Matrix Multiplication via Wavelets

Daniel Beatty

March 20, 2004

Contents

Chapter 1

Introduction

Some overwhelming questions that drive computational science are how fast can the answer be computed, how accurate is the answer, and how stable is the method for obtaining the answer. In this thesis, these questions are applied to matrix multiplication. Of course there are already conventional algorithms to compute matrix multiplication. This thesis contributes a simple analysis of how the wavelet operator can be applied to matrix multiplication.

The two most desired qualities in the computation of matrices are sparseness and the condition number. Sparseness for a matrix states that a good majority of the elements in such a matrix are zero. The condition number indicates how accurately an operation will be performed on the system. Wavelets contribute to numerical methods by providing a stable preconditioning technique which produces a more sparse and better conditioned than the original matrix. Various forms of the wavelet transform are key to applying wavelets as a preconditioning tool.

Matrix multiplication has applications in simulations, computer vision, and almost all areas of computational science. Classic matrix multiplication has a computational complexity of order N^3 operation, which makes it costly in the number of instructions that are needed to carry out the operation. Faster matrix multiplication techniques depend on the matrix being sparse to reduce the number elements that need to be multiplied. As shown in the results section, there exists various levels of acceptable sparseness that are generated by the wavelet transform.

For matrix multiplication, wavelets provide a preconditioning method that transforms a dense matrix into a sparse one. Thus sparse matrix multiplication can be applied more effectively to general class of matrices. Wavelet preconditioning may or may not help for matrices that are already sparse.

In this thesis, an overview is provided to define wavelets and how to apply them. This overview uses image processing to demonstrate the qualities of a two-dimensional wavelet transform. The next chapter describes the wavelet matrix multiplication procedure. Chapter 5 demonstrates matrix multiplication in the wavelet domain and shows the results of different threshold levels on the fidelity of the resulting product. Finally, conclusions are presented.

Chapter 2

Overview on Wavelets

Wavelets evolved from atomic function theory where they were developed as basic atoms or building blocks of all functions. This chapter provides an overview of wavelets. This is done first by providing the standard definitions and concepts of the wavelet basis function, the wavelet transform and multiresolution representation. Then a numerical example is utilized to demonstrate the concepts. The next section details the numerical implementation of the wavelet transform. The last section represents the 2D wavelet transform.

2.1 The Wavelet Basis Function

This section provides the basic properties of a wavelet function. It first describes some general properties of functions. Then it presents the translation and dilation properties which can be used to build an entire bases. Finally, it presents the wavelet function itself.

2.1.1 Function Properties

There are a number of different properties that can be used to classify functions. These include integrability, symmetry, compactness and orthogonality. The properties will be introduced now and used in the rest of the chapter.

A function, $f(\cdot)$, is *square integrable* if its L_2 norm is finite. This can be expressed by

$$f(x) \in L_2(\mathbb{R}) \quad \text{if} \quad \int_{-\infty}^{\infty} |f(x)|^2 \, dx < \infty.$$

A one-dimensional function is *symmetric* about the y-axis if

$$f(x) = f(-x)$$

and it is *antisymmetric* if

$$f(x) = -f(-x).$$

A set is considered *compact* in the n-dimensional real space, $\mathbb{R}^n$ if it is both closed and bounded. A function has *compact support* if it is zero outside of a compact set. This implies that there exists n-dimensional sphere, S^n, where

$$f(x) = 0 \qquad \forall x \notin S^n.$$

Orthogonality is another concept necessary for this thesis. Two different basis functions are *orthogonal* if their inner product is zero. The inner product of two functions, f and g, can be represented by

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)g(x)\ dx.$$

They are *orthonormal* if they are both orthogonal and have a norm of one.

2.1.2 Translation and Dilation

Translation shifts the basis function along the variable axis. The translation of a function, f, can be represented by f_k where

$$f_k(t) = f(t-k).$$

Notice that translation does not alter the shape of the function, only the position of the function along the number space of t [31].

Dilation (also called contraction) transforms a one dimensional function in width, and the output in height. The dilation of a function, f, can be presented by $f_{k,j}$ where

$$f_k^{\alpha}(t) = kf(\alpha t).$$

If $k > 1$ then the height of the function is increased. If $\alpha < 1$ then the width of the function is increased. Otherwise it is decreased.

These definitions can be combined into the translation-dilation representation upon dyadic intervals. A function f can be expressed by a dilation of j and a translation of k through

$$f_{j,k}(t) = 2^{j/2} f(2^j t - k) \qquad \forall j, k \in \mathbb{Z}.$$

A function is said to have the orthonormal translation-dilation property when for two translation-dilations, $f_{j,k}$ and $f_{l,m}$, the result can be expressed by [1]

$$\langle f_{j,k}, f_{l,m}\rangle = \delta_{j,l}\delta_{k,m}.$$

2.1.3 The Wavelet Basis Function

The two mandatory properties of a wavelet basis function are:

- it must square integrable, and
- must have a zero average, i.e.:

$$\int_{-\infty}^{+\infty} \psi(x)\ dx = 0.$$

In addition, the first generation wavelets were restricted to satisfy the orthonormal translation-dilation property. This restriction has been removed through the lifting scheme [**?**]. An example of a strict definition of a wavelet is given by Charles Chui [1]:

> "A function $\psi \in L_2(\mathbb{R})$ is called an orthonormal wavelet if the family $\{\psi_{j,k}\}$ defined
>
> $$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)\forall j, k \in Z$$
>
> is an orthonormal basis of $L_2(\mathbb{R})$ where $\langle\psi_{j,k}, \psi_{l,m}\rangle = \delta_{j,l}\delta_{k,m}$, $\forall j, k, l, m \in Z$ and every $f \in L_2(\mathbb{R})$ can be written as
>
> $$f(x) = \sum_{j,k=-\infty}^{\infty} c_{j,k}\psi_{j,k}(x)$$

where the series convergences and is in $L_2(\mathbb{R})$ such that

$$\lim_{M_1,M_2,N_1,N_2} ||f - \sum_{j=-M_2}^{N_2} \sum_{k=-M_1}^{N_1} c_{j,k}\psi_{j,k}|| = 0$$

The simplest example of orthonormal wavelets is the Haar Transform."

There is more than one wavelet basis function. Additional classifications used include the symmetry and compactness. The Haar Wavelet Basis function actually fulfills the strictest definition of a wavelet basis function, and has additional properties. The Haar Wavelet Basis Function has compact support. It is also symmetric. Furthermore, it has been stated by Walker[**?**] and Chui [1] that the Haar Wavelet Basis Function is the only wavelet basis function in $L_2(\mathbb{R})$ that is orthogonal, compact and symmetric. The Haar wavelet basis function is defined below.

$$\psi(x) = \begin{cases} 1 & 0 \leq x < \frac{1}{2} \\ -1 & \frac{1}{2} \leq x < 1 \\ 0 & otherwise \end{cases} \tag{2.1}$$

2.2 The Wavelet Transform

The wavelet transform is comprised of two parts. An average sample and a difference sample. This is expressed by first looking at pairs of wavelets and then looking at the actual representation.

2.2.1 Wavelet Pairs: The Averaging Basis Function

The Wavelet Basis Function section defined the Haar Wavelet Basis Function defined the Haar Wavelet Basis function, in equation 2.1. However, there exists a concept of a wavelet pair. These pairs exist as averaging and differencing basis. The wavelet basis function and differencing basis are synonymous. The averaging basis concept was derived from classic multi-resolution which is described in section 2.3.

This section only defines the wavelet basis pair in terms of a wavelet averaging basis and shows an example with the Haar Averaging Basis Function. A wavelet pair is a set of two basis functions containing one wavelet basis function and one averaging basis function which meet the following criteria:

- both must be square integrable,
- both must satisfy the orthonormal translation-dilation property, and
- the wavelet basis function must be orthogonal with the average basis function.

The simplest form of the averaging filter is the Haar Averaging Filter, and it is a pair to the Haar Wavelet Basis Function. Like the Haar Wavelet Basis Function, the Haar Averaging Filter also satisfies the orthogonal translation-dilation property and is in $L_2(\mathbb{R})$ The mother function for the Haar Averaging Filter is defined by

$$\phi(x) = \begin{cases} 1 & 0 \leq x < \frac{1}{2} \\ 0 & otherwise. \end{cases}$$

2.2.2 Transform Representation

This derivation uses the wavelet pair to define the wavelet transform. Any wavelet transform uses a satisfactory wavelet pair to transform an array into two halves which constitute the average filter sub-array and the difference filtered sub-array. This definition commonly refers to the average filtered sub-array as the average terms and difference terms respectively. The concatenated array of average and difference terms constitutes a similar array to the original array, and the same properties as similar vectors. Also, for every wavelet transform there is a straight forward method to restore the original array from the wavelet transformed array called the inverse wavelet transform.

One form of the wavelet transform is the integral wavelet transform (IWT) described by Chui. Another two are the discrete wavelet transform (DWT) and the fast wavelet transform (FWT). The convolution wavelet transform (CWT) is a general form of the FWT. In the case of the CWT, any proper wavelet pairs can used to generate corresponding average and difference terms. Application of the CWT is described in section 2.5.

Convolution of the wavelet pairs with an array form the starting point to the convolution wavelet transform. The following constitute the steps of the CWT:

1. Convolute the original array S with the wavelet pair. The results of this are the average filtered array, A, and the difference filtered array, D.

2. Selectively filter every other element of A and D into new arrays half the size of S. The results map $A \to A'$ and $D \to D'$ respectively, where A' and D' are the selectively filtered average filter array, and selectively filtered difference filtered array respectively.

3. Concatenate A' and D' to form $W(S) = (A'|D')$, which is the wavelet transform of S.

2.3 Multi-resolution Representation

In one dimension, multi-resolution analysis provides means to measure averages, differences of the original array or signal and of the sub-arrays generated in application of multi-resolution. In this section, techniques of multi-resolution are defined in terms of one-dimensional wavelet transforms. Two dimensional versions covered in section 2.6. Three methods for decomposing the signal are considered — multiresolution analysis (MRA), multiresolution expansion (MRE) and the ψ^n expansion.

2.3.1 MRA

MRA schemes generates averaging estimates of some signal such that the average representation is smaller than the original by some integer amount. MRA reapplies this averaging process until the process produces a small enough size for the analysis being conducted. MRA may maintain estimate correction factors for the purpose of translating an averaged array to the next larger average array to recover that particular function. The elements of average functions are averaging terms. Likewise, the elements of the estimate correction factor function are called differencing terms. To acquire each estimate by wavelets, the wavelet averaging basis is used. The wavelet basis function happens to be the basis function for the differencing function

There are simple ways to visualize this concept. A classic one dimensional form method is to consider a triangle. In this triangle, layers of functions are stacked up from the base to the apex. The original function is always placed on the bottom. Subsequent average terms are placed between the base and the apex. For every level between the apex and the base, there is an averaging method to map that level to the adjacent level closer to the apex. Also, there exist a set of difference elements such that when the average is expanded it can be mapped to the adjacent level toward the base. The apex has no adjacent level to average to. The base has no difference elements to map to a larger level.

The way to consider MRA with one-dimensional wavelet transforms is with a binary tree. At the root of the tree the entire original array is represented. Two branches exist on this tree, the average and difference branch. Each node in the tree represents a sub-array of the original array. Branches are generated on a node if and only if a wavelet transform is performed on the node. In case a wavelet transform is performed, both an average and difference branch are generated. Otherwise the node is a leaf. The following are rules for the one-dimensional wavelet transform binary tree in MRA:

- Transforms are only applied to leaves.
- Transforms are only allowed on the root or average leaves.
- If a node is not a leaf then a wavelet transformation has already been performed and is not permitted to be reapplied.
- If the leaf is a difference term, then a transformation is not allowed.
- The maximum height of this tree is $n = \lceil \log_2 |S| \rceil$ where $|S|$ is the cardinality of the signal S.

2.3.2 MRE

MRE extends this idea by considering the difference leaves, also. The rules for one-dimensional wavelet transform binary tree in MRA are as follows:

- Transforms are only applied to leaves.
- If a node is not a leaf then a wavelet transformation has already been performed and is not permitted to be reapplied.
- The maximum height of this tree is $n = \lceil \log_2 |S| \rceil$ where $|S|$ is the cardinality of the signal S.

Like in MRA, there is an analogues to time and frequency signal representation. Every sub-array represented at each node with in the wavelet transform binary tree represents activity in the time domain within a certain octave and note of frequencies. That octave's frequency range is specified by its position within the array defined here:

$$\left[\frac{\pi x_1}{2 \cdot |S|}, \frac{\pi (x_2 + 1)}{2 \cdot |S|}\right]$$

where x_1 is the starting index of the sub-array and x_2 is the ending index of the sub-array.

2.3.3 ψ^n Expansion

There is another expansion similar to MRE which places its sections in a different arrangement. This form is called the ψ^n expansion. It is a rather simple expansion of the wavelet transform. Each resolution, reapplies the wavelet transform to the whole array again. While it is simple to conceive and implement, the results are little more complicated to analyze. A mapping procedure does exist for mapping the ψ^n expansion to the MRE and vice versa.

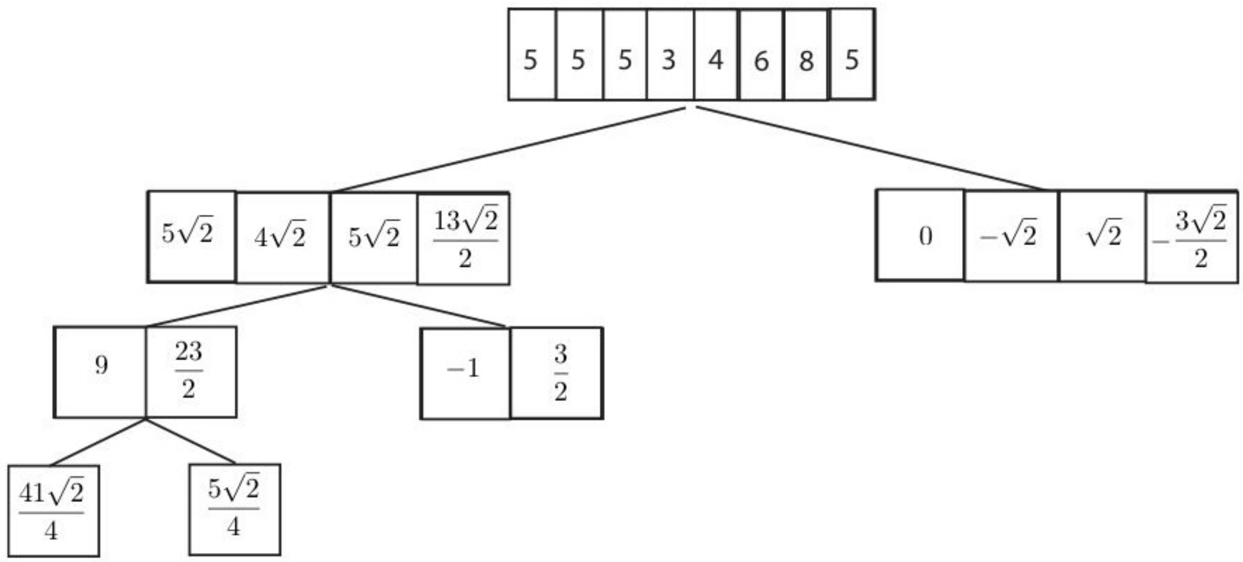

Figure 2.1: The Haar Transform performed a sample function show each step of the transform in multi-resolution analysis

2.4 Numerical Example

To illustrate the concepts of the wavelet transform, two simple numerical examples are provided. The first one utilities MRA and is shown in Figure 2.1. The second one utilizes MRE and is demonstrated in Figure 2.2.

