



IMAGE REGISTRATION USING  
REDUNDANT WAVELET TRANSFORMS

THESIS

Richard K. Brown Jr., First Lieutenant, USAF

AFIT/GE/ENG/01M-21

DEPARTMENT OF THE AIR FORCE  
AIR UNIVERSITY

***AIR FORCE INSTITUTE OF TECHNOLOGY***

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

20010706 147

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

AFIT/GE/ENG/01M-21

# Image Registration Using Redundant Wavelet Transforms

THESIS

Presented to the Faculty

Department of Computer and Electrical Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the  
Degree of Master of Science in Electrical Engineering

Richard K. Brown, Jr., BS Systems Engineering

First Lieutenant, USAF

March, 2001

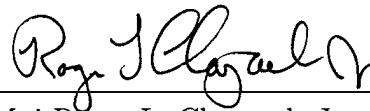
Approved for public release; distribution unlimited

## Image Registration Using Redundant Wavelet Transforms

Richard K. Brown, Jr., BS Systems Engineering

First Lieutenant, USAF

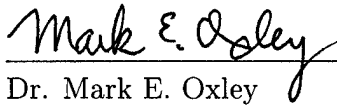
Approved:



Maj Roger L. Claypoole Jr.  
Thesis Advisor

5 MAR 01

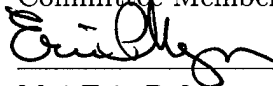
Date



Dr. Mark E. Oxley  
Committee Member

5 Mar 01

Date



Maj Eric P. Magee  
Committee Member

5 MAR 01

Date

## *Acknowledgements*

First, I would like to thank my thesis advisor, Maj Roger Claypoole, and thesis committee members, Maj Eric Magee and Dr. Mark Oxley, for providing useful commentary, criticism, and technical expertise.

Next, I would like to thank Harris Hall, Kristi Marcum, Ryan Thomas, Kyle Freundl, Alan Dean, Keven Golla, and the communications posse: Chris Bradley, Ben Crossley, Brian Dixon, Jason the Gregga, and Randy Klein. Their insight and support made the AFIT experience more enjoyable.

Special thanks to \_\_\_\_\_, \_\_\_\_\_, and \_\_\_\_\_, and \_\_\_\_\_, \_\_\_\_\_ for being great friends throughout the thesis process.

Finally, I want to give special thanks to Mike Mendenhall, who challenged me every step of the way and excelled as my thesis "coadvisor."

Richard K. Brown, Jr.

## *Table of Contents*

	Page
Acknowledgements . . . . .	iii
List of Figures . . . . .	viii
List of Tables . . . . .	x
Abstract . . . . .	xi
I. Introduction . . . . .	1-1
1.1 Problem Statement . . . . .	1-1
1.2 Scope . . . . .	1-2
1.3 Thesis Organization . . . . .	1-3
II. Background . . . . .	2-1
2.1 Overview . . . . .	2-1
2.2 Filter Banks . . . . .	2-1
2.2.1 Building Blocks . . . . .	2-1
2.2.2 Noble Identities . . . . .	2-3
2.3 The Wavelet Transform . . . . .	2-4
2.3.1 Wavelet Subspaces . . . . .	2-4
2.3.2 The Wavelet Recursion Equations . . . . .	2-5
2.3.3 Construction of the Discrete Wavelet Transform	2-6
2.3.4 Extension to Biorthogonal Wavelets . . . . .	2-9
2.3.5 The Two-Dimensional Discrete Wavelet Transform	2-10
form . . . . .	
2.3.6 Properties of Wavelets . . . . .	2-14
2.4 Image Registration . . . . .	2-15

	Page
2.4.1	Fundamentals of Image Registration . . . . . 2-15
2.4.2	Fourier Registration Techniques . . . . . 2-16
2.4.3	Limitations of Fourier Registration Techniques 2-17
2.4.4	Wavelet Registration Techniques . . . . . 2-18
2.4.5	Challenges Using Wavelet Techniques . . . . . 2-19
2.5	Summary . . . . . 2-21
III.	Methodology . . . . . 3-1
3.1	Overview . . . . . 3-1
3.2	The Wavelet Registration Concept . . . . . 3-1
3.3	Our Image Registration Algorithm . . . . . 3-3
3.3.1	Algorithm Flow for Translation Estimation . . 3-4
3.3.2	Shift-Invariant Wavelet Transform . . . . . 3-5
3.3.3	Subband Choice . . . . . 3-5
3.3.4	Masking . . . . . 3-8
3.3.5	Initial Estimation . . . . . 3-8
3.3.6	Final Estimation . . . . . 3-10
3.3.7	Algorithm Flow for Rotation Estimation . . . 3-10
3.3.8	Polar Redundant Wavelet Transform . . . . . 3-11
3.4	Summary . . . . . 3-16
IV.	Results . . . . . 4-1
4.1	Introduction . . . . . 4-1
4.2	Design of the Algorithm Validation Study . . . . . 4-1
4.2.1	Test Images . . . . . 4-1
4.2.2	Peak Signal-to-Noise Ratio . . . . . 4-2
4.2.3	Selection of Wavelet . . . . . 4-3
4.2.4	The Validation Study . . . . . 4-3

	Page
4.3 Translation Performance . . . . .	4-6
4.3.1 Redundant versus Standard Wavelet Transforms . . . . .	4-6
4.3.2 Subband Comparison . . . . .	4-9
4.4 Rotation Performance . . . . .	4-11
4.4.1 Redundant versus Standard Wavelet Transforms . . . . .	4-11
4.4.2 Subband Comparison . . . . .	4-13
4.5 Empirical Determination of $N$ . . . . .	4-15
4.6 Robustness . . . . .	4-17
4.7 Summary . . . . .	4-17
V. Discussion and Future Work . . . . .	5-1
5.1 Contributions of this Thesis . . . . .	5-1
5.2 Potential for Future Research . . . . .	5-1
5.2.1 Develop a More Robust Polar Wavelet Transform . . . . .	5-1
5.2.2 Leverage Multiscale Properties . . . . .	5-2
5.2.3 Calculate Translation and Rotation Simultaneously . . . . .	5-2
Appendix A. Additional Information on Filter Banks . . . . .	A-1
A.1 Polyphase Representation . . . . .	A-1
A.2 Perfect Reconstruction . . . . .	A-2
Appendix B. Additional Information on Wavelets . . . . .	B-1
B.1 Reconstruction . . . . .	B-1
B.2 Biorthogonal Wavelet Recursion Equations . . . . .	B-2
B.3 Construction of the Biorthogonal Discrete Wavelet Transform . . . . .	B-3
B.4 Reconstruction with the Biorthogonal Discrete Wavelet Transform . . . . .	B-3
B.5 Selection of Wavelet Filters . . . . .	B-4

	Page
Bibliography . . . . .	BIB-1
Vita . . . . .	VITA-1

## *List of Figures*

Figure		Page
2.1.	<i>M-channel filter bank. . . . .</i>	2-2
2.2.	<i>Examples of decimation and expansion. . . . .</i>	2-3
2.3.	<i>Nested subspaces of the orthogonal wavelet transform. . . . .</i>	2-5
2.4.	<i>Filter bank implementation for three iterations of the discrete wavelet transform. . . . .</i>	2-9
2.5.	<i>Example of biorthogonal wavelet spaces in <math>\mathcal{R}^3</math>. . . . .</i>	2-10
2.6.	<i>One iteration of the discrete wavelet transform of Lenna. . . . .</i>	2-11
2.7.	<i>Frequency responses for the filters used to create the subbands of the discrete wavelet transform. . . . .</i>	2-13
2.8.	<i>Directionality of the subbands and the multiscale effect of the discrete wavelet transform. . . . .</i>	2-14
3.1.	<i>Shift-invariance test. . . . .</i>	3-6
3.2.	<i>Comparison of the shift-invariant wavelet transform and the discrete wavelet transform. . . . .</i>	3-7
3.3.	<i>Camerman image sampled on a rectangular grid and on a polar grid. . . . .</i>	3-13
3.4.	<i>Rotation-invariant wavelet transform (one iteration). . . . .</i>	3-14
3.5.	<i>Rectangular versus polar sampling grids. . . . .</i>	3-15
3.6.	<i>Test for behavior of the polar redundant wavelet transform. . . . .</i>	3-16
4.1.	<i>Lenna and cameraman images. . . . .</i>	4-2
4.2.	<i>Examples of different PSNR values for the Lenna and cameraman images. . . . .</i>	4-4
4.3.	<i>Representation of the parsimony of several common wavelets. . . . .</i>	4-5
4.4.	<i>Translated and rotated versions of Lenna. . . . .</i>	4-7

Figure		Page
4.5.	<i>Translated and rotated versions of cameraman. . . . .</i>	4-8
4.6.	<i>Registration accuracy of the shift-invariant and discrete wavelet transforms. . . . .</i>	4-10
4.7.	<i>Registration accuracy of the detail subbands for the shift-invariant wavelet transform. . . . .</i>	4-12
4.8.	<i>Registration accuracy of the rotation-invariant and polar discrete wavelet transforms. . . . .</i>	4-14
4.9.	<i>Registration accuracy of the detail subbands for the rotation-invariant wavelet transform. . . . .</i>	4-16
A.1.	<i>Polyphase representation. . . . .</i>	A-3
A.2.	<i>Polyphase representation of an M-channel, maximally decimated uniform filter bank. . . . .</i>	A-4

*List of Tables*

Table		Page
3.1.	<i>Masking process. . . . .</i>	3-9
4.1.	<i>Algorithm robustness for Lenna. . . . .</i>	4-18
4.2.	<i>Algorithm robustness for cameraman. . . . .</i>	4-19

*Abstract*

Imagery is collected much faster and in significantly greater quantities today compared to a few years ago. Accurate registration of this imagery is vital for comparing the similarities and differences between multiple images. Image registration is a significant component in computer vision and other pattern recognition problems, medical applications such as Medical Resonance Images (MRI) and Positron Emission Tomography (PET), remotely-sensed data for target location and identification, and super-resolution algorithms. Since human analysis is tedious and error prone for large data sets, we require an automatic, efficient, robust, and accurate method to register images.

