1985.

[46] Martin Vetterli. A theory of filter banks. In *Signal Processing III: Theories and Applications (Proc. EUSIPCO-86)*, pages 61–64, Elsevier Science Publishers B.V. (North-Holland), 1986.

[47] Martin Vetterli. Filter banks allowing perfect reconstruction. *Signal Processing*, 10(3):219–244, April 1986.

[48] P. P. Vaidyanathan. Theory and design of M-channel maximally decimated quadrature mirror filters with arbritrary M, having the perfect-reconstruction property. *IEEE Trans. ASSP*, ASSP-35(4):476–492, April 1987.

[49] P. P. Vaidyanathan and Phuong-Quan Hoang. The perfect-reconstruction QMF bank: new architectures, solutions, and optimization strategies. In *Proceedings ICASSP*, pages 50.3.1–50.3.4, 1987.

[50] Tor A. Ramstad and Ole Foss. Sub-band coder design using recursive quadrature mirror filters. In *Signal Processing: Theories and Applications (Proc. EUSIPCO-80)*, pages 747–752, North-Holland Publishing Company, 1980.

[51] Thomas P. Barnwell, III. Subband coder design incorporating recursive quadrature filters and optimum ADPCM coders. *IEEE Trans. ASSP*, ASSP-30(5):751–765, October 1982.

[52] P. C. Millar. Mirror filters with minimum delay responses for use in subband coders. In *Proceedings ICASSP*, pages 11.5.1–11.5.4, 1984.

[53] R. Ansari and C.-L. Lau. Two-dimensional IIR filters for exact reconstruction in tree-structured sub-band decomposition. *Electronics Letters*, 23(12):633–634, June 1987.

[54] P. P. Vaidyanathan. On power-complementary FIR filters. *IEEE Trans. Circuits and Systems*, CAS-32(12):1308–1310, December 1985.

[55] J. D. Johnston. A filter family designed for use in quadrature mirror filter banks. In *Proceedings ICASSP*, pages 291–294, 1980.

[56] Vijay K. Jain and Ronald E. Crochiere. A novel approach to the design of analysis/synthesis filter banks. In *Proceedings ICASSP*, pages 5.10–5.10, 1983.

[57] Vijay K. Jain and Ronald E. Crochiere. Quadrature mirror filter design in the time domain. *IEEE Trans. ASSP*, ASSP-32(2):353–360, April 1984.

[58] Martin Vetterli. Multi-dimensional sub-band coding: some theory and algorithms. *Signal Processing*, 6(2):97–112, February 1984.

[59] John W. Woods and Sean D. O'Neil. Subband coding of images. *IEEE Trans. ASSP*, ASSP-34(5):1278–1288, October 1986.

[60] H. Gharavi and Ali Tabatabai. Sub-band coding of digital images using two-dimensional quadrature mirror filtering. In *Proceedings of SPIE*, pages 51–61, 1986.

[61] H. Gharavi and A. Tabatabai. Application of quadrature mirror filters to the coding of monochrome and color images. In *Proceedings ICASSP*, pages 32.8.1–32.8.4, 1987.

[62] Anh Tran, Kwun-Min Liu, Kou-Hu Tzou, and Eileen Vogel. An efficient pyramid image coding system. In *Proceedings ICASSP*, pages 18.6.1–18.6.4, 1987.

[63] W. Chen and W.K. Pratt. Scene adaptive coder. *IEEE Trans. Communications*, COM-32(3):225–232, March 1984.

[64] William F. Schreiber and Donald E. Troxel. Transformation between continuous and discrete representations of images: a perceptual approach. *IEEE Trans. PAMI*, 7(2):178–186, March 1985.

[65] Soo-Chang Pei and Sy-Been Jaw. Design of 2-D quadrature mirror FIR filters for image subband coding. *IEEE Trans. Circuits and Systems*, CAS-34(4):438–441, April 1987.

[66] Hermann Weyl. *Symmetry*. Princeton University Press, Princeton, NJ, 1952.