2.4.1 MRA Example

MRA starts with the original array. The wavelet transform is applied to the array. After the first application of the wavelet transform, an average and difference array now exist. This is shown in figure 2.1 as the first two children in the wavelet binary tree.

For MRA, the analysis is continued on the average branch. The difference branch is unchanged. This procedure is repeated on the average branch. The stopping point is determined by the size of array. The finite limit, n, is

$$n = \lceil \log_2 |S| \rceil \tag{2.2}$$

where $|S|$ is the size of the array.

Once completed, the arrays are joined into an array which is the same size as the original. The energy is generally concentrated at the beginning of the array. Also, the energy of the array tends to be ordered from strongest to weakest. Terms representing change are kept as pairs to each section.

Notice that the fine difference terms are left alone, and the dynamics of the function in the time domain are accentuated. Likewise the lower frequency components are separated allowing a closer analysis in both position and frequency.

2.4.2 MRE Example

Like MRA, MRE starts with an original array, and applies a wavelet transform to that array. The decomposition is also represented by binary tree, with an average and difference branch. The limit and height of the tree, n, remains the same as equation 2.2.

What is not the same is how the wavelet transform is reapplied. In addition to being applied on the average branch again, but the wavelet transform is also applied to the difference branch also. An example is provided in Figure 2.2. The sub arrays are reinserted into the array in order from the left branch to the right branch.

In the time and frequency analogues, each transformation filters the array with high pass and low pass filter. The frequency width of these sub arrays is proportional to length of the original array. The center frequency is relative to the position of the sub-array within the array. The lowest

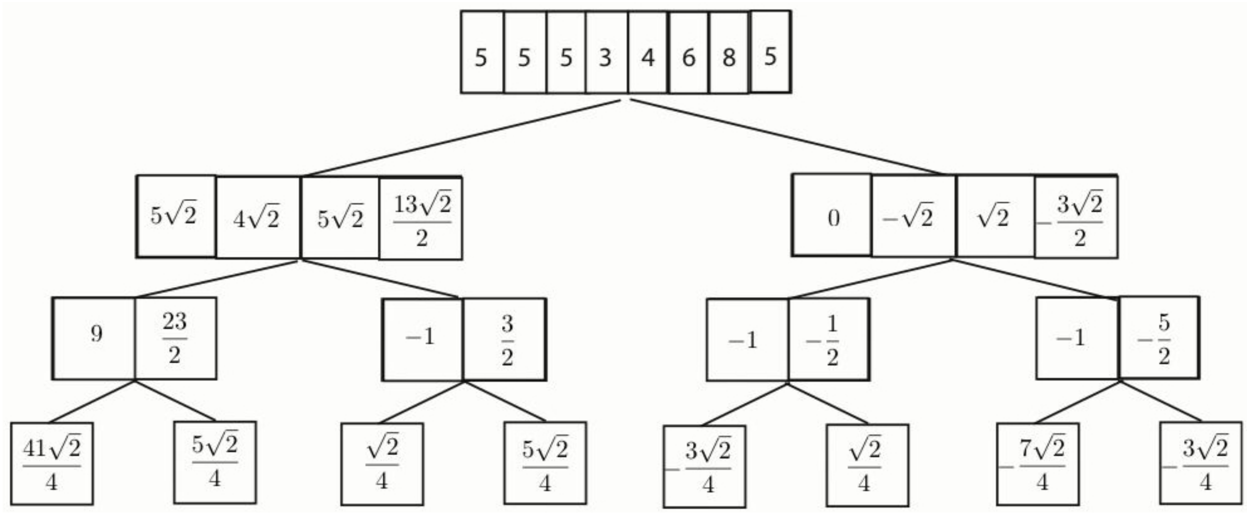

Figure 2.2: The Haar Transform performed a sample function show each step of the transform in multi-resolution expansion

frequency components are on the average side of the array, and the highest frequency components are on the difference section of the array. One special case exists for arrays of length 2^n. If the array is of length 2^n then the full MRE yields entire frequency domain. In case that the array is of odd size larger than one, then the array must be padded and then normal decomposition can continue.

2.4.3 ψ^n Expansion

One other form of the multi-resolution expansion that can be used is almost trivial by its nature, ψ^n expansion. This particular form has the same branches as the MRE. However, the sub arrays are place back in the array in a different order. Figure illustrates the difference in ordering from the ψ^n form and the MRE form. This form makes for analysis such time and frequency difficult, but contributes advantages linear operations.

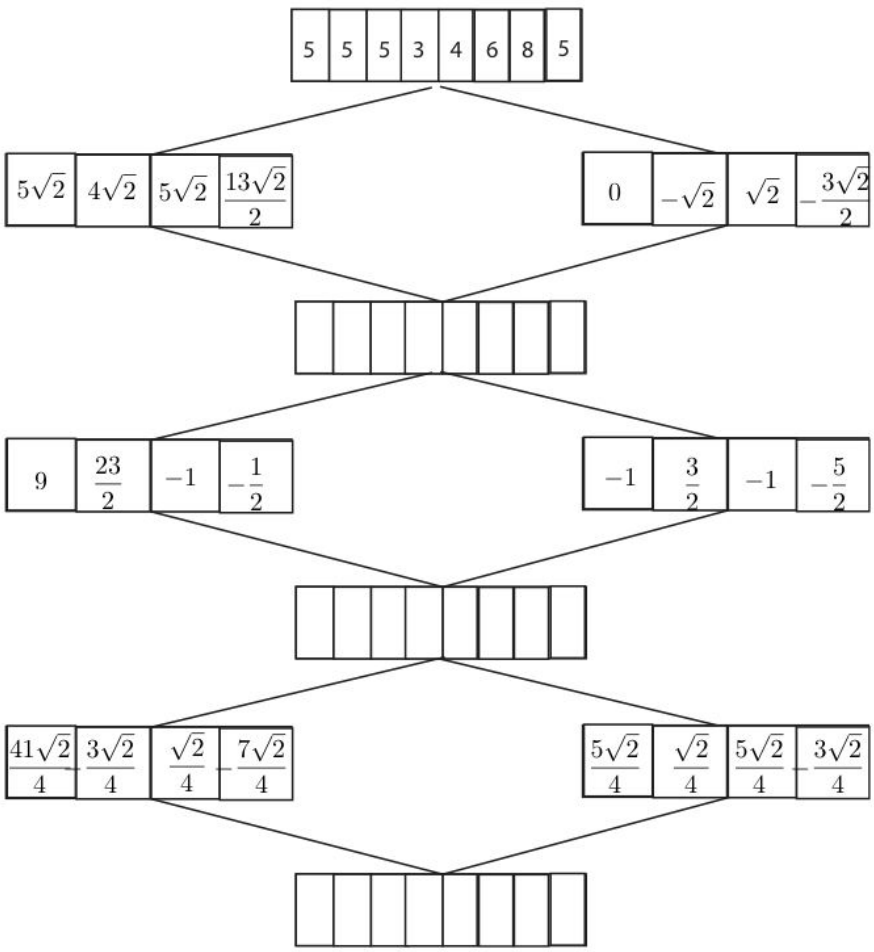

Figure 2.3: The Haar Transform performed a sample function show each step of the transform in ψ^n expansion

2.5 Implementation of the Wavelet Transform

As stated before, the wavelet transform is defined in terms of average and difference components. A signal is taken and transformed into the two base objects. Typically, the transform has the form $S \rightarrow (A'|D')$ where S is the original signal, A is the average component, D is the difference component and $(A'|D')$ the signal A' concatenated with D'. This transformation can be modeled with the convolution operator. Despite the decomposition of the signal into the two components, the original signal can be reconstructed.

Many mathematicians such as Walker[3], use a form that eliminates half of the values. Thus a form can be defined which has the same number of elements as the original. The rules for choosing the member elements are dependent on the wavelet filter choice. Another useful property of wavelets via convolution is the simplicity of the operation. The general case works for all. Such an algorithm requires one nested loop as seen in Algorithm 1.

Algorithm 1 Convolution of two signals $x(\cdot)$ and $h(\cdot)$.

```
for i = 0 to M − 1 do
  for j = 0 to N − 1 do
    n = i − j
    if (0 ≤ n < M) then
      y(i) += x(j) · h(n)
    end if
  end for
end for
```

This filter simply equates to the mathematical function

$$y(k) = (x * h)(k) = \sum_{l} x(l)h(k-l)$$

which is the convolution operation. It is obvious that the operation is $O(NM)$. For practical use, the filter is made smaller than the actual signal being analyzed. In some cases, the filter may be much smalled than the signal. Filter size matters in extracting features from the original signal.

To perform a wavelet transform via convolution, the signal is first convolved with the average operator and then with the difference operator.This can be represented by

$$A = S * \phi \qquad \text{and} \qquad D = S * \psi.$$

where ϕ is the mother wavelet and ψ is the difference wavelet.

2.6 The 2D Wavelet Transform

For the 2D Wavelet Transform, it has to be determined how to represent average and difference components. This can be done by creating one average components and three difference component - vertical, horizontal and diagonal. This can be used in a multiresolution fashion to provide several levels of decomposition.

A complete transform method returns a result matrix which is the same size as the source matrix. The result contains the four components. Each component resides on 4 corners of the matrix. Given a matrix B, the transform is to yield the following form:

$$B \Rightarrow \begin{pmatrix} H & D \\ A & V \end{pmatrix}$$

where A is the average component, H is the horizontal component, V is the vertical component, and D is the difference component. There is another form which is also used as an example:

$$B \Rightarrow \begin{pmatrix} A & V \\ \\ H & D \end{pmatrix}$$

The first version is simple in concept, but provides a few more possibilities for error and confusion. Regardless of the case, the four components have the following definitions:

1. Average component: produced by filtering the row vectors and the column vectors with the averaging filter.

2. Vertical Component: produced by applying the average filter to the column vectors and the difference filter to the row vectors.

3. Horizontal component: produced by applying the average filter to the row vectors and the difference to the column vectors.

4. Diagonal component: produced by applying the difference filter to both the row and column vectors.

2.6.1 Multi-Resolution in Two Dimensions

Three methods are presented for decomposing 2D matrices — 2D MRE, 2D MRA and 2D ψ^n Expansion. The first of these methods for matrices is the 2D MRE. In this form, the wavelet transform is applied recursively to each of the sub-matrices. The inverse of this MRE applies the inverse wavelet transform is applied to the lowest level adjacent sub-matrices.

The 2-D MRA is a special case of the 2-D MRE. The application of the wavelet transform is only applied to the average sub-matrix, only. This method is also called wavelet pyramids.

The ψ^n is rather more simple to implement, but is harder to analyze. In this case, the wavelet transform is simply reapplied to the whole matrix. The result of this method contains the same elements as the 2-D MRE (wavelet transform full decomposition.) However, the arrangement of the elements is not the same. This arrangement will be used for matrix multiplication.

The method for visualizing 2D MRE and 2D MRA is similar to the one-dimensional version. Here a quadtree is used instead. The four branches represent the four sub-matrix types (average, vertical, horizontal, and diagonal.) The rules for this tree is similar for the 1D wavelet transform binary tree.

The following are rules for the 2D wavelet transform quad tree in MRA.

- Transforms are only applied to leaves.
- Transforms are only allowed on the root or average leaves.
- If a node is not a leaf then a wavelet transformation has already been performed and is not permitted to be reapplied.
- If the leaf is a vertical, diagonal or horizontal term, then a transformation is not allowed.
- The maximum height of this tree is $n = \min\left(\lceil N_r \rceil, \lceil N_c \rceil\right)$ where N_r is the number of rows in the matrix and N_c is the number of columns in the matrix.

MRE extends this idea by considering the difference leaves, also. The rules for the 2D wavelet transform quad tree in MRA are below.

- Transforms are only applied to leaves.
- If a node is not a leaf then a wavelet transformation has already been performed and is not permitted to be reapplied.
- The maximum height of this tree is $n = \min\left(\lceil N_r \rceil, \lceil N_c \rceil\right)$ where N_r is the number of rows in the matrix and N_c is the number of columns in the matrix.

2.6.2 Methods of Implementation

Two methods of convolving a matrix are easily conceived. First is to use 1D wavelet. The other is to apply the convolution scheme straight to the matrix. Included in the wavelet experiment are both. Realistically, both can and do achieve the same result. However, the direct method achieves speed advantages by the lack of overhead. The direct method only stores a temporary vector resident in memory. Also, there are two fewer transfers per row and column.

1D to 2D Method

Both rows 1D and 2D and columns 1D and 2D transform are performed similarly. The obvious difference is the indexing of rows and columns.

Algorithm 2 Wavelet Row Transform

Require: Wavelet Transform, , and Wavelet Pair (ψ and ϕ)
Require: Matrix, $S \in \mathbb{R}^{N \times M}$
  for $i = 0$ to $N - 1$ **do**
    $R = S.getRow(i)$
    $S.row(i) = [R * \phi, R * \psi]$
  end for

This principle of this algorithm is simple. Only three intuitive steps are necessary per row or column. Two of these steps are array transfers (row/column transfer to an array). These arrays are fed into the 1D transforms.

However, the 1D wavelet transform itself includes a series of memory allocation and deallocation operations. Each memory call is at the minimum a system call.

Vector - Matrix Method

The principle of this algorithm is more complicated. All functionality, such as convolution, is built into the method. There are fewer calls and passing of structures to external functions to compute the transform.

This method has a few givens. The source matrix, the Haar average filter, and the Haar difference filter are given arguments. The result argument is the return argument. The transform signals sub-function row transforms and column transforms to perform the work.