Wavelet transforms have proven useful for a variety of signal and image processing tasks, including image registration. In our research, we present a fundamentally new wavelet-based registration algorithm utilizing redundant transforms and a masking process to suppress the adverse effects of noise and improve processing efficiency. The shift-invariant wavelet transform is applied in translation estimation and a new rotation-invariant polar wavelet transform is effectively utilized in rotation estimation. We demonstrate the robustness of these redundant wavelet transforms for the registration of two images (i.e., translating or rotating an input image to a reference image), but extensions to larger data sets are certainly feasible. We compare the registration accuracy of our redundant wavelet transforms to the “critically sampled” discrete wavelet transform using the Daubechies (7,9) wavelet to illustrate the power of our algorithm in the presence of significant additive white Gaussian noise and strongly translated or rotated images.

# Image Registration Using Redundant Wavelet Transforms

## *I. Introduction*

### *1.1 Problem Statement*

Image registration is the process which determines the best match of two or more images acquired at the same or different times by identical or different sensors. It is a necessary intermediate step when image analysts need to compare the similarities and differences between multiple images. Examples of areas where image registration is a significant component include computer vision and other pattern recognition problems, medical applications such as Medical Resonance Images (MRI) and Positron Emission Tomography (PET), and remotely-sensed data for target location and identification. Image registration also serves as the front-end for super-resolution algorithms, which increase resolution by overlaying multiple coarse images of the same object.

The military has a need for timely, accurate analysis of imagery to be used in intelligence gathering and operations planning. Commanders must have the best information as quickly as possible on which to base decisions regarding troop, ship, and aircraft movements. For example, during the conflict in Yugoslavia in 1999, unmanned aerial vehicles (UAVs) were used extensively to collect data concerning the location of enemy surface-to-air missile sites, communication centers, and troops. Operations planners relied heavily on this data to decide which enemy targets to strike. The quality of the analysis of the imagery depends greatly on the accuracy of the image registration.

Image analysts have typically had to analyze imagery manually by comparing multiple pictures of the same object(s) and attempting to extract the differences.

Imagery is collected much faster and in significantly greater quantities today compared to a few years ago. Additionally, multiple images of the same object(s) taken from satellites, UAVs, or other aircraft necessarily have different perspectives since the observing satellite, UAV, or aircraft is not stationary and the object(s) of interest may also be moving. These challenges make human analysis intractable and error prone for large data sets. An automatic, efficient, robust, and accurate method to register images is necessary.

Our research analyzes the use of redundant wavelet transforms in image registration. The shift-invariant wavelet transform and the rotation-invariant polar wavelet transforms are derived and applied. A fundamentally new, robust method to register images is developed, resulting in a more accurate registration of imagery and providing a sound front-end for super-resolution algorithms.

## *1.2 Scope*

Our research will demonstrate the robustness of applying redundant wavelet transforms to the image registration problem. Although our registration algorithm only addresses the registration of two images (i.e., translating or rotating an input image to a reference image), the extension to a larger data set is certainly feasible but not necessary to test the validity of our algorithm. Emphasis is on validation of the algorithm using the shift-invariant wavelet transform and the newly developed rotation-invariant polar wavelet transform, ensuring accurate registration is achieved in the presence of noise. The algorithm registers images that have been strongly translated or rotated, but not both. Minimal effort is made to optimize the speed of registration. Validation of our registration method is accomplished by applying it to strongly translated or rotated noisy versions of the “Lenna” and “cameraman” images commonly used in image processing.

### *1.3 Thesis Organization*

Chapter 2 provides the background of this thesis. Fundamentals of filter banks and wavelet theory are discussed. The relationship between filter banks and the discrete wavelet transform is analyzed. A review of current Fourier and wavelet registration techniques and their limitations provides a baseline for comparing these techniques to our registration algorithm.

Chapter 3 describes our image registration algorithm in detail. The shift-invariant wavelet transform used in translation estimation and the rotation-invariant polar wavelet transform utilized in rotation estimation are developed. The new concept of masking, which suppresses the adverse effects of noise and increases computational efficiency, is presented. Our process of generating initial estimates from the wavelet bandpass subbands and then refining these estimates to achieve the final estimate is explored.

Chapter 4 provides a validation study on the accuracy of our registration algorithm in the presence of mild and significant additive white Gaussian noise. Our shift-invariant and rotation-invariant wavelet transforms are shown to provide superior registration accuracy over the critically sampled discrete wavelet transform. Our decision to use the bandpass subbands over the highpass subband is shown to be sound and we empirically determine the best choice of significant wavelet coefficients to keep for feature matching for our test images. Finally, we show our algorithm is robust since it accurately estimates rotation when our test images are both translated and rotated.

Chapter 5 summarizes the results and outlines recommendations for future research.

## II. Background

### 2.1 Overview

The basics of filter banks, which are used in the implementation of the discrete wavelet transform are presented first. Next, we discuss how the discrete wavelet transform is constructed from orthogonal subspaces and we derive the wavelet recursion equations, which govern all orthogonal wavelets. We then move to biorthogonal wavelets which have useful properties, but do not obey the wavelet recursion equations as orthogonal wavelets do. The two-dimensional discrete wavelet transform is then developed and we finish our discussion on wavelets by stating some properties that make them attractive for registration. Finally, we present some Fourier and wavelet registration techniques and their limitations. These serve as a baseline for comparison to our registration algorithm described in Chapter 3.

### 2.2 Filter Banks

We provide a quick introduction to filter banks required to understand their use in the implementation of both the discrete wavelet transform and the redundant wavelet transform, which is presented in the next chapter. For more rigorous descriptions of filter banks, refer to (5, 6, 29, 32, 33). Additional information is also located in Appendix A.

*2.2.1 Building Blocks.* A filter bank is a collection of filters with a common input  $x(n)$  to  $M$  channels and a single output  $\hat{x}(n)$  as shown in Figure 2.1. The set of filters  $H_k(z)$  are referred to as the analysis bank and the set of filters  $F_k(z)$  are known as the synthesis bank (32, 33). For the purpose of implementing the discrete wavelet transform, we are only interested in maximally decimated uniform filter banks. That is, filter banks which have the same number of inputs and outputs

(input one image, output one image) and follow the relationship  $\sum_k 1/M_k = 1$ , where  $M_k$  is the decimation ratio in the  $k^{th}$  channel (32).

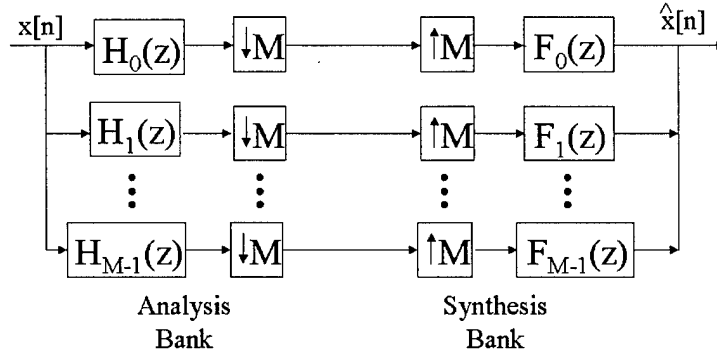


Figure 2.1. *M-channel filter bank. The analysis bank has one input and M outputs. The synthesis bank takes the M channels to a single output.*

The two primary building blocks of filter banks are the decimator and expander, both of which are linear, shift-varying elements (32). The decimator is defined as

$$y_D(n) = x(Mn)$$

where  $M$  is a positive integer known as the decimation ratio. The expander is defined as

$$y_E(n) = \begin{cases} x(n/L) & , \text{ if } n \text{ is an integer multiple of } L \\ 0 & , \text{ else} \end{cases}$$

where  $L$  is a positive integer known as the expansion ratio. Decimation by  $M = 2$  keeps every other point of a sequence  $x(n)$ , decimation by  $M = 3$  keeps every third point, etc. Similarly, expansion by  $L = 2$  places a zero between every point of a sequence  $x(n)$ , expansion by  $L = 3$  places 2 zeros between every point, etc. See Figure 2.2 for examples of decimation and expansion.

For  $y(n) = x(Mn)$ , where  $M$  is the decimation ratio, the corresponding notation is  $y(n) = (x(n)) \downarrow_M$ . Likewise, for  $y(n) = x(n/L)$ , where  $L$  is the expansion

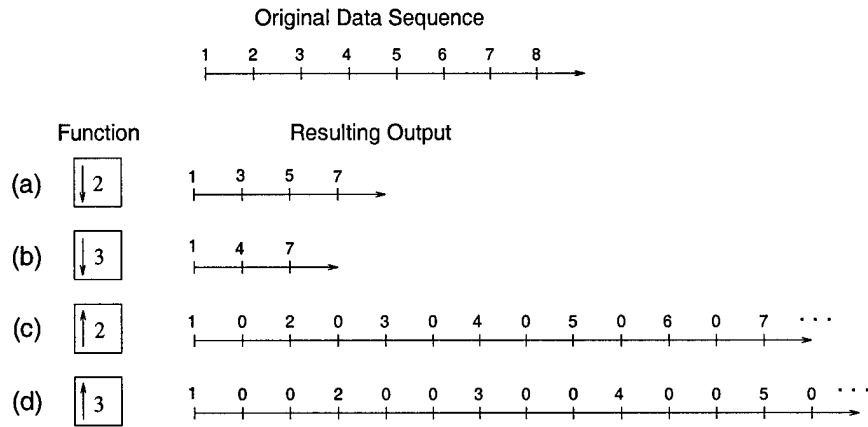


Figure 2.2. Examples of decimation and expansion. (a) Decimation by 2. (b) Decimation by 3. (c) Expansion by 2. (d) Expansion by 3.

ratio, the corresponding notation is  $y(n) = (x(n)) \uparrow_L$ . Figure 2.2 shows the block diagram for an expander and decimator.

**2.2.2 Noble Identities.** Having presented the fundamental building blocks of filter banks, we now discuss how to connect them (5, 32). Using the notation defined in Section 2.2.1, the following equations describe the different ways in which decimators may be combined

$$\alpha\{x(n)\} \downarrow_M = \{\alpha x(n)\} \downarrow_M, \alpha \text{ any scalar} \quad (2.1)$$

$$\{x_1(n) + x_2(n)\} \downarrow_M = \{x_1(n)\} \downarrow_M + \{x_2(n)\} \downarrow_M \quad (2.2)$$

$$\{x_1(n) \times x_2(n)\} \downarrow_M = \{x_1(n)\} \downarrow_M \times \{x_2(n)\} \downarrow_M. \quad (2.3)$$

These equations still hold when the decimators are replaced by expanders (32).

Cascaded decimators and expanders may only be interchanged when the decimation ratio  $M$  and the expansion ratio  $L$  are relatively prime (i.e., the lowest common denominator between  $M$  and  $L$  is 1). Thus, the following becomes valid when  $M$  and  $L$  are relatively prime

$$\{\{x(n)\} \downarrow_M\} \uparrow_L = \{\{x(n)\} \uparrow_L\} \downarrow_M .$$

For a sequence  $x(n)$  and digital filter  $h(n)$  with z-transforms  $X(z)$  and  $H(z)$ , respectively, this leads directly to

$$\{X(z)H(z^M)\} \downarrow_M = \{X(z) \downarrow_M\}H(z) \quad (2.4)$$

$$\{X(z)H(z)\} \uparrow_L = \{X(z) \uparrow_L\}H(z^L) \quad (2.5)$$

These equations, which are only valid when  $H(z)$  is comprised of integer powers of  $z$ , are known as the Noble Identities. They are the rules we must follow when interchanging filters with our building blocks (32). Armed with this basic knowledge of filter banks, we move on to wavelets.

### 2.3 The Wavelet Transform

We present some useful properties and provide a concise overview of the discrete wavelet transform in this section. See (3, 6, 8, 22, 29) for more extensive information on wavelets.

*2.3.1 Wavelet Subspaces.* Wavelet transforms provide an efficient multi-scale representation by projection of a function into smaller orthogonal subspaces formed from shifts and dilations of a lowpass scaling function  $\phi(t)$  and a highpass wavelet function  $\psi(t)$  (8).

Let  $\{V_m\}_{m \in \mathcal{Z}}$  be a sequence of nested subspaces in  $L^2(\mathfrak{R})$  such that  $V_m \subset V_{m-1} \forall m \in \mathcal{Z}$ . Let  $f$  be a function which exists in the subspace  $V_{m-1}$  for some  $m$ . Let  $P_m$  be the operator which projects  $f$  into the nested subspace  $V_m \subset V_{m-1}$ .  $P_m$  preserves the portion of  $f$  that lies in  $V_m$  and eliminates the part of  $f$  that is not in  $V_m$ .

Let  $\{W_m\}_{m \in \mathcal{Z}}$  be a different sequence of nested subspaces in  $L^2(\mathbb{R})$  such that  $W_m \subset V_{m-1} \forall m \in \mathcal{Z}$ . Let the span of  $V_m \cup W_m$  equal  $V_{m-1}$ , where  $V_m$  and  $W_m$  are orthogonal subspaces. Thus,  $V_{m-1} = V_m \oplus W_m$  and  $V_m \cap W_m = \phi$ . See Figure 2.3.

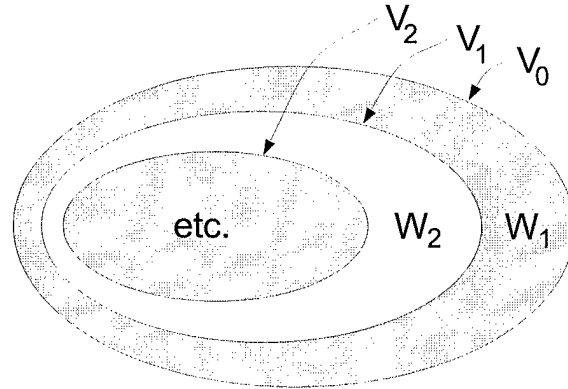


Figure 2.3. *Nested subspaces of the orthogonal wavelet transform. The  $V_m$  are spanned by shifts and dilations of a lowpass scaling function while the  $W_m$  are spanned by shifts and dilations of a highpass wavelet function.*

Now let  $Q_m = I - P_m$  (where  $I$  is the identity operator) be the operator which projects  $f$  into the nested subspace  $W_m \subset V_{m-1}$ . As before,  $P_m$  preserves the portion of  $f$  that lies in  $W_m$  and eliminates the part of  $f$  that is not in  $W_m$ . Finally,  $P_m Q_m = Q_m P_m = 0$ , where  $0$  is the zero operator.

**2.3.2 The Wavelet Recursion Equations.** Let the set of functions  $\{\phi_{m,n}\}_{n \in \mathcal{Z}}$  be an orthonormal basis for  $V_m$  and let the set of functions  $\{\psi_{m,n}\}_{n \in \mathcal{Z}}$  be an orthonormal basis for  $W_m$ . We can expand  $\phi_{m,n}$  in terms of  $\{\phi_{m-1,k}\}$  since  $V_m \subset V_{m-1}$ . Thus,

$$\phi_{m,n}(t) = \sum_k c_{m-1,n}(k) \phi_{m-1,k}(t).$$

To find a particular coefficient  $c_{m-1,n}(k)$ , we take the inner product of  $\phi_{m,n}$  with  $\phi_{m-1,k}$ , where the inner product of two real valued functions  $f$  and  $g$  is defined as

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

If we let  $h_n(k) = \langle \phi_{m,n}, \phi_{m-1,k} \rangle$ , the expansion of  $\phi_{m,n}$  now becomes

$$\phi_{m,n}(t) = \sum_k h_n(k) \phi_{m-1,k}(t).$$

The set of functions  $\{\phi_{m,n}\}$  must satisfy the above equation for all  $m$  to achieve an orthogonal wavelet transform. To form the discrete wavelet transform, we must further impose that each  $\{\phi_{m,n}\}$  be constructed from integer shifts ( $n$ ) and dilations ( $2^m$ ) of our scaling function  $\phi(t)$ . Thus, our expression for  $\{\phi_{m,n}\}$  becomes

$$\phi(t - n) = \sum_k h(k - 2n) \phi(2t - k). \quad (2.6)$$

Performing the same derivation for  $\{\psi_{m,n}\}$ , we have

$$\psi(t - n) = \sum_k g(k - 2n) \phi(2t - k), \quad (2.7)$$

where  $g(k - 2n) = \langle \psi_{m,n}, \phi_{m-1,k} \rangle$ . Equations 2.6 and 2.7 are the wavelet recursion equations (3).

*2.3.3 Construction of the Discrete Wavelet Transform.* In Sections 2.3.1 and 2.3.2, we constructed a set of orthogonal subspaces to permit the multiscale decomposition of a signal  $f$  and we derived the wavelet recursion equations which govern our orthogonal wavelet transform. Now, we derive the discrete wavelet decomposition of  $f$ .

First, we assume that  $f$  exists entirely in  $V_{m-1}$  and may be completely represented as a linear combination of basis functions for subspace  $V_{m-1}$  for  $m \in \mathcal{Z}$ , where typically  $m = 1$ .

The projection of  $f$  into  $V_m$  can be expanded in terms of the basis functions of  $V_m$  and the projection of  $f$  into  $W_m$  can be expanded in terms of the basis functions

of  $W_m$ . Thus, we have

$$\begin{aligned} [P_m f](t) &= \sum_n c_m(n) \phi_{m,n}(t) \\ [Q_m f](t) &= \sum_n d_m(n) \psi_{m,n}(t), \end{aligned}$$

with  $c_m(n) = \langle P_m f, \phi_{m,n} \rangle$  and  $d_m(n) = \langle Q_m f, \psi_{m,n} \rangle$ .