The algorithm is as follows for the row transform (and is similar for the column transform):

Algorithm 3 Wavelet Transform: Vector - Matrix Method: Row Transform

Require: Wavelet Pair
Require: Temporary Vector S
Require: Matrix, $A \in \mathbb{R}^{M \times N}$
  for $i = 0$ to M **do**
    load S from A_i where A_i is the row vector at row i
    $S \xrightarrow{\psi_1} R$
    Load R into result matrix α at α_i
  end for
  Return α

To perform a wavelet transform with this method, the driving wavelet transform method calls on of the wavelet transform methods and feeds the results into the other. In the implementation used for this thesis, the row transform is called first, then its results α are fed into the column transform

Algorithm 4 Wavelet Transform: Vector - Matrix Method: Column Transform

Require: Wavelet Pair
Require: Temporary Vector S
Require: Matrix, $A \in \mathbb{R}^{M \times N}$
  for $j = 0$ to N **do**
    load S from A_j where A_i is the column vector at column j
    $S \xrightarrow{\psi_1} R$
    Load R into result matrix α at α_j
  end for
  Return α

as the column transforms A. The results from the column transform are returned as the wavelet transform of the original matrix.

2.6.3 2-D MRA: Wavelet Pyramids

This version of the wavelet transform (multiresolution) uses private members of the class $(hA, hD, xD/yD, xA/yA)$. Both Haar filters are maintained this way. Also both row and column transforms have average and difference myVector classes for temporary storage. All of these members are allocated and destroyed by the wavelet transform method itself. The simplified algorithm of the row transform is shown in Algorithm 5, and the column transform is shown in Algorithm 6 .

Note: A_i names the row vectors and A_j names the column vectors, and $A_{i,j}$ is the element from the i^{th} row and j^{th} column.

The computational cost of this algorithm $\frac{1}{2n} \cdot C_\psi$ where C_ψ is the cost of the wavelet transform at large, and n is the number of resolutions performed. The reason for this constant is that the size of the matrix being transformed shrinks by a factor of 2 each time. There may be additional overhead for the setup of the matrices, and transfer of matrix S to T, which is N^2. Since $C_\psi \propto O(N^2)$ then this a cost of another constant multiplied by $O(N^2)$.

Algorithm 5 Wavelet Transform: Wavelet Pyramid Method: Row Transform

Require: Wavelet Pair hA and hD of length w_l
Require: Temporary Vector S
Require: Matrix, $A \in \mathbb{R}^{M \times N}$
Require: Limits of Rows and Columns to traverse: M' and N'
Require: Temporary vectors xA and xD for row Average vector and row Difference Vector.

```
Initialize vector xA and xD
for i = 0 to M' do
  Initialize xA and xD
  for k = 0 to N' do
    for l = 0 to w_l do
      n = k − l
      if n ∈ [0, N'] then
        xA_k = A_{i,n} · hA_l
        xD_k = A_{i,n} · hD_l
      end if
    end for
  end for
  Transfer to α. α_i ← xA'|xD'
end for
Return α
```

Algorithm 6 Wavelet Transform: Wavelet Pyramid Method: Column Transform

Require: Wavelet Pair hA and hD of length w_l
Require: Temporary Vector S
Require: Matrix, $A \in \mathbb{R}^{M \times N}$
Require: Limits of Rows and Columns to traverse: M' and N'
Require: Temporary vectors yA and yD for column Average vector and column Difference Vector.

```
for j = 0 to N' do
  Initialize yA and yD
  for k = 0 to M' do
    for l = 0 to w_l do
      n = k − l
      if n ∈ [0, M'] then
        yA_k = A_{n,j} · hA_l
        yD_k = A_{n,j} · hD_l
      end if
    end for
  end for
  Transfer to α. α_j ← yA'|yD'
end for
Return α
```

Algorithm 7 Wavelet Transform: Wavelet Pyramid Method: Driving Transform

Require: Wavelet Pair hA and hD of length w_l
Require: Matrix A of size $M \times N$
Require: Number of resolutions, r
  Initialize matrix α to $M \times N$ and set equal to A
  Initialize matrix β to $M \times N$ and set equal to zero.
  for $k = 0$ to r **do**
    $M' = \frac{M}{2^k}$
    $N' = \frac{N}{2^k}$
    call row transform for matrix α, with dimension limits M' , N' store in β
    call column transform for matrix β, with dimension limits M' , N' store in α
  end for
  Return α as result

2.6.4 2-D MRE via Quad Tree

There are two ways to implement the 2-D MRE and it depends on if other stopping conditions are desired. The one stopping condition of size can be handled in a simple loop. To enforce an energy limit stop, then a means to stop any further transforms on that branch must be used. Two such means exist. One way is to mark that branch and do so to its children. Another way is to use a queue structure to designate nodes to be transform. If the node represents a stopping condition, simply do not enter that node into the queue.

In either case, the matrix information can be stored in this quad-tree. Either by referencing position within the working matrix, or by storing a matrix in the node itself. In this thesis, references were used for space efficiency. However, the storage of the matrices in the nodes is just as valid. Also a hybrid of only storing the final matrices in the leaves is acceptable as well.

As stated, a straight forward loop can acquire the full decomposition. The limit of the times required for is defined $v_l = 4^{r-1}$ where r is the height of the quad-tree. In this implementation, the quad tree is implemented in an array, and those properties are exploited in using a simple for loop to perform the 2-D MRE.

Algorithm 8 Wavelet Transform: MRE Quad Tree Loop

Require: Wavelet Transform, ψ, and Wavelet Pair
Require: Matrix, $S \in \mathbb{R}^{N \times M}$
  for $i = 0$ to v_l **do**
    draw S from quad-tree at index i
    $S \overset{\psi_r}{\to} T$
    $T \overset{\psi_c}{\to} X$
    Wavelet Split and Store X in branch children.
  end for

Note on Algorithm 8 that the Wavelet Split and Store operation splits the matrix X into its four sub-matrices (average, horizontal, vertical, and diagonal) and stores them in their appropriate spot in the tree. Thus every-time that S is loaded a leaf branch of the tree is loaded to be transformed.

The queue based 2-D MRE with a quad-tree is a bit more complicated. The queue must be loaded with the root, then operation may be allowed to proceed. The end condition is when the queue is empty. Rules for loading the queue is must force the leaves on the last level to be ineligible for loading. Otherwise, the loop will never stop and segmentation fault is likely to arise. Energy rules and other arbitrary rules may also be imposed, but they are not as critical as the leaf limit rule. For this example, refer to Algorithm 9. The computational cost of this algorithm $n \cdot C_\psi$ where

Algorithm 9 Wavelet Transform: MRE with queue controlled visits of the Quad Tree

Require: Wavelet Transform, MRE, and wavelet pair (ψ and ϕ)
Require: Matrix S
  Insert root node representing the whole matrix in the queue.
  while as long as queue is not empty **do**
    Get the next element from the queue, and load matrix S
    $S \overset{\psi_r}{\to} T$
    $T \overset{\psi_c}{\to} X$
    Check X for terminating conditions
    if X does not have a terminating condition store X's four sections into the branch children of node S. Store theses children nodes in the queue in order of average, vertical, horizontal, and difference.
  end while

C_ψ is the cost of the wavelet transform at large, and n is the number of resolutions performed. In some special cases, the cost may be slightly to significantly less if additional termination rules are applied. The reduced size of the sub-matrices is canceled by quantity of sub-matrices to transform. There may be additional overhead for the setup of the matrices. Since $C_\psi \propto O(N^2)$ then this a cost of this algorithm is on the order N^2 also.

2.6.5 2-D ψ^n Implementation

As stated the ψ^n expansion is simpler to implement, but produces a matrix that is harder to analyze for features. However, it is scientific merit since it is literally a wavelet transform of a wavelet transform. Also, the limits imposed on the MRA and MRE for tree height are not required, but are rather a good idea. A good practical limit would be for the resolution limit to be defined $n = \log_2 |S| - 1$ where $|S|$ is the size of the the matrix S. For this example, refer to Algorithm 10.

Algorithm 10 Wavelet Transform: MRE with queue controlled visits of the Quad Tree

Require: Wavelet Transform, MRE, and wavelet pair (ψ and ϕ)
Require: Matrix S
  Load S into temporary matrix T
  for $i = 1$ to n **do**
    $T \xrightarrow{\psi_c} X$
    $X \xrightarrow{\psi_r} T$
  end for
  return T as the transformed matrix

The computational cost of this algorithm $n \cdot C_\psi$ where C_ψ is the cost of the wavelet transform at large, and n is the number of resolutions performed. There may be additional overhead for the setup of the matrices, and transfer of matrix S to T, which is N^2. Since $C_\psi \propto O(N^2)$ then this a cost of another constant multiplied by $O(N^2)$.

Chapter 3

Image Processing Example of Wavelets

This chapter is for showing graphically examples of the one-dimensional and two dimensional wavelet transform. For one dimensional wavelet transform a simple sinusoid signal is used as the source input to measure correctness of the implementation. Likewise, a simple pictorial image is used as the test input to show correctness of these implementations of the two dimensional wavelet transform.

3.1 Results — 1D Wavelet Transform

Testing of the 1D wavelet was performed on a sinusoidal wave form of 128 elements:

$$y(n) = 10\sin\left(\frac{n}{128}\right) - 5\sin\left(\frac{n}{64}\right) + 2\sin\left(\frac{3n}{128}\right) - \sin\left(\frac{n}{32}\right). \tag{3.1}$$

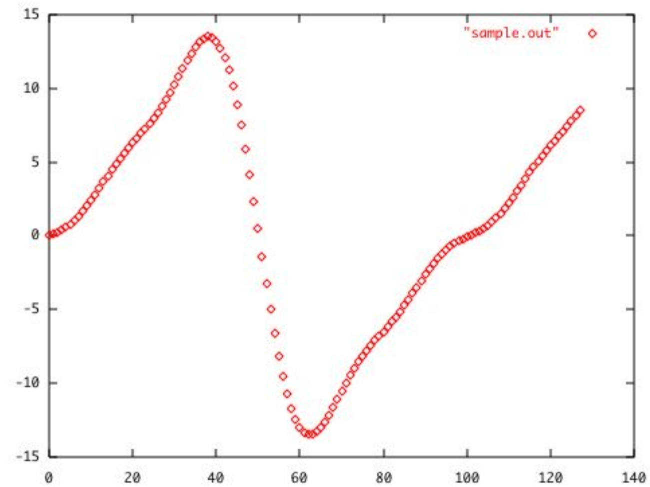

Figure 3.1: Sample function. The x-axis is the array index (index n). The y value is simple – the value $y(n)$.

This function is the input function used for the one-dimensional implementations and is shown graphically in Figure 3.1. The first test of the 1D transform used the even elements of both convolutions to generate the wavelet transform. These even elements came from the over-complete form and naturally allow the potential to have complete information. However, in doing so, a fundamental flaw appears.

In order to evaluate the effectiveness of the wavelet transform three tests have been devised. First, energy equivalence is used to determine how much energy is retained in the transform from the original. The general shape is used on a the first resolution to test if the average signal has the same general shape as the original. Lastly, the inverse transform is used to recover the original signal. A comparison is made between the original and the recovered signal.

After one resolution, the transformed signal has the same energy as the original. This is good since it allows the original to be recovered from the transform. Also, the average component of

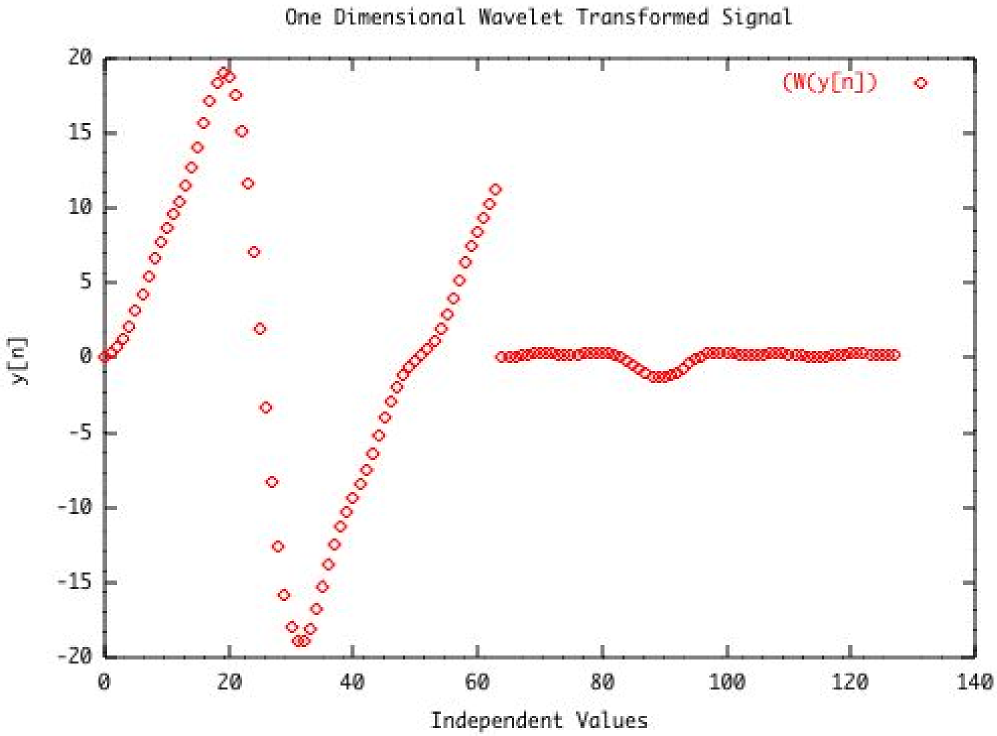

Figure 3.2: Signal after the wavelet transformation.

the transform has the same shape as the original, which is good. However, the recovered signal is missing the last element. Refer to Figure 3.3. The secret is in which elements are used from the over-complete to make the complete. The over-complete form is defined from the average and difference components which are simply the result of convolution.