Since  $P_m + Q_m = I$ , we can write

$$f = P_m f + Q_m f,$$

and any coefficient  $c_m(n)$  in  $V_{m-1}$  can be written as

$$\begin{aligned} c_m(n) &= \langle P_m f, \phi_{m,n} \rangle \\ &= \langle (f - Q_m f), \phi_{m,n} \rangle \\ &= \langle f, \phi_{m,n} \rangle - \langle Q_m f, \phi_{m,n} \rangle \\ &= \langle f, \phi_{m,n} \rangle. \end{aligned}$$

If  $f$  is expanded in terms of the basis functions for  $V_{m-1}$  and substituted into the expression above, we have

$$\begin{aligned} c_m(n) &= \left\langle \left( \sum_k c_{m-1}(k) \phi_{m-1,k} \right), \phi_{m,n} \right\rangle \\ &= \sum_k c_{m-1}(k) \langle \phi_{m-1,k}, \phi_{m,n} \rangle. \end{aligned}$$

Similarly, it can be shown that

$$d_m(n) = \sum_k c_{m-1}(k) \langle \phi_{m-1,k}, \psi_{m,n} \rangle.$$

Recall from Section 2.3.2 that for the case of the discrete wavelet transform, the inner products in the expressions for  $c_m(n)$  and  $d_m(n)$  are digital filters  $h$  and  $g$  independent of decomposition level

$$\begin{aligned} h(k - 2n) &= \langle \phi_{m-1,k}, \phi_{m,n} \rangle \\ g(k - 2n) &= \langle \phi_{m-1,k}, \psi_{m,n} \rangle. \end{aligned}$$

It follows that if we know the coefficients  $c_{m-1}(n)$  of  $f \in V_{m-1}$ , then the coefficients of the projection of  $f$  into  $V_m$  and  $W_m$ , respectively, are given by

$$c_m(n) = \sum_k c_{m-1}(k)h(k - 2n) \quad (2.8)$$

$$d_m(n) = \sum_k c_{m-1}(k)g(k - 2n). \quad (2.9)$$

Now that we have the coefficients of the projection of  $f$  into  $V_m$  and  $W_m$ ,  $f$  can be further decomposed into  $V_{m-1}$  and  $W_{m-1}$ , the orthogonal subspaces of  $V_m$ , using the same formulas since the digital filters  $h$  and  $g$  are independent of decomposition level  $m$ . We now have a recursive routine to decompose  $f$  into smaller orthogonal subspaces. This makes the discrete wavelet transform multiscale: it consists of a set of scaling (coarse) coefficients  $c_m(n)$ , which represent coarse signal information at scale  $m = M$ , and a set of wavelet (detail) coefficients  $d_m(n)$ , which represent detail signal information at scales  $m = 1, 2, \dots, M$  (8).

Equations 2.8 and 2.9 provide insight into the filter bank implementation of the discrete wavelet transform (Figure 2.4). We see that the discrete wavelet transform can be implemented using cascades of two channel filter banks with analysis bank digital filters  $h$  and  $g$ , where  $h$  and  $g$  are traditionally lowpass and highpass, respectively (3). If  $h$  and  $g$  satisfy their respective wavelet decomposition requirements,

$$h(k - 2n) = \langle \phi_{m-1,n}, \phi_{m,n} \rangle$$

$$g(k - 2n) = \langle \phi_{m-1,n}, \psi_{m,n} \rangle,$$

then they form a perfect reconstruction set (the transform is invertible using these filters) (32). A detailed explanation of reconstruction of the original signal from its wavelet decomposition is contained in Appendix B.1.

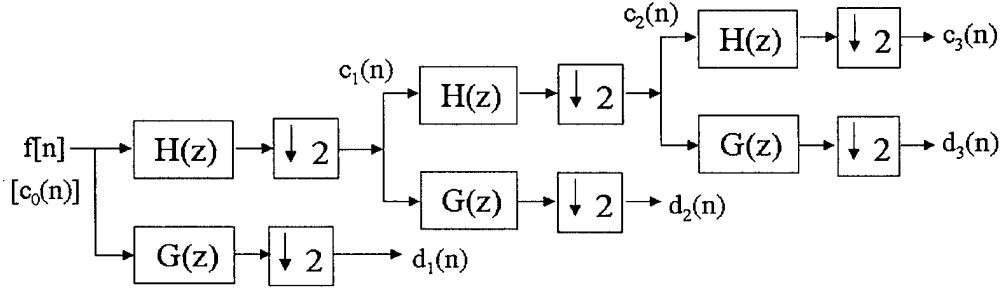


Figure 2.4. *Filter bank implementation for three iterations of the discrete wavelet transform. Coarse coefficients  $c_{m-1}(n)$  are created by convolving the original signal  $f[n]$  with lowpass filter  $h$  and downsampling by two. Detail coefficients  $d_m(n)$  are produced by convolving the original signal with highpass filter  $g$  and downsampling by two. Lower scales are formed by iterating on the coarse coefficients of the next highest scale.*

2.3.4 *Extension to Biorthogonal Wavelets.* Everything to this point in our analysis has assumed orthogonal spaces, which forces the analysis and synthesis filters to be identical. Now, we loosen this restriction and form a biorthogonal wavelet transform with dual spaces  $\tilde{V}_m$  and  $\tilde{W}_m$  where  $\tilde{V}_m = \text{span}\{\tilde{\phi}_{m,k}\}$  and  $\tilde{W}_m = \text{span}\{\tilde{\psi}_{m,k}\}$  (8). For biorthogonal spaces, we have  $V_m \cap \tilde{W}_m = \phi$  and  $\tilde{V}_m \cap W_m = \phi$ . Also, the basis functions must satisfy the biorthogonality condition,

$$\langle \phi_{m,k}, \tilde{\phi}_{m,n} \rangle = \delta(k - n) \quad (2.10)$$

$$\langle \psi_{m,k}, \tilde{\psi}_{m,n} \rangle = \delta(k - n). \quad (2.11)$$

Figure 2.5 demonstrates the biorthogonal decomposition of  $V_0 = \mathbb{R}^3$ .  $V_1$  is spanned by  $\phi_0$  and  $\phi_1$ , while  $\tilde{V}_1$  is spanned by  $\tilde{\phi}_0$  and  $\tilde{\phi}_1$ .  $V_1$  is orthogonal to its dual

space  $\tilde{W}_1$ . Likewise,  $\tilde{V}_1$  is orthogonal to  $W_1$ . All basis functions satisfy Equations 2.10 and 2.11, the biorthogonality condition.

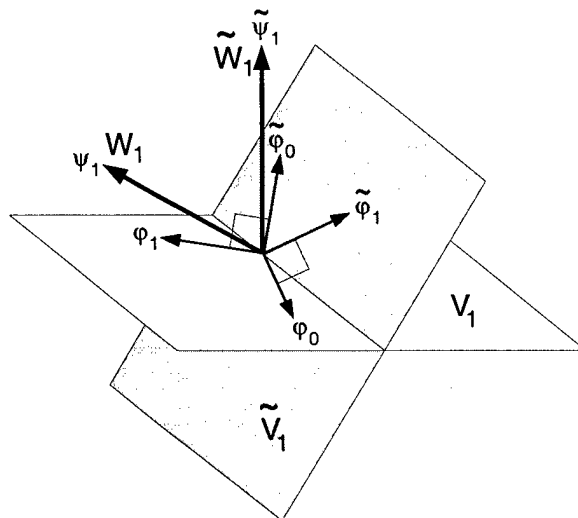


Figure 2.5. *Example of biorthogonal wavelet spaces in  $\mathbb{R}^3$ .  $\text{Span}(V_1 \cup W_1) = \mathbb{R}^3$  and  $\text{span}(\tilde{V}_1 \cup \tilde{W}_1) = \mathbb{R}^3$ .  $V_1 \perp \tilde{W}_1$  and  $\tilde{V}_1 \perp W_1$ . All basis functions satisfy the biorthogonality condition.*

Biorthogonal wavelets are more desirable than orthogonal wavelets for two reasons. First, they allow for the easy design of linear phase filters, which are important because phase information is more significant than magnitude in image reconstruction (19). All orthogonal wavelets are composed of even length filters. Filters that produce biorthogonal wavelets may be of odd length. Thus, we may design symmetric biorthogonal filters to produce linear phase. Next, biorthogonal wavelets allow for larger size linear phase filters. The only orthogonal wavelet filters that have linear phase are the Haar filters, which are of length two. In general, larger size filters correspond to smoother wavelet functions and a more parsimonious signal representation (29). More specific details concerning biorthogonal wavelets can be found in Appendix B.2.

*2.3.5 The Two-Dimensional Discrete Wavelet Transform.* We transform two-dimensional signals (images) by first processing the rows and then the columns

in a separable fashion. Figure 2.6 a and b show the Lenna image and one iteration of the discrete wavelet transform of Lenna, respectively. The discrete wavelet transform is a dyadic decomposition; the rows and columns of the original image are filtered and downsampled by a factor of two. Each of the four quadrants is known as a subband. The subbands preserve certain characteristics of the original image.

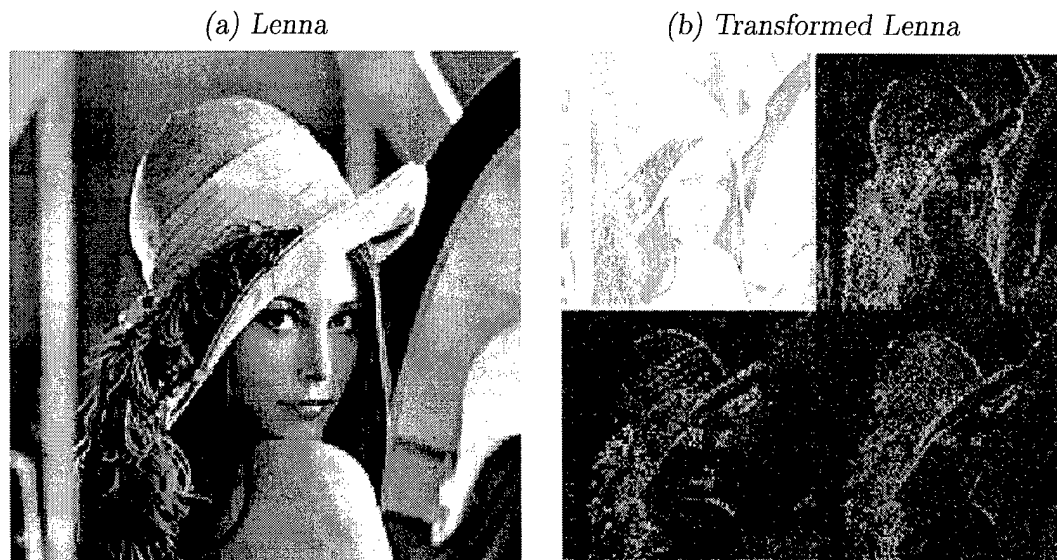


Figure 2.6. *Original Lenna image and one iteration of the discrete wavelet transform of Lenna. The discrete wavelet transform is a dyadic decomposition. Each of the four subbands contains different characteristics of the original image.*

We refer to the top left subband as the Low-Low (LL) subband, where LL refers to the type of image characteristics the subband preserves. The convention is that first letter of the subband refers to processing performed along the rows and the second letter refers to processing performed along the columns. Thus, for the LL subband, we first lowpass filter the rows and then lowpass filter the columns. The result is a coarse approximation of our signal since lowpass filtering both the rows and columns blurs the edges (high frequency components) of the original image in both directions.

As we examine the type of information provided by the other subbands, we see that the discrete wavelet transform is highly directional. The bottom right subband is the highpass or High-High (HH) subband. It preserves edges oriented at  $45^\circ$  (the "cross-hatch") because both the rows and columns are highpass filtered. The top right subband is known as the Low-High (LH) subband. It blurs horizontal lines because the rows are lowpass filtered, but it preserves vertical lines because the columns are highpass filtered. The lower left subband is known as the High-Low (HL) subband. It preserves horizontal lines because the rows are highpass filtered, but it blurs vertical lines because the columns are lowpass filtered. The LH and HL subbands are collectively known as the bandpass subbands; the highpass and bandpass subbands together are referred to as the detail subbands. Thus, the discrete wavelet transform of an image is a coarse approximation (LL subband) of the original image and a series of details (LH, HL, and HH subbands).

The frequency responses of the filters used to form the subbands are given in Figure 2.7. Although the Daubechies (7,9) wavelet was used to create the frequency response plots, the general shape of the plots is the same for all wavelets. The white areas indicate where frequency components of the image are preserved. The black areas represent which frequency components of the image are attenuated. For example, we see in Figure 2.7 b that the filters that create the LH subband preserve the high frequency components of the columns, but do not pass high frequency components of the rows. Hence, the LH subband preserves vertical edges in an image. Similar observations may be made from the other frequency response plots, which clearly illustrate why the subbands of the discrete wavelet transform are highly directional.

Lower scales are formed by iterating on the coarse approximation (LL subband) of the next highest scale. Each new scale is given by a coarse approximation of the LL subband of the next highest scale and a series of details. Figure 2.8 illustrates the multiscale effect of the discrete wavelet transform.

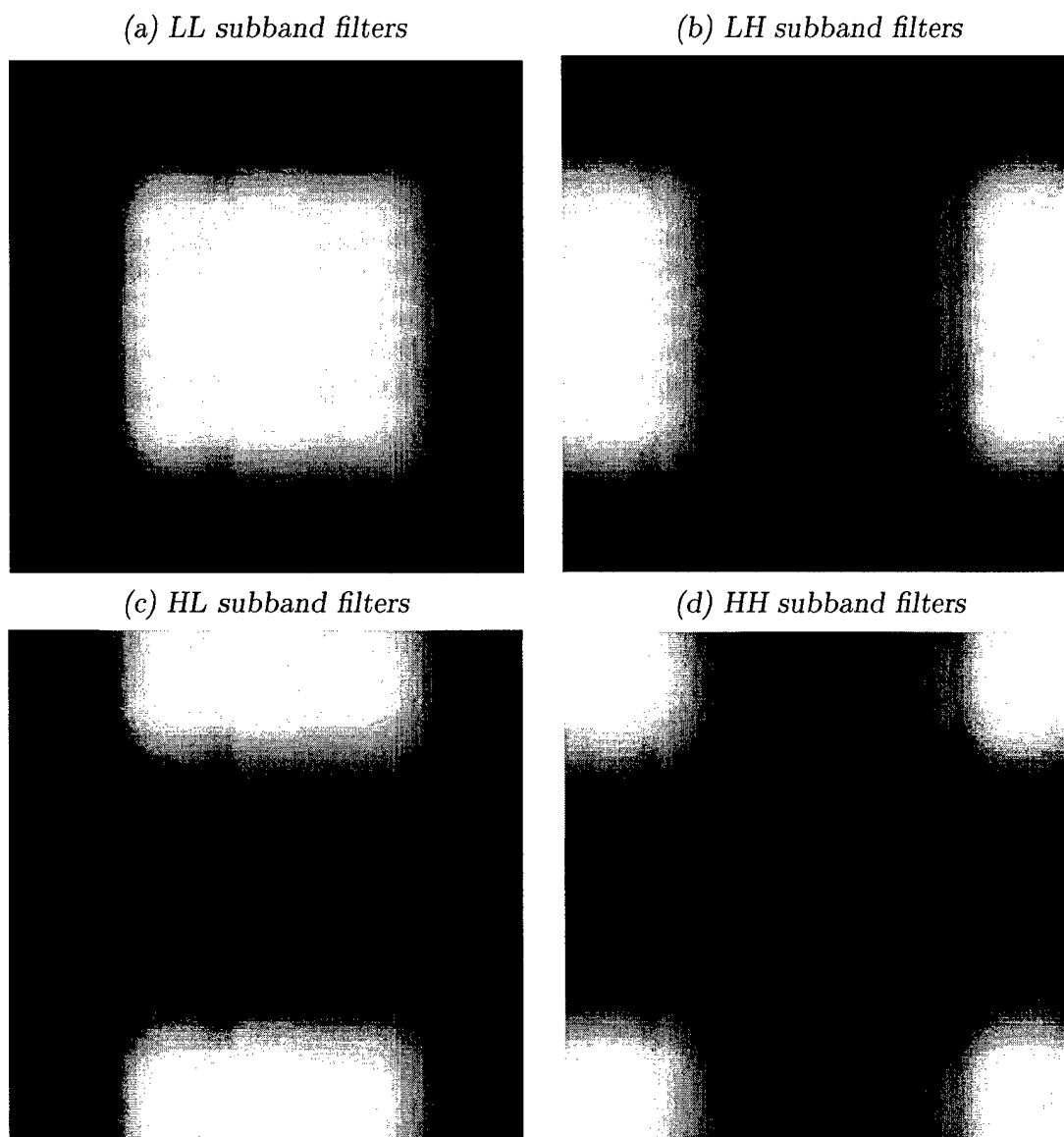


Figure 2.7. Frequency responses for the filters used to create the (a) LL subband, (b) LH subband, (c) HL subband, and (d) HH subband of the discrete wavelet transform. In each plot, the origin (DC value) is at the center. The x-axis corresponds to the spatial frequencies along the rows and the y-axis corresponds to the spatial frequencies along the columns of the original image. White areas pass frequencies and black areas attenuate. Although the lowpass and highpass filters of the Daubechies (7,9) wavelet were used for these plots, the general shape of the frequency responses of other wavelets look similar.

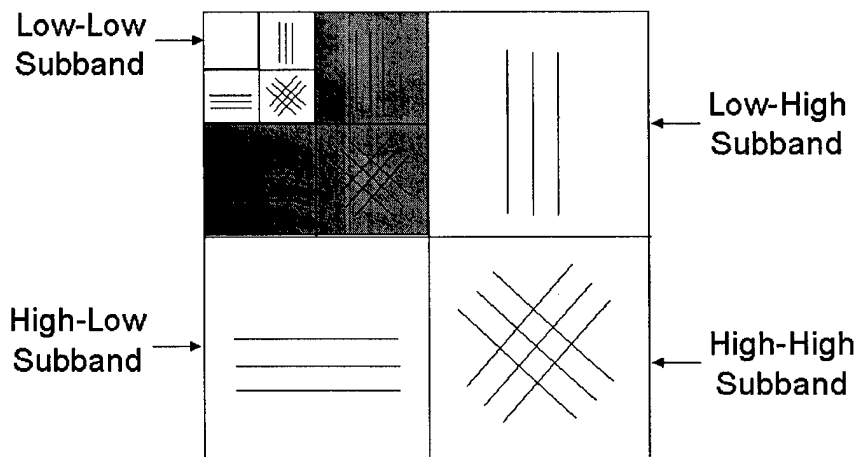


Figure 2.8. *Three iterations of the discrete wavelet transform. The second scale is shaded in gray to distinguish it from the other two. The signal characteristics preserved by each subband are shown.*

*2.3.6 Properties of Wavelets.* Now that we have constructed the discrete wavelet transform, we mention three primary and two secondary properties that make it attractive for image registration. A detailed explanation of the application of these properties to the registration problem is given in Chapter 3. Primary properties are the following (22):

1. Locality - wavelet coefficients are localized simultaneously in space and spatial frequency
2. Parsimony - wavelet coefficients of real world signals (images) tend to be sparse
3. Multiresolution - wavelet coefficients are shifted and dilated into a nested set of scales

Secondary properties of importance are the following (20, 21):

1. Clustering - given a large/small wavelet coefficient, adjacent wavelet coefficients tend to also be large/small
2. Persistence - large/small values of wavelet coefficients tend to propagate across scales

Having presented an overview of filter banks and wavelets, we now discuss their application in image registration.

## *2.4 Image Registration*

Image registration is the process which determines the best match of two or more images acquired at the same or different times by identical or different sensors. The set of input images are matched relative to a set of reference images. Registration is a necessary first step in image processing when comparing the similarities and differences between multiple images.

*2.4.1 Fundamentals of Image Registration.* In (2), Brown describes the four basic components of image registration:

(a) The feature space which contains the characteristics extracted from the reference and input images that are used for matching.