The convolution means is at the heart of the issue. The convolution operator in this case starts with the first element of the filter against the first element of the signal. In the simple Haar Wavelet case, there is a transformation pairing

$$(S_i, S_{i-1}) \rightarrow A_i, \text{ and } (S_i, S_{i-1}) \rightarrow D_i.$$

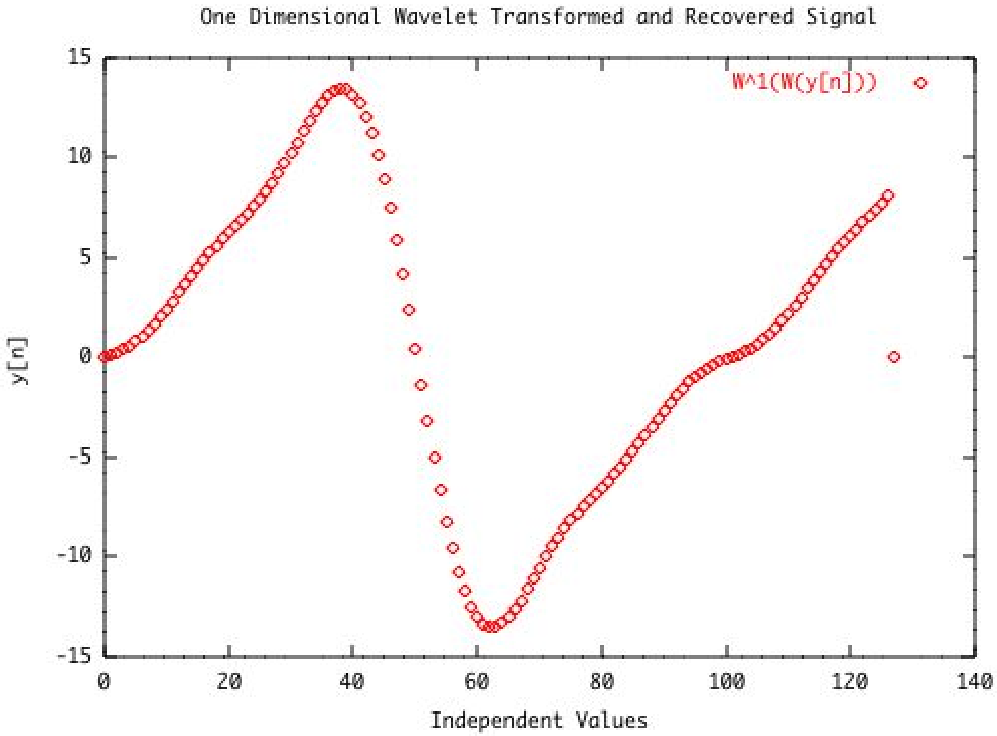

Figure 3.3: Recovered function. The x-axis is the array index (index n). The y value is simply the value $y(n)$. The function was recovered from an even indexed wavelet transform.

In this pairing with zero indexed signals, the odd indexed elements from the over-complete must be used to have all elements of the original accounted for.

Also this produces a functional difference between wavelet inverse transform for odd and even versions. The difference is slight; however, the last element is lost in the even indexed form.

$$\text{Odd:} \qquad R_{2i} = (A_i - D_i)\sqrt{1/2}, \qquad R_{2i+1} = (A_i + D_i)\sqrt{1/2}$$

$$\text{Even:} \qquad R_{2i} = (A_i + D_i)\sqrt{1/2}, \qquad R_{2i-1} = (A_i - D_i)\sqrt{1/2}$$

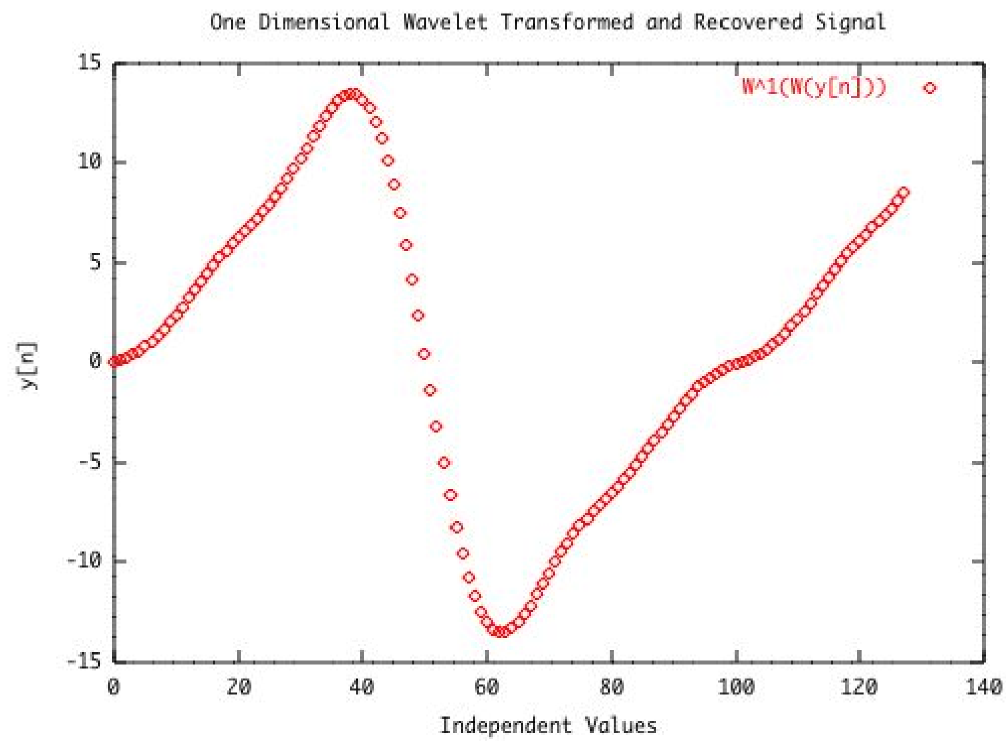

Figure 3.4: Recovered function. The x-axis is the array index (index n). The y value is simply the value y[n]. The function was recovered from an odd indexed wavelet transform.

An odd indexed wavelet transform yields the same energy. However, all of the values are accounted for. Refer to Figure 3.4.

3.2 Results: 2D Wavelet Transform

A simple room picture shows the difference that correct indexing produces in the wavelet transform and its inverse. The 1D to 2D method shows the incorrectly indexed case. A correctly indexed version is shown in the vector-matrix method.

The 1D to 2D implementation has a serious issue with memory leak errors. Memory is allocated and deallocated quickly, and on some platforms shows up as an error. Some platforms are not forgiving of this error and will force the program to terminate (Macintosh OSX 10.2, using gcc 3.1). On other platforms, the error is tolerated and performance is degraded (IRIX, SGI Octane2 using gcc 2.9). An example image of 720 x 486 required nearly 10 minutes to compute the wavelet transform by this method on an SGI Octane2. However, it does eventually return a correct result.

The matrix-vector method also yields the correct result. However, there is less memory overhead in this method as compared to the 1D to 2D method. As a result, both the row wavelet transform and column wavelet transforms are performed more quickly, with fewer memory transfers and allocations. Obviously, this also allows for the operation to be conducted almost entirely in cache memory on both the SGI Octane2 and Macintosh G4 based machines. A Macintosh G3 based machine still requires main memory at a minimum to execute the same operation.

A correct result must also be matched to a correct inverse method. The indexing order matters. The inverse transform method is a forward inverse transform method. In the case of 1D to 2D transform, the ordering was reverse indexed (Figure 3.7). As a result of an error in indexing, ringing is seen on edges in this method(Figure 3.6) for a case in point. Caution is incredibly important when matching both forward and reverse indexing, since matching the mathematics to the actual ordering can be obscure and tricky.

A correct result is shown in Figure 3.9. In this case, the indexing was matched up and ringing is not present. It is clear that the recovered image and the original (Figure 3.5) are nearly indistinguishable.

Figure 3.5: Original Image. This image is the original image.

Figure 3.6: Recovered Image. This image is the recovered image. Depending on whether the image was saved as a picture first can affect the white spots in the picture. Ringing is also an issue.

3.2.1 Multiresolution Results

The expected result is a picture within a picture. Each average component has a further transform on it. The three resolution transform has the form:

$$\begin{pmatrix} A_3 & V_3 & V_2 & V_1 \\ H_3 & D_3 & & \\ H_2 & & D_2 & \\ & & & \\ H_1 & & & D_1 \\ & & & \end{pmatrix}$$

Refer to Figure 3.10 for the image transform results.

To obtain the inverse, an exact reverse procedure is necessary, otherwise the distortion is hideous.

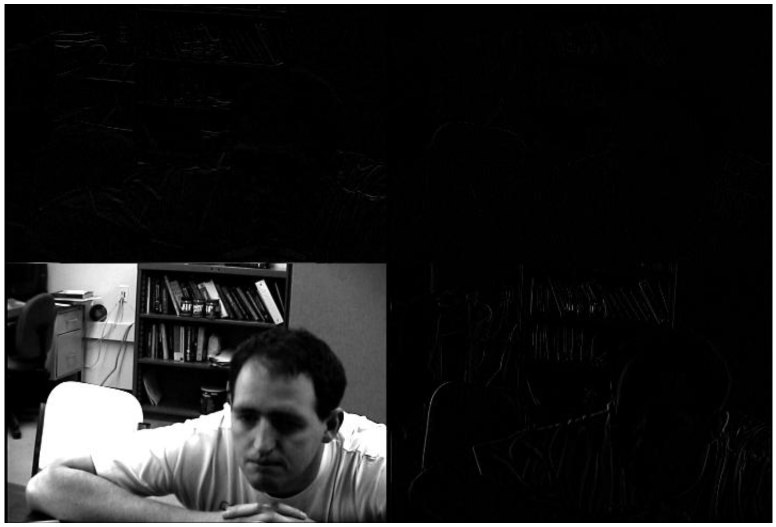

Figure 3.7: Wavelet Transform Image. This image is divided in to average, horizontal, vertical and diagonal components.

The first attempt of the wavelet inverse transform was out of order, refer to Figure 3.11. A correct picture was obtained during the second attempt. Correct order yielded correct results, refer to Figure 3.6.

3.2.2 Threshold Filtering

After a triple resolution, a 0.02 threshold will eliminate 81.1706 percent of elements in the original sample picture. Also at this point, the effects of removing these elements becomes visually evident (Figure 3.12). At a 0.01 threshold, 66.0205 percent of the elements are removed. Visually, the recovered sample and the original appear to be the same (Figure 3.15). At a threshold of 0.1, 92.9987 percent of the elements are reduced to zero. However, the distortions are clearly visible at this level of thresholding (Figure 3.13). Even at a threshold of 0.001 which is below the numerical precision of the original, 16.0814 elements are reduced to zero. At a threshold of 0.002, 28.9683 percent is removed.

Figure 3.8: Wavelet Transform Image. This image is divided in to average, horizontal, vertical and diagonal components, using the vector-matrix version.

Consequently after a triple resolution, nearly 29% of the data was irrelevant for the image's brightness resolution (which also applies to color). Subjective examination reveals that removing 60% to 85% of the data was not noticeable to human perception. Which leaves only 15% to 40% of the data actually contributing or being necessary to reconstruct the image.

Figure 3.9: Recovered Image (Vector-Matrix Method). This image is the recovered image. This version avoids the ringing by using the vector-matrix version which is more aligned for the inverse wavelet transform.

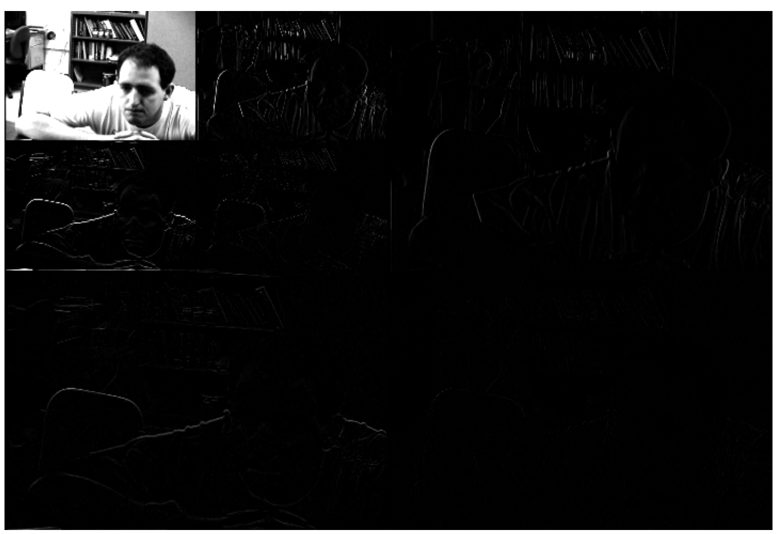

Figure 3.10: Wavelet Transform Image. This image is divided in to average, horizontal, vertical and diagonal components using multiresolution wavelet transform. Note the the average component was transformed one step further.

Figure 3.11: Recovered Image - Wrong Order (Multiresolution). This image shows a 2D wavelet transform after it was recovered out of order. Obviously, the distortion is hideous.

Figure 3.12: Recovered Image - 2% threshold (Multiresolution). This image had nearly 83% of its elements removed in the triple resolution wavelet transform.

Figure 3.13: Recovered Image - 10% threshold (Multiresolution). This image had nearly 93% of its elements removed in the triple resolution wavelet transform.

Figure 3.14: Recovered Image - 5% threshold (Multiresolution). This image had nearly 85% of its elements removed in the triple resolution wavelet transform.

Figure 3.15: Recovered Image - 1% threshold (Multiresolution). This image had nearly 60% of its elements removed in the triple resolution wavelet transform.

Chapter 4

Matrix Multiplication via Wavelets

In this chapter, the general concept of matrix multiplication via wavelets is introduced. In addition, the chapter provides proof that the Haar Wavelet Transform can precondition matrix, and produces the same answer. A little bit of matrix multiplication and sparse matrix multiplication itself is reviewed for completeness.

4.1 Matrix Multiplication

Matrix multiplication is this ...

What is sparse matrix multiplication?

Matrix multiplication is one of the fundamental operations in linear algebra. It is defined for two matrices A and B, denoted $A \cdot B$. Matrix multiply requires that the row length of A to be the same as the column length of B. Lastly, the matrix multiplication operation is defined:

$$c_{i,j} = \sum_k a_{i,k} b_{k,j}.$$

Matrix multiplication also obeys the following properties

- Associative law: $(AB)C = A(BC)$
- Left Distributive Law $A(B + C) = AB + AC$
- Right Distributive Law $(B + C)A = BA + CA$
- Distribution of a constant

[18]

Consider a 2×2 matrix, this example matrix multiplication:

$$A \cdot B = \begin{pmatrix} a_1^1 & a_1^2 \\ a_2^1 & a_2^2 \end{pmatrix} \begin{pmatrix} b_1^1 & b_1^2 \\ b_2^1 & b_2^2 \end{pmatrix} = \begin{pmatrix} a_1^1 b_1^1 + a_1^2 b_2^1 & a_1^1 b_1^2 + a_1^2 b_2^2 \\ a_2^1 b_1^1 + a_2^2 b_2^1 & a_2^1 b_1^2 + a_2^2 b_2^2 \end{pmatrix} \tag{4.1}$$

There is a fast matrix multiply which was devised in 1969 by Strassen. It tends to obtain a computational complexity of $O(N^{2.807})$. This works for square even dimensioned matrices. There is a variation of the Strassen called the Winograd obtains slightly better performance. [21]

Another view on matrix multiplication is sparse multiplication, is the use of sparse matrix multiply. One method is included in the Sparse Matrix Multiplication Package, called SYMBMM [20]. An other is the one contained Basic Linear Algebra Subprogram for sparse matrix to dense matrix multiply.