(b) The search space which contains the class of potential transformations that establish the correspondence between the reference and input images. The most common transformations used are rigid (scaling, translation, and rotation of the input image relative to the reference image), affine (a rigid transformation plus a shear and aspect ratio change), and polynomial (reference image pixels are transformed according to a polynomial function).

(c) The search strategy which chooses the transformations to be used. Examples include local versus global search, optimization techniques, and the multiresolution search.

(d) The similarity metric which measures how well the match is between the reference and input images for the selected search space. Correlation between the reference and input images is most commonly used, but other metrics may be utilized.

Of the four basic components, the feature space is most important. Proper feature selection is critical for accurate registration. According to Li and Zhou (18), features to be used in registration should be:

1. Present in the same position in the image (consistency);
2. Located in high contrast regions;
3. Distributed proportionally across the image;
4. Unique in their surrounding areas.

Manjunath *et al* discuss the importance of feature detection in terms of generality and robustness (23). Feature detection must be general in the sense that it can be applied to many different registration problems. To achieve robustness, the same feature locations must be consistently identified regardless of translation, rotation, and minor scaling and deformation between the input and reference images. It is for this reason we use redundant wavelet transforms.

There are two main types of feature detectors, image-based and edge-based. Image-based detectors select points based on image grayscale intensity, which is not consistent in imagery acquired from multiple sensors. Edge-based methods perform better than image-based techniques because they do not rely on intensity. Rather, edge-based methods look for certain levels of energy, not the amount of energy present. It is the differences between contrasting regions within images that the human eye detects more readily, not absolute intensity values. Thus, edge-based methods are preferred since they appeal more to the human visual system. In Chapter 3, we examine the role wavelets play in our edge-based feature detection scheme.

Having presented the key aspects of image registration, we now discuss existing methods and their limitations to provide a baseline for comparison to our algorithm.

*2.4.2 Fourier Registration Techniques.* Fourier techniques are the precursor to modern wavelet registration techniques. Fourier methods have been used for

many years in signal representation and analysis. Images may be represented as the summation of a series of sinusoids (Fourier series) in the spatial domain or transformed into their frequency components using the Discrete Space Fourier Transform, which is defined as

$$X(\omega_1, \omega_2) = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} x(n_1, n_2) e^{-j\omega_1 n_1} e^{-j\omega_2 n_2}$$

where  $x(n_1, n_2)$  is a spatial domain image (19).

Brown states that nearly all Fourier image registration techniques utilize phase correlation, which results from the shift theorem (translation property) of the Fourier Transform (2). To account for rotation in images, phase correlation is still applied except that we first convert our images from rectangular coordinates to polar coordinates. See (1, 4, 12, 24) for specific implementations of Fourier registration using phase correlation.

*2.4.3 Limitations of Fourier Registration Techniques.* A significant limitation utilizing Fourier techniques to register images is that image decomposition is localized in frequency only. Having resolution only in the frequency domain makes it impossible to determine where the edges (high frequency components) of the image are spatially located. Knowledge of the location of edges is critical because they produce the best features to use in matching. The Short Time (Windowed) Fourier Transform (STFT) partially alleviates this problem. By performing the Fourier Transform on a fixed window (usually Gaussian) of the image, it is possible to obtain some localization in space and frequency that is not possible with the Fourier Transform. Because the STFT is based on a fixed window, it does not provide enough spatial information for high frequencies. This deficiency is overcome using wavelet techniques.

2.4.4 *Wavelet Registration Techniques.* Previous work has shown the promise of wavelet image registration. We explain a few of these methods that directly influenced the development of our registration algorithm.

The largest positive wavelet coefficients obtained from an orthogonal decomposition using Daubechies filters have been used effectively (13, 14, 15; 16). When using cross correlation as a feature matching technique, features larger than twice the size of the wavelet filters were correctly registered using the bandpass subbands of the Daubechies wavelet decomposition (28).

Le Moigne and Zavorin compare the robustness of Daubechies filters and Simoncelli steerable filters to translation, rotation, and noise parameters over large ranges (17). They propose a three step approach for the registration of remotely-sensed imagery. First, the wavelet transform is performed on the reference and input images. Next, domain independent features are extracted at each scale of the wavelet decomposition. Finally, correlations are performed at each scale between the wavelet domain images to find the best rigid transformation. This iterative process improves accuracy when going from coarse to fine resolution of the wavelet coefficients. Additionally, Simoncelli steerable filters are shown to perform better than the orthogonal Daubechies filters because of the invariance of Simoncelli steerable filters. Invariance is achieved because Simoncelli's steerable pyramid decomposition is a redundant representation of the wavelet transform (27).

Djamjdi and Corvi also use the largest positive wavelet coefficients as features (7, 9). Unlike Le Moigne and Corvi, Djamjdi utilizes an *à-trous*, non-pyramidal, wavelet decomposition scheme which is computationally slower than the pyramidal wavelet decomposition. However, both Djamjdi and Corvi match features locally one by one whereas Le Moigne computes a global transformation over the whole image. A unique aspect of Corvi's method is that it employs a clustering technique to provide a significantly more reliable initial estimate for strongly translated and/or rotated images. Also, rather than exclusively using the largest positive wavelet coef-

ficients as features for matching, it leverages both the largest positive and negative wavelet coefficients. One final difference is that Djamjdi computes a polynomial transformation versus a rigid transformation computed by Le Moigne.

A fast method to register images using the wavelet modulus maxima, which is defined as the highpass subband of any scale in a wavelet decomposition, is presented by Sharman (26). Although Sharman utilizes the largest magnitude wavelet coefficients as features like the previously described methods, his technique differs in that the bandpass subbands of the wavelet decomposition are disregarded. Rather, the highpass subband is exclusively utilized because his method utilizes edge-based feature detection and significant edge information is contained in the highpass subband. Using correlation as the similarity metric, this point-to-point feature matching method has been shown to be near perfect for rigid transformations in the presence of little noise.

Sharman's technique is meant for quick registration of images that are not strongly translated, rotated, or scaled. Zheng and Chellappa present a more robust image registration algorithm that estimates translation, rotation, and scale using a point-to-point edge-based feature extraction method (34). Rotation is estimated first, then an initial estimate of translation and scale are computed from a small number of features extracted using the Gabor wavelet model for detecting local curvature discontinuity. Finally, hierarchical feature matching is used to compute translation, rotation, and scale precisely. This technique has been shown to accurately register images in the presence of large translation and rotation and when the images lack significant features.

*2.4.5 Challenges Using Wavelet Techniques.* Wavelets show great promise for use in registration algorithms, but there are issues that must be addressed. The following are the three most crucial challenges to overcome: (1) lack of invariance of

the discrete wavelet transform, (2) choice of wavelet (dependent on the image), and (3) the use of a single global transformation may not be acceptable.

In terms of our basic requirements for an image registration algorithm as described in Section 2.4.1, wavelets certainly pass the generality requirement as they may be applied to many different images. They do not, however, meet the robustness specification because they are sensitive to translation and rotation. Translation or rotation-invariance occurs when the wavelet transform commutes with the translation or rotation operator, respectively. Simoncelli *et al* state that translation or rotation invariance cannot be expected in a system based on convolution and subsampling such as the filter bank implementation of the discrete wavelet transform (27). The critical sampling condition of the discrete wavelet transform must be eliminated to achieve the desired translation or rotation invariance for image registration. Invariance comes at the costs of slower computation time and a more memory intensive process.

Proper choice of the wavelet function is another issue to consider when utilizing wavelet transforms for image registration. Sharman *et al* note that the energy content of a given subband is different depending on the choice of the wavelet basis functions (25). For example, an image dominated by squares and rectangles would require a wavelet basis that best approximates images dominated by arcs and circles. More wavelet coefficients will be produced since the wavelet basis will not approximate the image well. Thus, registration accuracy is improved because the number of features that may be used for registration increases.

Finally, the problems of using a single global transformation to map an image onto a common standard must be considered. Unser *et al* state that the two principal difficulties encountered when using a single global transformation are (1) a proper deformation model is rarely known for the reference image, and (2) analysis methods lack a strong statistical background because pixels in the spatial domain are correlated in some manner and not statistically independent as desired (31).

## *2.5 Summary*

The basics of filter banks, which are used in the implementation of the discrete wavelet transform, were presented. The wavelet recursion equations, which all orthogonal wavelets obey, were derived. From these equations, the discrete wavelet transform was constructed from integer shifts and dilations of the wavelet basis functions. Biorthogonal wavelets do not obey the wavelet recursion equations like orthogonal wavelets, but are useful because they could be constructed from large, linear phase filters. Next, we discussed the two-dimensional discrete wavelet transform, which is a separable transform since it is a product of two one-dimensional transforms; we first process the rows and then the columns. We concluded our wavelet discussion by stating some useful properties of wavelets that make them attractive for registration. Finally, we presented some Fourier and wavelet registration techniques and their limitations to serve as a baseline for comparison to our registration algorithm, which we describe in the next chapter.

### *III. Methodology*

#### *3.1 Overview*

We require our image registration algorithm to be automatic, general, robust, and accurate. First, a computer must be able to perform the registration of images with little or no human interaction (automatic). Second, our algorithm must be general enough to handle many different kinds of images. Third, it must be robust enough to handle strongly translated or rotated images in the presence of additive white Gaussian noise. Finally, accuracy is still the primary concern since image registration serves as a crucial first step in image processing of large data sets.

We discussed Fourier and wavelet techniques and their limitations in Chapter 2. Our algorithm improves on some of the key ideas from each including, but not limited to, the following:

1. Matching the largest magnitude wavelet coefficients of the bandpass subbands;
2. Performing a global rigid transformation (a translation or rotation) of the data;
3. Generating initial estimates of translation or rotation and then refining;
4. Converting from rectangular to polar coordinates to more accurately estimate rotation.