The whole point of sparse matrix multiplication is to take a matrix such that the matrix composition is mostly zero, and exploit that composition to reduce the number of multiplications required. Such multiplication schemes optimal limit for matrix multiplication is $O(N^2)$.

4.2 Wavelet Matrix Multiplication

The point of this thesis is not show the efficiencies of these above algorithms. Rather, it is to show a pre-conditioner using the wavelet transform that can be used in each of them. Wavelet based matrix multiply is sound for the Haar Wavelet Transform and the ψ^n expansion. In other words, All terms contributed by the wavelet transform on the operation of matrix multiplication are cancelled in the inverse. If the Haar Wavelet Transform produces a sparse and well conditioned matrix, then the Haar Wavelet Transform proves itself as a useful preconditioner. In this chapter, the general concept of matrix multiplication via wavelets is introduced, and the linearity principle is shown. The key point for wavelet matrix multiplication is the proof that $W(A) \times W(B) = W(A \times B)$ If this is the case, then it is obvious that $W(A) \times W(B) = W(\Gamma) = W(A \times B = \Gamma)$.

4.2.1 A 2×2 example

The feasibility of multiplication in the wavelet domain is demonstrated directly using a 2×2 matrix. The coefficients of the matrices are multiplied both according to normal matrix multiplication and the modified wavelet multiplication operator. In the end the resulting coefficients are seen to be the same.

Wavelet Transforms of the Matrices

For a wavelet transform, the result on matrix A is

$$W(A) = \frac{1}{2}\begin{pmatrix} \left(a_1^1 + a_2^1 + a_1^2 + a_2^2\right) & \left(a_1^1 + a_2^1 - a_1^2 - a_2^2\right) \\ \left(a_1^1 - a_2^1 + a_1^2 - a_2^2\right) & \left(a_1^1 - a_2^1 - a_1^2 + a_2^2\right) \end{pmatrix}, \tag{4.2}$$

and for matrix B it is

$$W(B) = \frac{1}{2}\begin{pmatrix} \left(b_1^1 + b_2^1 + b_1^2 + b_2^2\right) & \left(b_1^1 + b_2^1 - b_1^2 - b_2^2\right) \\ \left(b_1^1 - b_2^1 + b_1^2 - b_2^2\right) & \left(b_1^1 - b_2^1 - b_1^2 + b_2^2\right) \end{pmatrix}. \tag{4.3}$$

Product of A and B in wavelet space

The conventional product of A and B (equation 4.1) can be transformed into wavelet space. The wavelet transform of this matrix is represented by

$$W(A \cdot B) = \frac{1}{2}\begin{pmatrix} \psi(A) & \psi(V) \\ \psi(H) & \psi(D) \end{pmatrix}$$

where

$$\psi(A) = (a_1^1 b_1^1 + a_1^2 b_2^1 + a_1^1 b_1^2 + a_1^2 b_2^2) + (a_2^1 b_1^1 + a_2^2 b_2^1 + a_2^1 b_1^2 + a_2^2 b_2^2) \tag{4.4}$$

$$\psi(V) = (a_1^1 b_1^1 + a_1^2 b_2^1 - a_1^1 b_1^2 - a_1^2 b_2^2) + (a_2^1 b_1^1 + a_2^2 b_2^1 - a_2^1 b_1^2 - a_2^2 b_2^2) \tag{4.5}$$

$$\psi(H) = (a_1^1 b_1^1 + a_1^2 b_2^1 + a_1^1 b_1^2 + a_1^2 b_2^2) - (a_2^1 b_1^1 + a_2^2 b_2^1 + a_2^1 b_1^2 + a_2^2 b_2^2) \tag{4.6}$$

$$\psi(D) = (a_1^1 b_1^1 + a_1^2 b_2^1 - a_1^1 b_1^2 - a_1^2 b_2^2) - (a_2^1 b_1^1 + a_2^2 b_2^1 - a_2^1 b_1^2 - a_2^2 b_2^2). \tag{4.7}$$

The product of the waveletized matrices

Straight forward multiplication of $W(A) \cdot W(B)$ represented by equations 4.2 and 4.3 works out as follows:

$$W(A) \cdot W(B) = \frac{1}{4} \begin{pmatrix} W_A & W_V \\ W_H & W_D \end{pmatrix}$$

where

$$\begin{aligned}
W_A &= (a_1^1 + a_2^1 + a_1^2 + a_2^2)(b_1^1 + b_2^1 + b_1^2 + b_2^2) + (a_1^1 + a_2^1 - a_1^2 - a_2^2)(b_1^1 - b_2^1 + b_1^2 - b_2^2) \\
W_V &= (a_1^1 + a_2^1 + a_1^2 + a_2^2)(b_1^1 + b_2^1 - b_1^2 - b_2^2) + (a_1^1 + a_2^1 - a_1^2 - a_2^2)(b_1^1 - b_2^1 - b_1^2 + b_2^2) \\
W_H &= (a_1^1 - a_2^1 + a_1^2 - a_2^2)(b_1^1 + b_2^1 + b_1^2 + b_2^2) + (a_1^1 - a_2^1 - a_1^2 + a_2^2)(b_1^1 - b_2^1 + b_1^2 - b_2^2) \\
W_D &= (a_1^1 - a_2^1 + a_1^2 - a_2^2)(b_1^1 + b_2^1 - b_1^2 - b_2^2) + (a_1^1 - a_2^1 - a_1^2 + a_2^2)(b_1^1 - b_2^1 - b_1^2 + b_2^2)
\end{aligned}$$

which simplifies to

$$\begin{aligned}
W_A &= a_1^1 b_1^1 + a_2^1 b_1^1 + a_1^2 b_2^1 + a_2^2 b_2^1 + a_1^1 b_1^2 + a_2^1 b_1^2 + a_1^2 b_2^2 + a_2^2 b_2^2 \\
W_V &= a_1^1 b_1^1 + a_2^1 b_1^1 + a_1^2 b_2^1 + a_2^2 b_2^1 - a_1^1 b_1^2 - a_2^1 b_1^2 - a_1^2 b_2^2 - a_2^2 b_2^2 \\
W_H &= a_1^1 b_1^1 - a_2^1 b_1^1 + a_1^2 b_2^1 - a_2^2 b_2^1 + a_1^1 b_1^2 - a_2^1 b_1^2 + a_1^2 b_2^2 - a_2^2 b_2^2 \\
W_D &= a_1^1 b_1^1 - a_2^1 b_1^1 + a_1^2 b_2^1 - a_2^2 b_2^1 - a_1^1 b_1^2 + a_2^1 b_1^2 - a_1^2 b_2^2 + a_2^2 b_2^2
\end{aligned}$$

These can then be compared to the coefficients of $W(A \cdot B)$ in equations 4.4-4.5 and seen to be identical. This asserts that $W(A) \cdot W(B) = W(A \cdot B)$ in the case of 2×2 matrices.

4.2.2 Haar Wavelets and Vector Inner Products

One of the other crucial keys for the Haar Wavelet Transform Matrix Multiply to work is vector inner product. The issue is whether or not

$$\langle f', g' \rangle = \langle f, g \rangle \tag{4.8}$$

is true for vectors f and g which both of length p, and the wavelet transformed versions of f and g denoted f' and g' respectively. To show this, $\langle f', g' \rangle$ is expanded algebraically to establish its identity.

$$\langle f', g' \rangle = \sum_{k=0}^{p/2-1} (\frac{f_{2k} + f_{2k+1}}{\sqrt{2}} \cdot \frac{g_{2k} + g_{2k+1}}{\sqrt{2}}) + \sum_{k=0}^{p/2-1} (\frac{f_{2k} - f_{2k+1}}{\sqrt{2}} \cdot \frac{g_{2k} - g_{2k+1}}{\sqrt{2}}) \tag{4.9}$$

When equation 4.9 is expanded further, the terms that emerge expose the identity of the inner product.

$$\begin{aligned} \langle f', g' \rangle = \frac{1}{2} \sum_{k=0}^{p/2-1} (&f_{2k}g_{2k} + f_{2k+1}g_{2k} + f_{2k}g_{2k+1} + f_{2k+1}g_{2k+1} \\ &+ f_{2k}g_{2k} - f_{2k+1}g_{2k} - f_{2k}g_{2k+1} + f_{2k+1}g_{2k+1} \end{aligned} \tag{4.10}$$

When simplified the $f_{2k}g_{2k+1}$ and $f_{2k+1}g_{2k}$ terms cancel. The $f_{2k}g_{2k}$ and $f_{2k+1}g_{2k+1}$ terms combine to yield equation 4.11 and further simplifies to equation vectinwtsimple2.

$$\langle f', g' \rangle = \frac{1}{2} \sum_{k=0}^{p/2-1} (f_{2k}g_{2k} + f_{2k+1}g_{2k+1}) \tag{4.11}$$

$$\langle f', g' \rangle = \frac{1}{2} \sum_{k=0}^{p-1} (f_k g_k) = \langle f, g \rangle \tag{4.12}$$

4.2.3 Matrix Multiply with the Haar Transform: Formal Proof

Given two arbitary matrices A of size $m \times p$, B of size $m \times p$, the Haar Wavelet Pair, and the Haar Wavelet Transform. The wavelet pair to be used here is

$$\psi_H(x) = \sqrt{\frac{1}{2}} \begin{cases} 1 & x = 0 \\ -1 & x = 1 \\ 0 & otherwise \end{cases}$$

$$\phi_H(x) = \sqrt{\frac{1}{2}} \begin{cases} 1 & x = 0 \\ 0 & otherwise \end{cases}$$

Also, this notation is used for this proof.

- $\psi_{1R}(S)$ is the row transform of matrix S.
- $\psi_{1C}(S)$ is the column transform of matrix S.
- ψS is the 2-D wavelet transform of matrix S.
- $\langle f, g \rangle = \langle \psi_1(f), \psi_1(g) \rangle$
- $A' = \psi(A)$
- $B' = \psi(B)$
- $A^R = \psi_{1R}(A)$
- $B^C = \psi_{1C}(B)$
- A^R_{ri} is the row vector i of the row transform of A.

- A^R_{ri+1} is the row vector $i+1$ of the row transform of A.

- B^C_{cj} is the column vector j of the column transform of B

- B^C_{cj+1} is the column vector $j+1$ of the column transform of B

Required show that

$$\psi(A)\psi(B) = \psi(AB)$$

is a true statement.

Proof

There are two formulae required for transforming either A or B into the wavelet domain.

$$\psi(A) = \psi_{1R}(\psi_{1C}(A)) = \psi_{1C}(\psi_{1R}(A)) \tag{4.13}$$

$$\psi(B) = \psi_{1R}(\psi_{1C}(B)) = \psi_{1C}(\psi_{1R}(B)) \tag{4.14}$$

Thus, this is simply a re-write of the definition of the wavelet transform by the CWT. Next, define a matrix Γ' so that

$$\Gamma' = \psi(A)\psi(B)$$

. From, the definition Γ', a series of re-writes and the properties of the Haar Wavelet Transform can expose the solution for every element of Γ' and Γ'. In general, the elements of Γ' defined as follows:

$$\Gamma'_{i,j} = \langle \psi_{1C}(\psi_{1R}(A))_{ri}, \psi_{1R}(\psi_{1C}(B))_{ci} \rangle \tag{4.15}$$

$$\Gamma'_{i,j} = \langle \psi_{1R}(\psi_{1C}(A))_{ri}, \psi_{1C}(\psi_{1R}(B))_{ci} \rangle \tag{4.16}$$

In this case, each element of Γ' is defined as nothing more than the inner product of row i of $\psi(A)$ and column j of $\psi(B)$. This is in agreement with standard matrix multiplication.

Another identity that is important what the wavelet pair for transform will do in either a row or column transform.

$$\psi_{1C}(A) = \begin{cases} \frac{A_{i,j}+A_{i,j+1}}{\sqrt{2}} & j < col \\ \frac{A_{i,j}-A_{i,j+1}}{\sqrt{2}} & j \geq col \end{cases}$$

$$\psi_{1R}(A) = \begin{cases} \frac{A_{i,j}+A_{i+1,j}}{\sqrt{2}} & i < row \\ \frac{A_{i,j}-A_{i+1,j}}{\sqrt{2}} & i \geq row \end{cases}$$

In this case, there are four cases to show for:

1. $i < \frac{row}{2}$ and $j < \frac{col}{2}$
2. $i \geq \frac{row}{2}$ and $j < \frac{col}{2}$
3. $i < \frac{row}{2}$ and $j \geq \frac{col}{2}$
4. $i \geq \frac{row}{2}$ and $j \geq \frac{col}{2}$

Case 1 $i < \frac{row}{2}$ and $j < \frac{col}{2}$. The base equation $\Gamma'_{i,j}$ in this cases is as follows:

$$\Gamma'_{ij} = \left\langle \frac{A^R_{ri} + A^R_{ri+1}}{\sqrt{2}}, \frac{B^C_{cj} + B^C_{cj+1}}{\sqrt{2}} \right\rangle \tag{4.17}$$

This re-write must be expanded to expose an equality. This expansion is valid for vectors sums within inner products. Thus $\Gamma'_{i,j}$ in this case is expanded to:

$$\Gamma'_{i,j} = \frac{1}{2}(\langle A^R_{ri}, B^C_{cj}\rangle + \langle A^R_{ri+1}, B^C_{cj}\rangle + \langle A^R_{ri}, B^C_{cj+1}\rangle + \langle A^R_{ri+1}, B^C_{cj+1}\rangle)$$

Next, $\Gamma'_{i,j}$ must be compared to $\psi(\Gamma)_{i,j}$ where

$$\psi(AB) = \psi(\Gamma)$$

By definition, every element in C is defined:

$$C_{i,j} = \langle A_{ri}, B_{cj}\rangle$$

The next step is to show what every element within this case is for $\psi(\Gamma)$. This is done by analyzing what the column transform will do at $C_{i,j}$. To do this, the column transform is applied to both column j and $j+1$ for columns between $[0, \frac{col}{2} - 1]$. The effects of these column transforms are defined in equations 4.18 and 4.19.

$$\psi_{1C}(\Gamma)_{i,j} = \frac{1}{\sqrt{2}} \langle A_{ri}, B_{cj}\rangle + \langle A_{ri+1}, B_{cj}\rangle \tag{4.18}$$

$$\psi_{1C}(\Gamma)_{i,j+1} = \frac{1}{\sqrt{2}} \langle A_{ri}, B_{cj+1}\rangle + \langle A_{ri+1}, B_{cj+1}\rangle \tag{4.19}$$

The result of the row transform on $\psi_{1C}(\Gamma)$ in this case is a sum of $\psi_{1C}(\Gamma)_{i,j}$ and $\psi_{1C}(\Gamma)_{i,j+1}$, and specified in equation 4.20.

$$\psi(\Gamma) = \frac{1}{\sqrt{2}}(\psi_{1C}(\Gamma)_{i,j} + \psi_{1C}(\Gamma)_{i,j+1}) \tag{4.20}$$

Expanded this equation is equation 4.21.

$$\psi(\Gamma) = \frac{1}{2}(\langle A_{ri}, B_{cj}\rangle + \langle A_{ri+1}, B_{cj}\rangle + \langle A_{ri}, B_{cj+1}\rangle + \langle A_{ri+1}, B_{cj+1}\rangle) \tag{4.21}$$

Case 2 $i \geq fracrow2$ and $j < \frac{col}{2}$. The base equation $\Gamma'_{i,j}$ in this cases is as follows:

$$\Gamma'_{ij} = \left\langle \frac{A^R_{ri,j} - A^R_{ri+1}}{\sqrt{2}}, \frac{B^C_{cj} + B^C_{cj+1}}{\sqrt{2}} \right\rangle$$

The next step is to show what every element within this case is for $\psi(\Gamma)$.

$$\Gamma'_{i,j} = \frac{1}{2}(\langle A^R_{ri}, B^C_{cj}\rangle - \langle A^R_{ri+1}, B^C_{cj}\rangle + \langle A^R_{ri}, B^C_{cj+1}\rangle - \langle A^R_{ri+1}, B^C_{cj+1}\rangle)$$

How does this compare with $\psi(\Gamma)$?

$$\psi(AB) = \psi(\Gamma)$$

$$(\Gamma) = \psi \langle A_{ri}, B_{cj}\rangle$$

This equation expands with wavelet operator to:

$$\psi_{1C}(\Gamma)_{i,j} = \frac{1}{\sqrt{2}} \langle A_{ri}, B_{cj}\rangle - \langle A_{ri+1}, B_{cj}\rangle$$

$$\psi_{1C}(\Gamma)_{i,j+1} = \frac{1}{\sqrt{2}} \langle A_{ri}, B_{cj+1}\rangle - \langle A_{ri+1}, B_{cj+1}\rangle$$

The result of the row transform on $\psi_{1C}(\Gamma)$ in this case is a sum of $\psi_{1C}(\Gamma)_{i,j}$ and

$$\psi(\Gamma) = \frac{1}{\sqrt{2}}(\psi_{1C}(\Gamma)_{i,j} + \psi_{1C}(\Gamma)_{i,j+1})$$

Expanded this equation is equation 4.21.

$$\psi(\Gamma) = \frac{1}{2}(\langle A_{ri}, B_{cj}\rangle - \langle A_{ri+1}, B_{cj}\rangle + \langle A_{ri}, B_{cj+1}\rangle - \langle A_{ri+1}, B_{cj+1}\rangle)$$

Case 3 $i < \frac{row}{2}$ and $j \geq \frac{col}{2}$. Like, before the base case for $\Gamma'_{i,j}$

$$\Gamma'_{ij} = \left\langle \frac{A^R_{ri,j} + A^R_{ri+1}}{\sqrt{2}}, \frac{B^C_{cj} - B^C_{cj+1}}{\sqrt{2}} \right\rangle$$

Again, the vector sums are expanded within their inner products.

$$\Gamma'_{i,j} = \frac{1}{2}(\langle A^R_{ri}, B^C_{cj}\rangle + \langle A^R_{ri+1}, B^C_{cj}\rangle - \langle A^R_{ri}, B^C_{cj+1}\rangle - \langle A^R_{ri+1}, B^C_{cj+1}\rangle)$$

Next, $\Gamma'_{i,j}$ is compared with $\psi(\Gamma)$

$$\psi(AB) = \psi(\Gamma)$$

Next, the column transform expansion is conducted:

$$(\Gamma) = \psi \langle A_{ri}, B_{cj}\rangle$$

$$\psi_{1C}(\Gamma)_{i,j} = \frac{1}{\sqrt{2}} \langle A_{ri}, B_{cj}\rangle + \langle A_{ri+1}, B_{cj}\rangle$$

$$\psi_{1C}(\Gamma)_{i,j+1} = \frac{1}{\sqrt{2}} \langle A_{ri}, B_{cj+1}\rangle + \langle A_{ri+1}, B_{cj+1}\rangle$$

Now the column transform results are combined, by subtraction in this case:

$$\psi(\Gamma) = \frac{1}{\sqrt{2}}(\psi_{1C}(\Gamma)_{i,j} - \psi_{1C}(\Gamma)_{i,j+1})$$

Lastly, the expanded form is expanded.

$$\psi(\Gamma) = \frac{1}{2}(\langle A_{ri}, B_{cj}\rangle + \langle A_{ri+1}, B_{cj}\rangle - \langle A_{ri}, B_{cj+1}\rangle - \langle A_{ri+1}, B_{cj+1}\rangle)$$

Case 4 $i \geq \frac{row}{2}$ and $j \geq \frac{col}{2}$. Like, before the base case for $\Gamma'_{i,j}$

$$\Gamma'_{ij} = \left\langle \frac{A^R_{ri,j} - A^R_{ri+1}}{\sqrt{2}}, \frac{B^C_{cj} - B^C_{cj+1}}{\sqrt{2}} \right\rangle$$

Again, the vector sums are expanded within their inner products.

$$\Gamma'_{i,j} = \frac{1}{2}(\langle A^R_{ri}, B^C_{cj}\rangle - \langle A^R_{ri+1}, B^C_{cj}\rangle - \langle A^R_{ri}, B^C_{cj+1}\rangle + \langle A^R_{ri+1}, B^C_{cj+1}\rangle)$$

Next, $\Gamma'_{i,j}$ is compared with $\psi(\Gamma)$

$$\psi(AB) = \psi(\Gamma)$$

$$(\Gamma) = \psi \langle A_{ri}, B_{cj}\rangle$$

Once the column wavelet transform expands Γ to

$$\psi_{1C}(\Gamma)_{i,j} = \frac{1}{\sqrt{2}} \langle A_{ri}, B_{cj}\rangle - \langle A_{ri+1}, B_{cj}\rangle$$

$$\psi_{1C}(\Gamma)_{i,j+1} = \frac{1}{\sqrt{2}} \langle A_{ri}, B_{cj+1}\rangle - \langle A_{ri+1}, B_{cj+1}\rangle$$

Now the column transform results are combined, by subtraction in this case:

$$\psi(\Gamma) = \frac{1}{\sqrt{2}}(\psi_{1C}(\Gamma)_{i,j} - \psi_{1C}(\Gamma)_{i,j+1})$$

Lastly, the expanded form is expanded.

$$\psi(\Gamma) = \frac{1}{2}(\langle A_{ri}, B_{cj}\rangle - \langle A_{ri+1}, B_{cj}\rangle - \langle A_{ri}, B_{cj+1}\rangle + \langle A_{ri+1}, B_{cj+1}\rangle)$$

Since the inner product of two vectors f, g is the same as their wavelet transformed vectors f', g', then all 4 cases produce a case where $\Gamma'_{i,j} = (\psi(\Gamma))_{i,j}\ \forall(i,j)$. Thus for the Haar Wavelet Transform (discrete mother basis),

$$\psi(A) \cdot \psi(B) = \psi(A \cdot B)$$

.

4.3 Multi Resolution Expansion Proof

In a previous section, wavelet based matrix multiplication was proven such that

$$\psi(A) \cdot \psi(B) = \psi(A \cdot B)$$

In this section the question of whether or not there is MRE form which is sound for matrix multiplication. There are few facts that are relevant and were made obvious in the empirical analysis in the results chapter.

- $\psi_{WPx}(A) \neq \psi^x(A)$ where $\psi_{WPx}(A)$ is the x resolution of wavelet transform packets (full decomposition) except for $x = 1$.
- $\psi_{Wx}(A) \neq \psi^x(A)$ where $\psi_{Wx}(A)$ is the x resolution of wavelet transform pyramids except for

$x = 1.$

The next hypothesis to be answered is can the wavelet transform be applied more than once to condition a matrix for matrix multiplication? This question is answered by the following lemma.

$$\psi^2(A) \cdot \psi^2(B) = \psi^2(A \cdot B)$$

Proof:

The theorem

$$\psi(A) \cdot \psi(B) = \psi(A \cdot B)$$

is proven as fact.

$$\psi^2(A) = \psi(\psi(A))$$

$$\psi^2(A) = \cdot\psi^2(B) = \psi(\psi(A)) \cdot \psi(\psi(B))$$

$$\psi(\psi(A)) \cdot \psi(\psi(B)) = \psi(\psi(A) \cdot \psi(B))$$

$$\psi(\psi(A) \cdot \psi(B)) = \psi(\psi(A \cdot (B))$$

$$\psi(\psi(A \cdot (B)) = \psi^2(A \cdot B)$$

Therefore:

$$\psi^2(A) \cdot \psi^2(B) = \psi^2(A \cdot B)$$

Chapter 5

Results

Each empirical result is from a series of matrix multiplications on 9 different images. These matrices were submitted preconditioning with the wavelet transform in the 3 basic method of decomposition, wavelet pyramid, wavelet packet (full decomposition) and the ψ^n decomposition.In the end, wavelet pyramids performed poorly for any set of resolutions deeper than one in matrix multiplication. Thus wavelet packets should not be used as a pre-conditioner to matrix multiplication. The wavelet transform by wavelet packets performs modestly in matrix multiplication tests, but to be effective must be reordered into ψ^n structure. Lastly, the psi^n form performs well on matrix multiply.

The matrix multiply used in each of these cases is $A^2 = A \cdot A$, where A is $M \times N$ The general equation for evaluating the wavelet transform pyramid method is the fidelity measurement:

$$\sqrt{E(\psi_{Pr}^{-x}((\psi_{Pr}^{x}(A))^2) - A^2)}$$

Relative Fidelity measurements provide a relative standard and is more useful for general observations. To obtain this measurement, the results are compared to standard.In this case, A^2 is the standard, and the energy difference is relative to that standard.

$$\sqrt{\frac{E(\psi_{Pr}^{-x}((\psi_{Pr}^{x}(A))^2) - A^2)}{E(A^2)}}$$

5.1 Empirical Analysis: Wavelet Transform Pyramids

Several standard images were used for measuring effect matrix multiply in each of these multiplications. Amongst them were the artificial images such as the waterfall and the clock. Others were actual images such as an arial view of the Pentagon, an in flight photo of an F-16, a fishing boat, an opera house, a river, and a set of bell peppers.

In first round analysis, matrix multiply was measured against the wavelet pyramids. These showed some undesirable results. As the decomposition proceeded deeper, the fidelity quick became unstable, and made the method unusable for matrix multiplication.

These tables are shown below.

Sample Two: Peppers 512×512 by wavelet transform pyramid

Resolution	Original Energy	Estimate Energy	Fidelity
1	12972.4	12972.4	$1.26464 \cdot 10^{-13}$
2	12972.4	12960.3	3.11211
3	12972.4	12947.8	5.44678
4	12972.4	12939.7	8.43738
5	12972.4	12926	13.1662
6	12972.4	12904.3	19.9257
7	12972.4	12877.9	25.4429

Sample Four: Waterfall 512×512 by wavelet transform pyramid

Resolution	Original Energy	Estimate Energy	Fidelity
1	6630.76	6630.76	$9.37836e - 14$
2	6630.76	6629.92	1.20049
3	6630.76	6626.6	2.37529
4	6630.76	6619.14	3.9194
5	6630.76	6595.14	6.63615
6	6630.76	6579.62	9.48774
7	6630.76	6567.17	12.1211

5.2 Empirical Analysis: Wavelet Transform Packets

The primary example shown is from the Peppers image. Although, all nine of the images used for this thesis could have been used, all of them exhibit similar qualities. After, two resolutions the 2-D MRE becomes unstable for matrix multiplication. It misplaces sections to where random noise in those sections would be just useful as the actual information contained in those sections.

Sample One: Peppers 512×512 by wavelet transform packet

Resolution	Original Energy	Estimate Energy	Fidelity
1	12972.4	12972.4	$1.26464e \cdot 10^{-13}$
2	12972.4	12972.4	0.059622
3	12972.4	12972.4	0.072957
4	12972.4	12972.4	0.0849579
5	12972.4	12972.4	0.09721
6	12972.4	12972.4	0.104528
7	12972.4	12972.4	0.119789

5.3 Empirical Results on the Psi-N Wavelet Transform Method

The Psi-N transform clearly produces better results as shown in all examples of the wavelet transform. Trade off issues start to appear when comparing the number of resolution levels to apply as opposed to a spare filter threshold.

In these next examples, resolution levels 1 through 7 are examined by removing data based on energy level. The method used is to strictly remove the lower threshold of energy from the matrix at large. The range of energy thresholds vary from 0 to $\frac{9}{10}$.

The reasoning for any of these schemes is to establish sparse representations of the matrix. In these sparse representation, only the above threshold elements are used. These representations are useful for sparse matrix multiplication since only the above threshold elements need be multiplied thus reducing the number of multiplications (operations at large) performed.

This filter is appropriately named for allowing the strongest elements to survive and weaker ones to be reduced to zero. This particular Darwin Filter is applied through out the matrix. In order to find the element that is the threshold element, each element in the matrix is squared, placed in a one dimensional array, and then sorted in the array from least to greatest. Lastly, two method exist to find the threshold value in the array of values. One is to sum up values until the threshold energy level has been reached. Another is to take element relative to the percentage threshold desired. The relative percentage method is simpler, by performing one division on the energy array size and a look up for the value in the energy array.

Once found, the filter threshold is used to filter out elements whose squared value is less than the threshold. Those elements that are less are made to be zero. Those that are above the threshold are left alone.

For this experiment both matrix operands have the filter applied. It should be noted that for the BLAS level three matrix multiply, this filtering scheme need only be applied to the first matrix operand. Applying this filter to the second is simply unnecessary since the method states that the second matrix can be dense.

The pattern of fidelity loss has a dramatic climb in the first $\frac{1}{20}$ of energy loss. From $\frac{1}{20}$ th to $\frac{1}{2}$ the fidelity stays roughly within the same order of magnitude. After this there is a sharp rise until all fidelity is lost.