We further expand on these ideas by introducing the new concepts of masking and polar wavelets. Before describing our registration algorithm, we must discuss why wavelets are well suited to tackle the image registration problem.

#### *3.2 The Wavelet Registration Concept*

Proper feature selection is critical for accurate registration. Recall from Section 2.4.1, we desire to select features per the criteria of Li and Zhou:

1. Present in the same position in the image (consistency);

2. Located in high contrast regions;
3. Distributed proportionally across the image;
4. Unique in their surrounding areas.

Recall from Section 2.3.6, we discussed primary and secondary properties of the discrete wavelet transform which make it especially useful for image registration. We now examine how these properties satisfy the above criteria so we may produce a general and robust registration algorithm.

Consistency of features is essential for registration success. We need to be able to select a feature in the input image and know that it will be in the same position in our reference image. The parsimony of the wavelet coefficients guarantees that there are a limited number of significant coefficients to use as features for matching. Since we have a limited number of features for each image, we can reasonably expect to select the same ones each time even if noise significantly degrades our images.

Edges in images typically serve as a dividing line between high contrast regions. Because of their localization in space and spatial frequency, wavelets are natural edge detectors. Together, locality and clustering of wavelet coefficients help to select groups of features located in high contrast regions.

Proportional distribution of features helps achieve a robust algorithm that can better handle the adverse effects of noise. Parsimony and clustering ensure that we have groups of significant features in our images. Locality ensures that these clusters of significant coefficients are spread throughout each image since edges are typically distributed throughout images.

Finally, we desire unique features for matching. Again, parsimony and clustering ensure that we have clusters of significant coefficients to choose. We can expect to select the same limited number of coefficients every time when using wavelets because of locality. It is the uniqueness of these features that makes wavelet registration more accurate than spatial registration.

For the last three criteria, redundant wavelet transforms and the discrete wavelet transform perform about the same. The real power of a redundant wavelet transform over the discrete wavelet transform lies in its ability to consistently extract the same features. This is a direct result of how each transform is created. The redundant wavelet transform is the product of an undecimated filter bank that contains no time-varying components. The discrete wavelet transform, however, is critically sampled using time-varying decimators which destroy the consistency we desire.

Additionally, orthogonal wavelet transforms compress signal information (parsimony) while keeping noise scattered and white. Biorthogonal wavelets also compress signal information, but they do not keep noise completely white. The noise becomes correlated with the signal information, but not nearly enough to negatively impact our feature matching. Redundant wavelet transforms can never be orthogonal, but like biorthogonal wavelets, they exhibit the ability to scatter noise enough for registration purposes.

From this point forward in this thesis, the term discrete wavelet transform refers to the critically sampled discrete wavelet transform. The term redundant wavelet transform refers to both the shift-invariant wavelet transform and the rotation-invariant polar wavelet transform. Note, we are not implying that these redundant wavelet transforms are continuous; they are not. Rather, we are simply distinguishing between redundant and critically sampled wavelet transforms.

### *3.3 Our Image Registration Algorithm*

We now present our image registration algorithm in detail, beginning with translation estimation and concluding with rotation estimation. In both cases, we provide a top level description of the algorithm first and then give detailed explanations of why we chose to build our algorithm in the way we did.

*3.3.1 Algorithm Flow for Translation Estimation.* When estimating translation, the flow of our registration algorithm is as follows:

1. User specifies the reference and input images, the number of significant coefficients to keep ( $N$ ), and that translation estimation is required.
2. Perform one iteration of the shift-invariant wavelet transform on each image.
3. Mask the LH subbands of the wavelet transformed reference and input images, keeping the largest  $N$  wavelet coefficients (both positive and negative values). For the rest of this thesis, we refer to these coefficients as the significant coefficients.
4. Compare the location of the largest significant coefficient in the masked reference subband to the locations of all  $N$  significant coefficients in the masked input subband. This produces  $N$  initial estimates of translation. Repeat this procedure for the middle and tenth largest significant coefficients in the masked reference subband so that we have  $3N$  initial estimates of translation.
5. Perform correlations only for the initial estimates of translation (a total of  $3N$  correlations). Do this by circularly shifting the coefficients of the masked reference subband according to the translation estimate and then comparing to the masked input subband. Keep the highest valued correlation and consider it the best estimate of translation produced by the LH subbands.
6. Beginning with the masking process, perform all the above steps again using the HL subband of the reference and input images.
7. Perform a Pearson correlation in the spatial domain for the translation estimate from the LH subband. Do the same for the translation estimate from the HL subband. The correlation producing the highest value is selected as the final estimate for translation.
8. Translate the input image to the reference image based on our final estimate to produce the registered image.

Having presented the general flow of the translation portion of our registration algorithm, we now discuss each of its important components.

*3.3.2 Shift-Invariant Wavelet Transform.* Shift-invariance is crucial in registering strongly translated images, especially in the presence of noise. It is vital to achieve the consistency of features in images as described in Section 3.2. Figure 3.1 illustrates what we mean by shift-invariance. We start with the original data sequence and the same sequence shifted one unit to the right. Next, we take the discrete wavelet transform of both sequences. If the discrete wavelet transform was shift-invariant, the transform of the shifted sequence would be the transform of the original sequence shifted one unit to the right. Clearly it is not and we see that the discrete wavelet transform is not shift-invariant.

We use a redundant wavelet transform to provide the shift-invariance we desire. The shift-invariant wavelet transform differs from the discrete wavelet transform in that we account for all possible shifts of an image, not just every other shift in the row and column directions.

Our shift-invariant wavelet transform is implemented as shown in Figure 3.2 a. By first pulling all the filters through the decimators using the Noble Identities and then removing the time-varying decimators, we achieve shift-invariance at the cost of doubling the image size. Because iterations for scale  $m$  require that we use the Noble Identities  $m$  times, we must expand the filters  $h$  and  $g$  by  $2^m$ .

For our algorithm, we only perform one iteration of the shift-invariant wavelet transform because each subband is the same size as the original image. There are no gains in speed by computing further iterations since the image size remains constant after the first iteration.

*3.3.3 Subband Choice.* The bandpass subbands contain the most useful information for registration purposes. According to Stone *et al*, the highpass subband

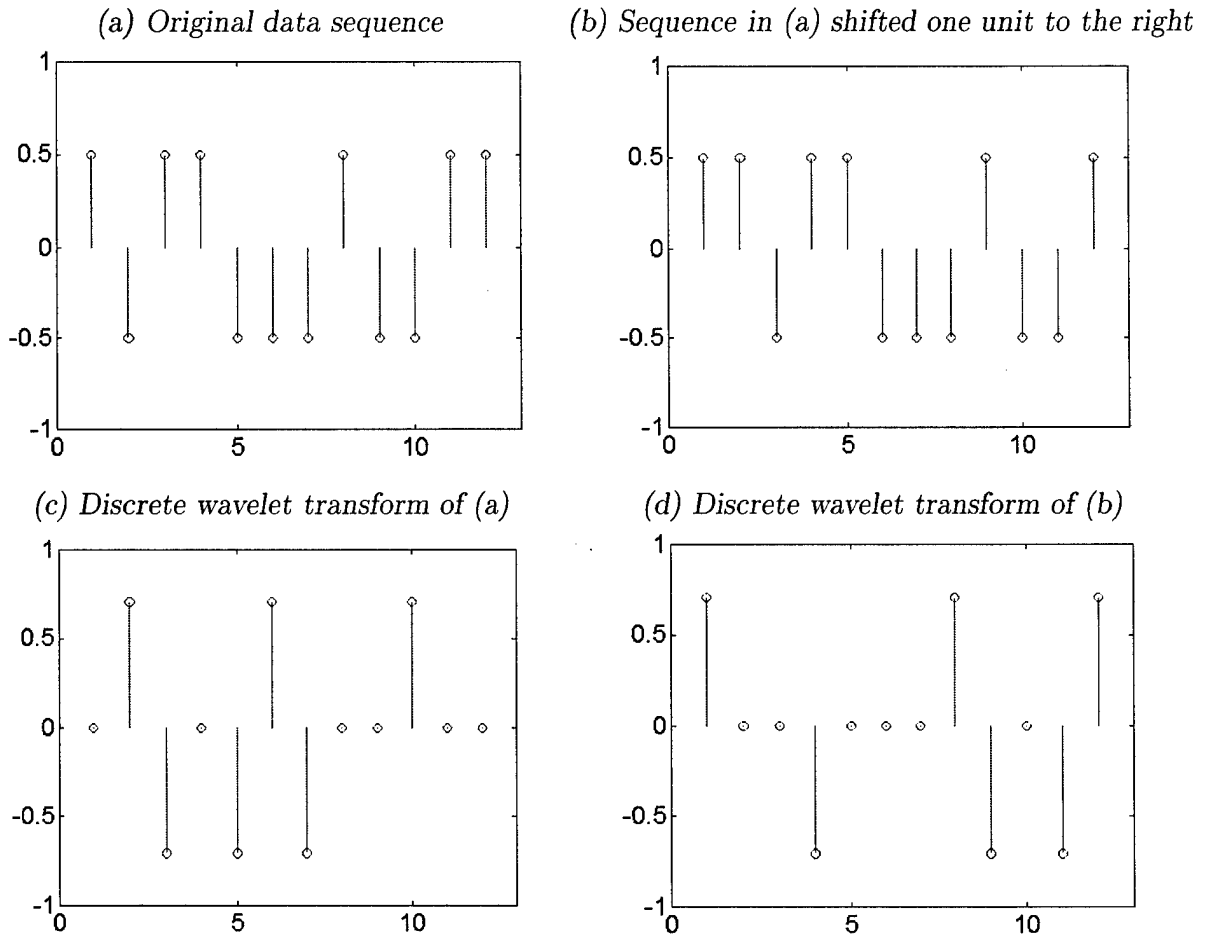
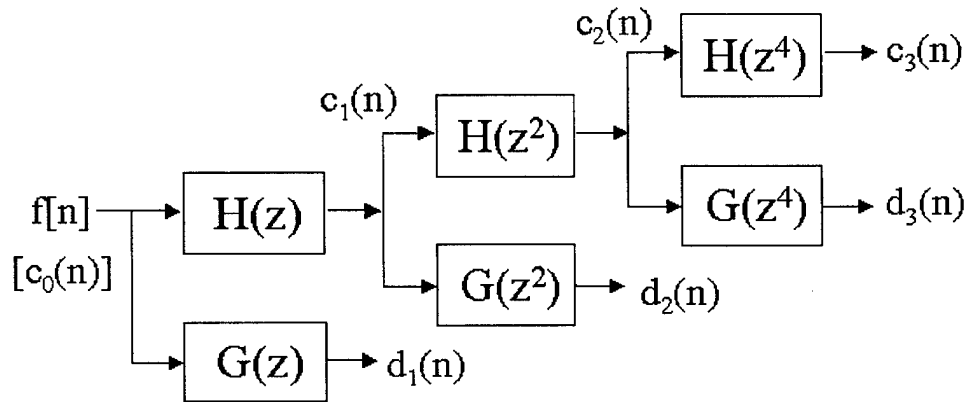