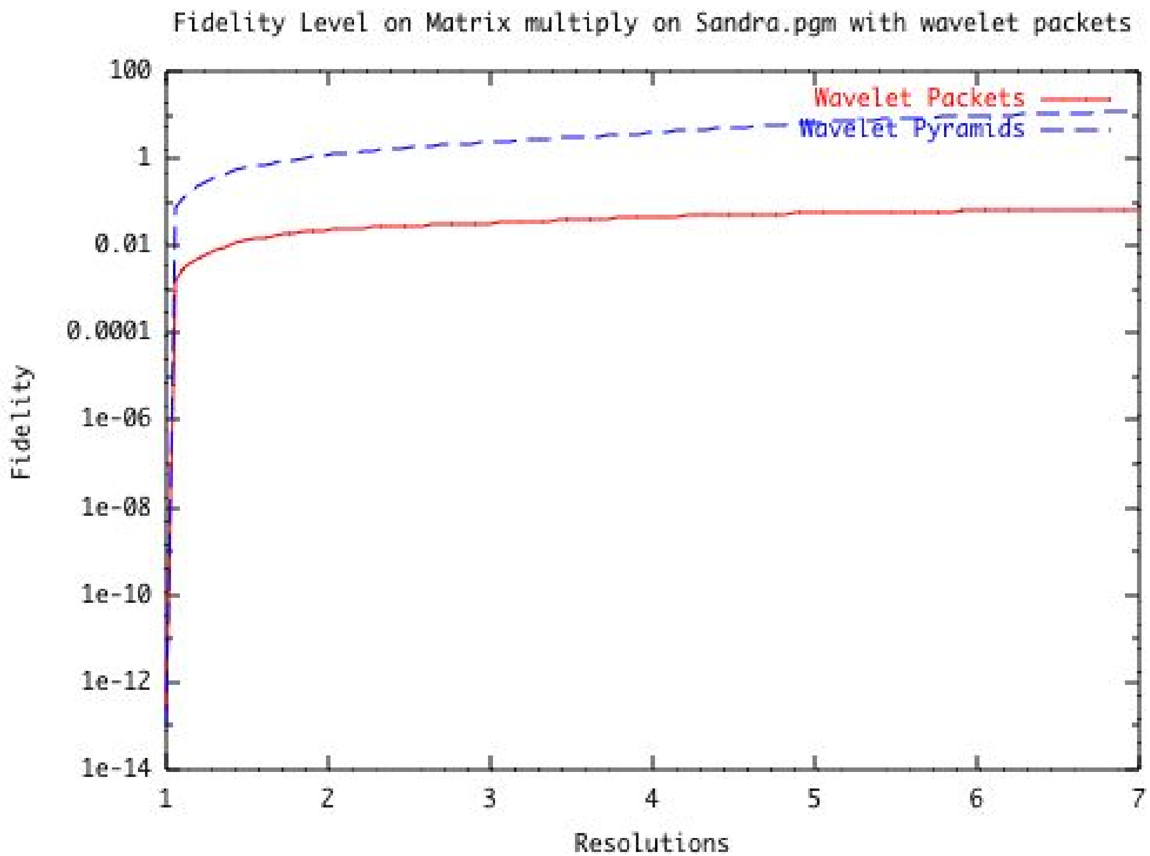

Figure 5.1: This fidelity level graph shows that the wavelet transform packet method as a preconditioner to matrix multiply is not reliable past the first resolution. The wavelet packet produces marginal results mostly due sections being placed out of order.

Figure 5.2: This image is a photo of an F-16 fighter jet provided Courtesy of the Signal and Image Processing Institute at the University of Southern California. [27]

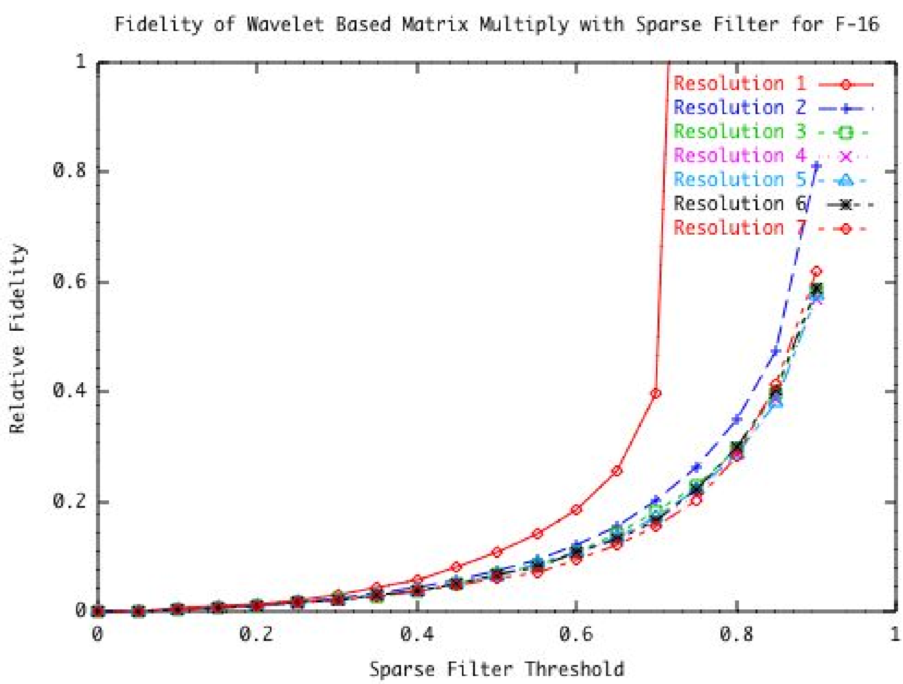

Figure 5.3: This fidelity level graph shows that Psi Wavelet Transform (full decomposition) approximation error in matrix multiplication. This graph in particular was obtained by multiplying the F-16 image by itself with various resolution levels of Wavelet Transforms applied. [27]

Figure 5.4: This image is a photo of an F-16 fighter jet provided Courtesy of the Signal and Image Processing Institute at the University of Southern California. [29]

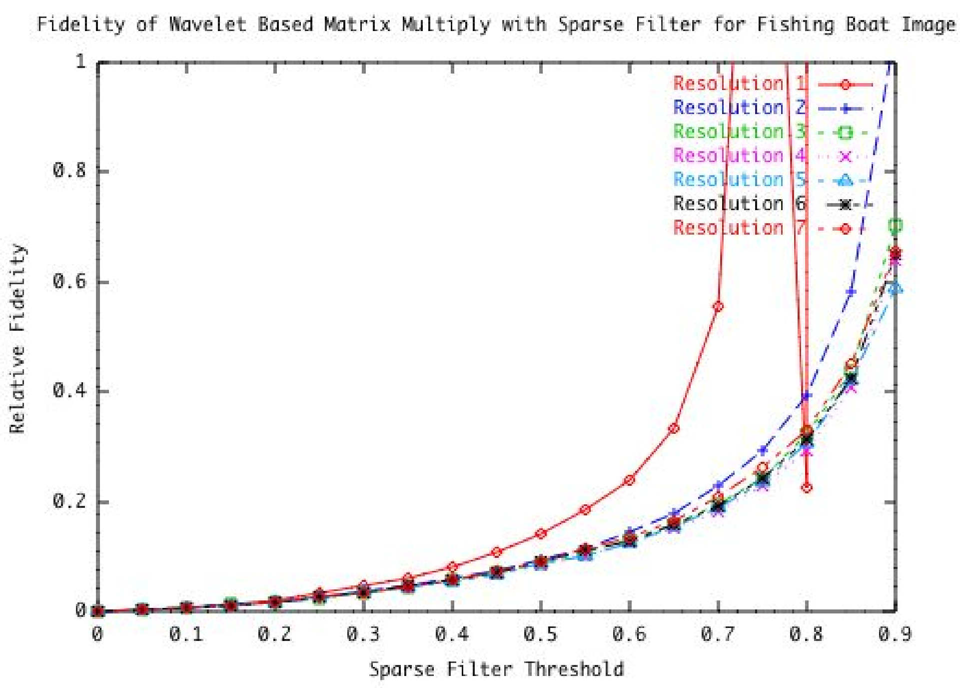

Figure 5.5: This fidelity level graph shows that Psi Wavelet Transform (full decomposition) approximation error in matrix multiplication. This graph in particular was obtained by multiplying the Fishing Boat image by itself with various resolution levels of Wavelet Transforms applied. [29]

Figure 5.6: This image is a photo of the Opera House in Lyon [25]

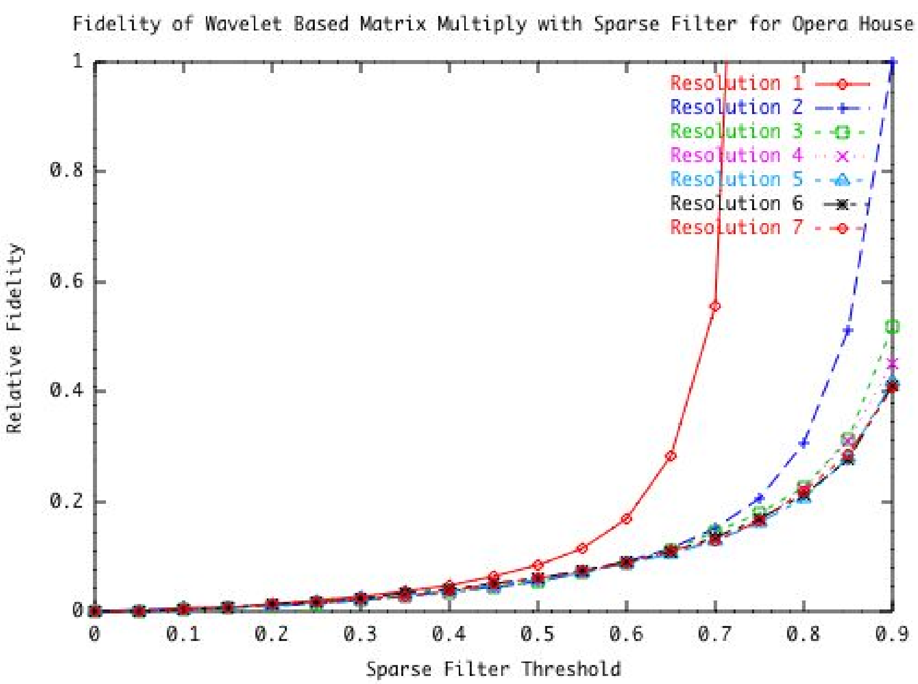

Figure 5.7: This fidelity level graph shows that Psi Wavelet Transform (full decomposition) approximation error in matrix multiplication. This graph in particular was obtained by multiplying the Opera House image by itself with various resolution levels of Wavelet Transforms applied.

Figure 5.8: This image is an ariel photo of an Pentagon provided Courtesy of the Signal and Image Processing Institute at the University of Southern California. [26]

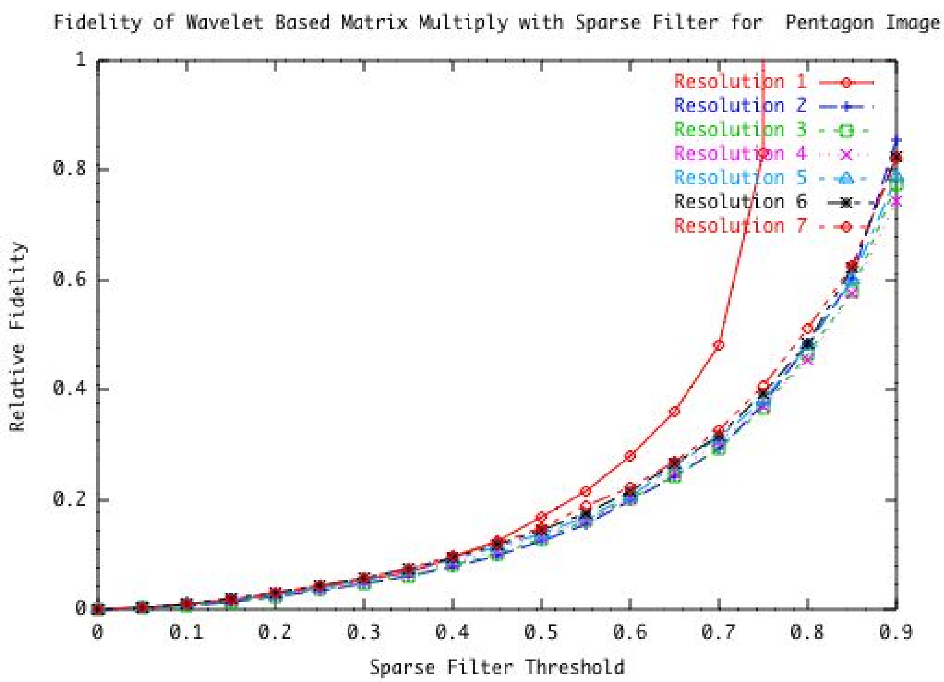

Figure 5.9: This fidelity level graph shows that Psi Wavelet Transform (full decomposition) approximation error in matrix multiplication. This graph in particular was obtained by multiplying the Pentagon image by itself with various resolution levels of Wavelet Transforms applied. [?]

Figure 5.10: This image is an ariel photo of “Peppers” provided Courtesy of the Signal and Image Processing Institute at the University of Southern California. [28]

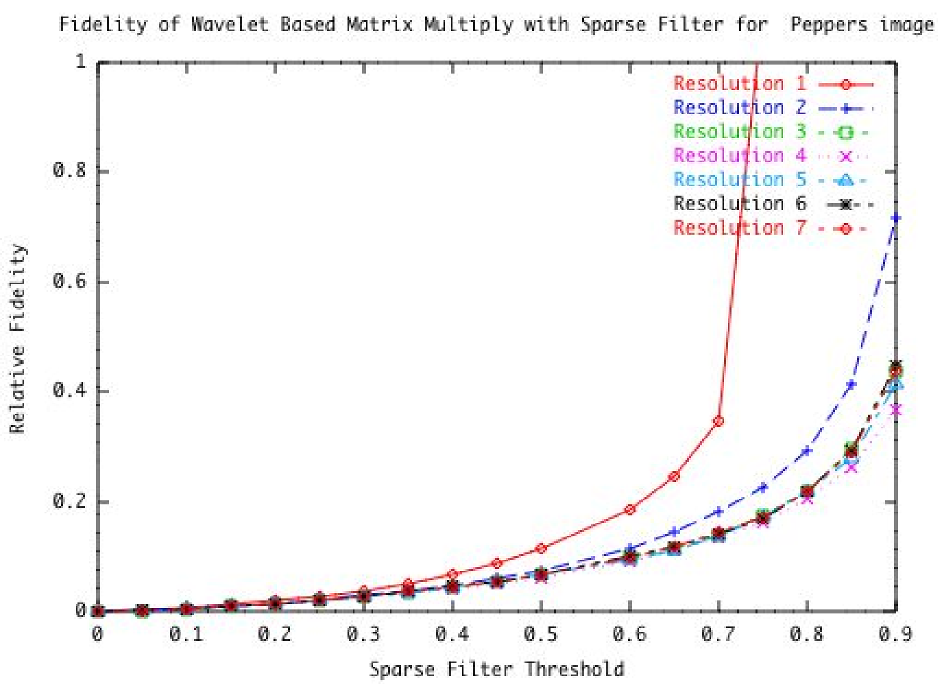

Figure 5.11: This fidelity level graph shows that Psi Wavelet Transform (full decomposition) approximation error in matrix multiplication. This graph in particular was obtained by multiplying the Peppers image by itself with various resolution levels of Wavelet Transforms applied. [28]

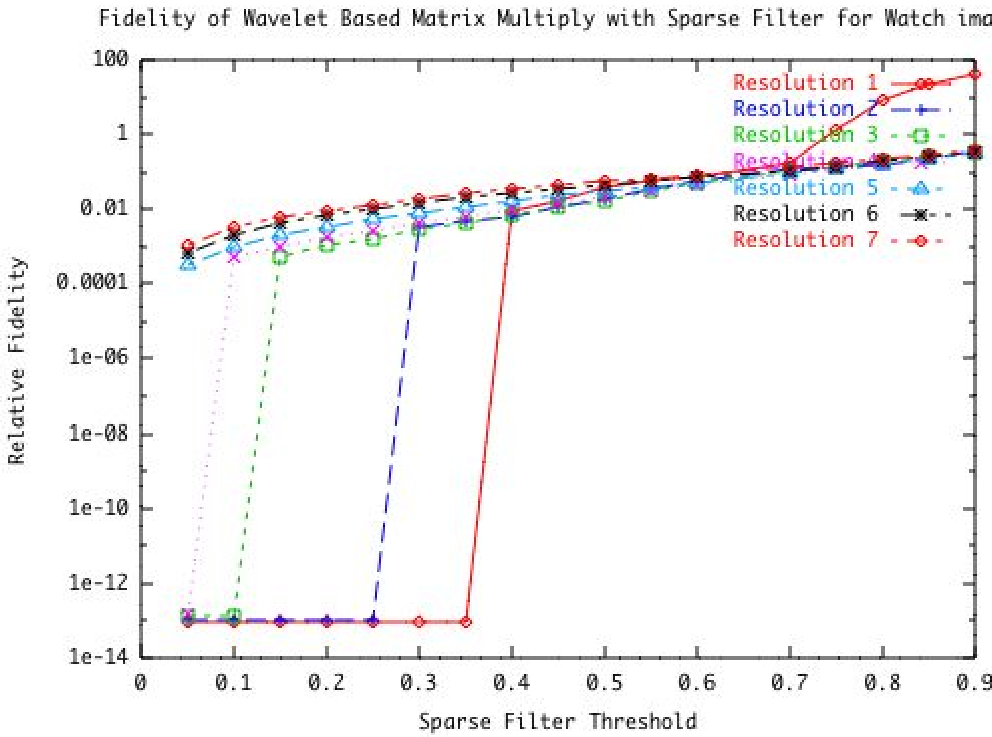

Figure 5.12: This fidelity level graph shows that Psi Wavelet Transform (full decomposition) approximation error in matrix multiplication. This graph in particular was obtained by multiplying the Watch image by itself with various resolution levels of Wavelet Transforms applied. [24]

Figure 5.13: "Pocket Watch on a Gold Chain. Copyright image courtesy of Kevin Odhner" [24]

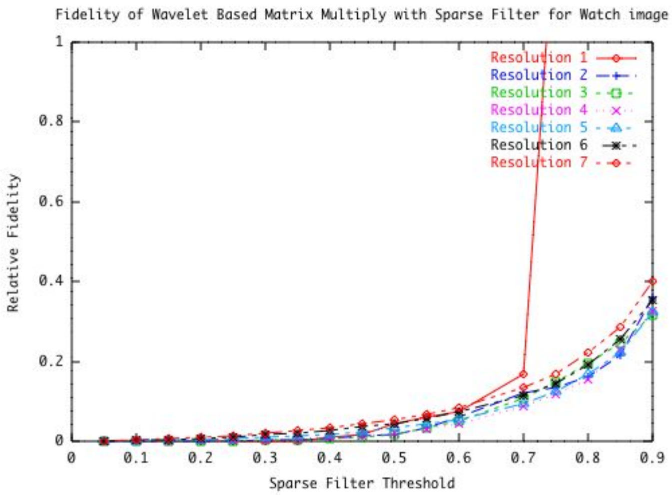

Figure 5.14: This fidelity level graph shows that Psi Wavelet Transform (full decomposition) approximation error in matrix multiplication. This graph in particular was obtained by multiplying the Watch image by itself with various resolution levels of Wavelet Transforms applied. [24]

Figure 5.15: This image is an ariel photo of “Always running, never the same....” provided Courtesy of the Jaime Vives Piqueres. [23]

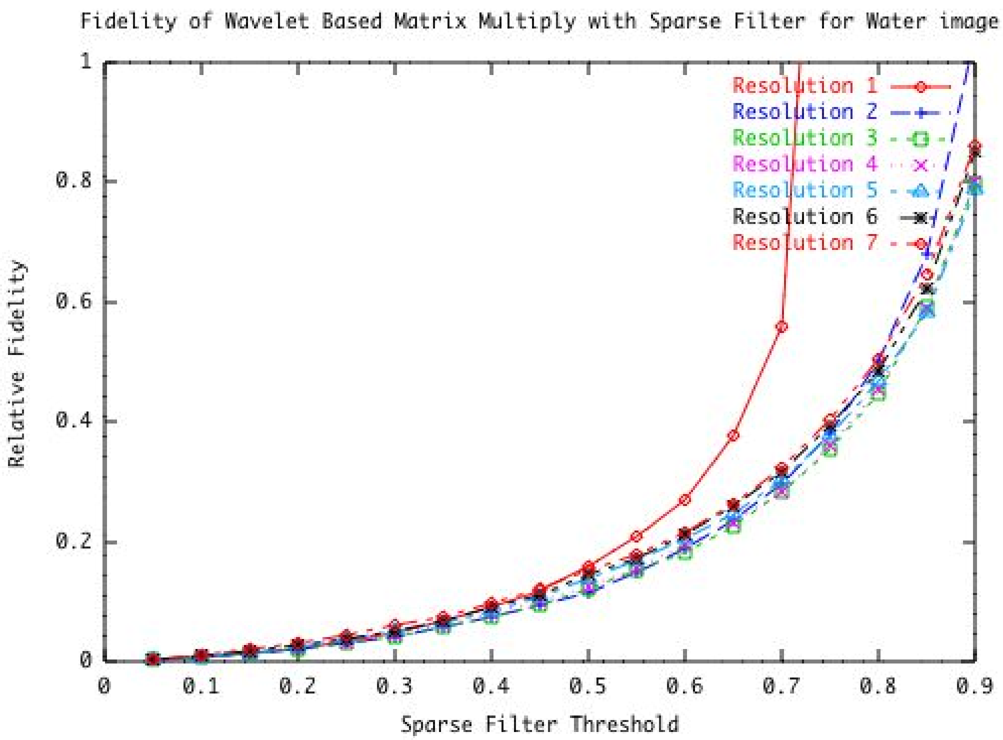

Figure 5.16: This fidelity level graph shows that Psi Wavelet Transform (full decomposition) approximation error in matrix multiplication. This graph in particular was obtained by multiplying the Water image by itself with various resolution levels of Wavelet Transforms applied. [24]

Figure 5.17: "This is a raytraced version of M.C.Escher's (1898-1972) famous lithography 'Waterfall' (1961), which again is based on a spatial illusion drawen by the mathematician Roger Penrose." [22]

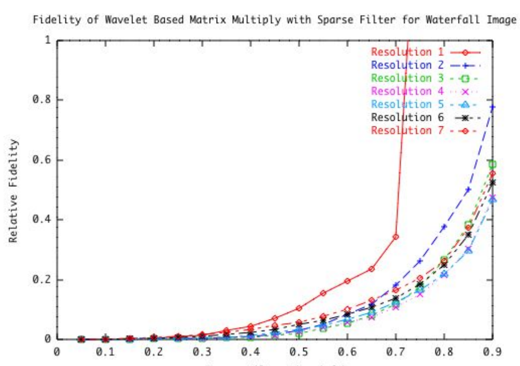

Figure 5.18: This fidelity level graph shows that Psi Wavelet Transform (full decomposition) approximation error in matrix multiplication. This graph in particular was obtained by multiplying the Waterfall image by itself with various resolution levels of Wavelet Transforms applied. [22]

Chapter 6

Conclusion

These following items shall be accomplished by this thesis: testing matrix multiplication and testing partial differential equations with the wavelet transform as a preconditioning function. The hypothesis for matrix multiplication is that a sparser matrix should emerge and therefore should be easier to multiply. In case of the partial differential equation, the hypothesis is that a better conditioned matrix should emerge.

In order to satisfy the matrix multiplication component, a prototype shall be developed. This prototype shall multiply two matrices together. Several wavelet bases shall be used to condition the matrix. The results shall be tested for their correctness.

For the wavelet based partial differential equation, another prototype shall be developed. This prototype shall solve the PDE within the standard implicit method to provide means of comparison. A few wavelet basis function shall be used to test their effectiveness in solving PDE. Also, a few multi-resolution methods shall be used to test their effectiveness.

Bibliography

[1] Charles Chui *An Introduction to Wavelets* published by Academic Press San Diego, CA 92101-4495 copyright 1992

[2] Howard L. Resnikoff. and Raymond O. Wells, Jr. *Wavelet Analysis: The Scalable Structure of Information* copyright Springer-Verlag New York, Inc. New York, NY 10010, USA, 1998

[3] James Walker *A Primer on Wavelets and Their Scientific Applications* copyright Chapman & Hall/CRC : Boca Raton, FL, USA, 1999

[4] Gavin Tabor "Wavelets - The Idiots Guide" Retrieved March 20, 2004 http://monet.me.ic.ac.uk/people/gavin/java/waveletDemos.html

[5] Amara Graps "Introduction to Wavelets" Retrieved March 20, 2004 from http://www.amara.com/current/wavelet.html

[6] Singiresu S. Rao *Applied Numerical Methods for Engineers and Scientists* published Prentice Hall Upper Saddle River, NJ 07458 copyright 2002

[7] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery *Numerical Methods in C* Published by the Press Syndicate of the University of Cambridge The Pitt

Building, Trumpington Street, Cambridge CB2 1RP 40 West 20th Street, New York, NY 10011-4211, USA, Copyright Cambridge University Press 1988, 1992

[8] G. Beylkin *Wavelets: Mathematics and Applications*, 1993 CRC Press , chapter "On wavelet-based algorithms for solving differential equations",

[9] G. Beylkin *Wavelets and Fast Numerical Algorithms AMS-93 Proceedings of Symposia in Applied Mathematics*, v.47, pp. 89-117, 1993

[10] G. Beylkin, R. Coifman, and V. Rokhlin *Fast Wavelet Transforms and Numerical Algorithms I, Article in Communications on pure and applied mathematics* copyright 1991 Wiley, New York

[11] G. Beylkin, D. Gines and L. Vozovoi "Adaptive Solution of Partial Differential Equations in Multiwavelet Bases " May 6, 1999 National Institute of Standards and Technology, Boulder, CO 80303-3328 and Department of Applied Mathematics, University of Colorado, Boulder, CO 80309-0526

[12] Stanly J. Farlow *Partial Differential Equations for Scientists and Engineers* copywrite 1993 Dover Publications, Inc Mineola, N.Y. 11501

[13] Murray R. Spiegel, *Theory and Problems of Advanced Mathematics for Engineers and Scientist* copy-write 1996 McGraw-Hill

[14] Stephane Jaffard, Yves Meyer, and Robert D. Ryan *Wavelets: Tools for Science and Technology* copyright 2001 by the Society for Industrial and Applied Mathematics Philadelphia, PA 19104-2688

[15] Mladen Victor Wickerhauser *Adapted Wavelet Analysis from Theory to Software* copyright 2001 by the Society for Industrial and Applied Mathematics Philadelphia, PA 19104-2688

[16] En-Bing Lin and Zhengchu Xiao *Multiwavelet Solutions for the Dirichlet Problem* Department of Mathematics, University of Toledo; Toledo, OH 43606, USA

[17] ian-Xiao He *Wavelet Analysis and Multiresolution Methods (Lecture Notes in Pure and Applied Mathematics)* published Marcel Dekker, Inc. New York, NY 10016 copyright 2000

[18] Seymour Lipschutz *Schaum's Outlines Theory and Problems of Linear Algebra, 2nd Edition* copywrite 1991, 1968 McGraw Hill New York

[19] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, editors. *Templates for the Solution of Algebraic Eigenvalue Problems: a Practical Guide*. SIAM, Philadelphia, 2000.

[20] Randolph E. Bank, Craig C. Douglas *Sparse Matrix Multiplication Package* April 23, 2001.

[21] Steven Huss-Lederman, Elaine M. Jacobson, J.R. Johnson, Anna Tsao, Thomas Turnbull *Strassen's Algorithm for Matrix Multiplication: Modeling, Analysis, and Implementation* November 15, 1996

[22] Sascha Ledinsky *Waterfall* http://oz.irtc.org/ftp/pub/stills/1998-10-31/waterfa1.txt Internet Raytracing Competition copyright October 31, 1998

[23] Jaime Vives Piqueres *Always running, never the same...* http://oz.irtc.org/ftp/pub/stills/1998-10-31/running.txt Internet Raytracing Competition copyright October 31, 1998

[24] Kevin Odhner *Pocket Watch on a Gold Chain*

[25] fabien a. p. petitcolas Opera House of Lyon open domain

[26] Signal and Image Processing Institute at the University of Southern California *Pentagon* The copyright status of this images in unclear.

[27] Signal and Image Processing Institute at the University of Southern California *F-16* The copyright status of this images in unclear.

[28] Signal and Image Processing Institute at the University of Southern California *Peppers* The copyright status of this images in unclear.

[29] Signal and Image Processing Institute at the University of Southern California *Fishing Boat* The copyright status of this images in unclear.

[30] Eric W. Weisstein et al. "Dual Space." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/DualSpace.html

[31] Eric W. Weisstein. "Translation." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Translation.html