Figure 3.1. *Shift-invariance test. If the discrete wavelet transform was shift-invariant, the sequence in (d) would be the same as (c) shifted to the right by one unit.*

(a) Shift-invariant wavelet transform



(b) Discrete wavelet transform

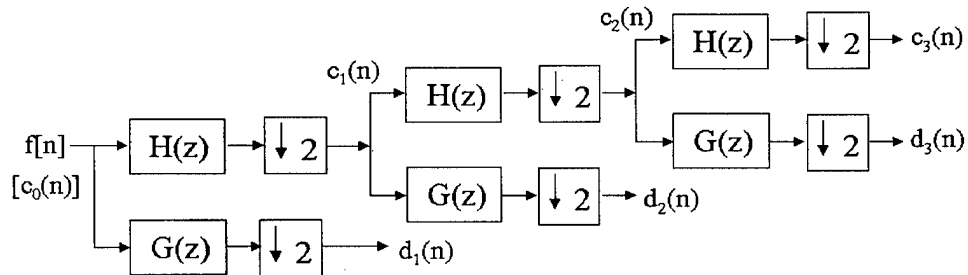


Figure 3.2. Comparison of the shift-invariant wavelet transform and the discrete wavelet transform. Three iterations of each transform are shown. The shift-invariant wavelet transform filters need to be expanded by  $2^m$ , where  $m$  is the scale, since we apply the Noble Identity for Decimators.

is much more sensitive to translation than the bandpass subbands (28). This prevents the highpass subband from preserving the consistent features we require. Also, the highpass subband is more adversely affected by noise since it preserves high frequency components and noise is essentially comprised of high frequency components. The bandpass subbands, however, suppress some of the negative effects of noise because each is lowpass filtered. They also prove useful because of the image characteristics they preserve. The LH subband performs exceptionally well for images dominated by vertical edges and the HL subband performs well for images where horizontal edges are prevalent.

*3.3.4 Masking.* The process of masking the bandpass subbands of the transformed reference and input images is unique to our algorithm. By assigning a value of “1” to the top  $N$  significant coefficients and a “0” to all other values, we weight all significant coefficients equally. This minimizes the impact of noise that may be present in the reference or input images. Masking also allows us to perform correlation using a logical comparison, which is much quicker than the more computationally intensive method of multiplying each matrix element point by point and summing over each index. The small cost of masking is that we mitigate potential positive effects of a high correlation between particularly large wavelet coefficients that may be present in the reference and input images. The masking process is powerful as it provides highly desired noise suppression and reduces computational complexity while still leveraging the advantage of using the highest magnitude wavelet coefficients. See Table 3.1 for an example of the application of the mask.

*3.3.5 Initial Estimation.* After masking the LH subband of the reference and input images, we are left with two matrices of “1’s” (the significant coefficients) and “0’s.” We select the location of the largest significant coefficient from the masked reference subband and compare it to the locations of all  $N$  significant coefficients in the masked input subband to obtain  $N$  initial estimates of translation. If we are

(a) *Original matrix*                      (b) *Masked matrix*

$\begin{bmatrix} 17 & 24 & 1 & 8 & 15 \\ 23 & 5 & 7 & 14 & 16 \\ 4 & 6 & 13 & 20 & 22 \\ 10 & 12 & 19 & 21 & 3 \\ 11 & 18 & 25 & 2 & 9 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$
---	---

Table 3.1. *Masking process. We create the mask by replacing the top  $N$  significant coefficients with a “1” and assigning a “0” to all others. For this case,  $N = 10$ .*

extracting the same features (i.e., using the same coefficient locations) from each masked subband and noise has not adversely affected us, then one of our  $N$  estimates should be the true translation. Thus, we only need to perform  $N$  correlations rather than  $m * n$  correlations for a  $m$ -by- $n$  image. For the case where little or no noise is present, this works well. However, to account for the situation where there is moderate or heavy noise, we need to select more significant coefficients from the masked reference subband to compare to all the significant coefficients of the masked input subband. Thus, we obtain more initial estimates. This reduces the probability that we select a significant coefficient more heavily influenced by noise than the others. For our algorithm, we select three significant coefficients from the masked reference subband: the largest, middle, and tenth significant coefficients. We select these three significant coefficients because we expect them to produce features that are distributed proportionally throughout the image as required by the criteria described in Section 3.2. At this point in our algorithm, we have a total of  $3N$  initial estimates of translation for the LH subband ( $N$  estimates for each of the three significant coefficients we selected). Next, we circularly shift the coefficients of the masked reference subband according to each of the  $3N$  initial translation estimates and correlate with the masked input subband. We perform a fast correlation by using a logical AND operation and then summing over all indices. We only care about the number of “hits” we receive since we have masked our true coefficient values. Thus,

the translated masked reference subband that correlates the best (highest number of “hits” after summing) with the masked input subband is declared the translation estimate. We perform the same procedure for the HL subband of the reference and input images.

*3.3.6 Final Estimation.* In this phase of the algorithm, we refine our initial estimates to achieve the final translation estimate. After initial estimation, we have a single estimate of translation for both the LH and HL subbands. We apply each translation estimate to the input image in the spatial domain and perform a Pearson correlation with the reference image. We keep the translation estimate which produces the highest correlation (value is between 0 and 1, with 1 representing perfect correlation and 0 representing no correlation). The Pearson correlation is defined in (30) as

$$r = \sum_{i=1}^N \frac{(x_i - \bar{x})(y_i - \bar{y})}{(N - 1)S_x S_y}$$

where  $N$  is the number of samples,  $x_i$  is the reference image data,  $y_i$  is the input image data,  $\bar{x}$  and  $\bar{y}$  are the means, and  $S_x$  and  $S_y$  are the standard deviations of the reference and input images, respectively. Having described translation estimation, we now describe rotation estimation.

*3.3.7 Algorithm Flow for Rotation Estimation.* When estimating rotation, the flow of the registration algorithm is as follows:

1. User specifies the reference and input images, the number of significant coefficients to keep ( $N$ ), and that rotation estimation is required.
2. Transform the reference and input images from rectangular coordinates to polar coordinates (the origin is the center of the image).
3. Perform one iteration of the shift-invariant wavelet transform on each image. The result is a rotation-invariant polar wavelet transform.

4. Mask the LH subband of the transformed reference and input images, keeping the largest  $N$  wavelet coefficients (both positive and negative values).
5. Compare the location of the largest significant coefficient in the masked reference subband to the locations of all  $N$  significant coefficients in the masked input subband. This produces  $N$  initial estimates of rotation. Repeat this procedure for the middle and tenth largest significant coefficients in the masked reference subband so that we have  $3N$  initial estimates of rotation.
6. Perform correlations only for the initial estimates of rotation (a total of  $3N$  correlations). Do this by circularly shifting the coefficients of the masked reference subband according to the rotation estimate and then comparing to the masked input subband. Keep the highest valued correlation and consider it the best estimate of rotation produced by the LH subbands.
7. Beginning with the masking process, perform all the above steps again using the HL subband of the reference and input images.
8. Perform a Pearson correlation in the spatial domain for the rotation estimate from the LH subband. Do the same for the rotation estimate from the HL subband. The correlation producing the highest value is selected as the final estimate for rotation.
9. Rotate the input image to the reference image based on our final estimate to produce our registered image.

The process of estimating rotation is exactly the same as the process for estimating translation. The only exception is that we first convert our reference and input images from rectangular coordinates to polar coordinates before processing. Thus, we only discuss the differences between translation and rotation estimation that result from this coordinate transformation.

*3.3.8 Polar Redundant Wavelet Transform.* The shift-invariant wavelet transform does not provide rotation-invariance because of the directionality of the

detail subbands. Rotations do not correspond to shifts of horizontal or vertical edges in a right-handed rectangular coordinate system. Complex wavelets (11) and steerable pyramids (27) were developed to provide greater directionality, but they still lack the necessary rotation-invariance needed to properly estimate rotation. Rather than attempt to calculate a marginal estimate for rotation based on the shift-invariant wavelet transform, we instead opt for a purely polar coordinate based solution to provide rotation-invariance.

The standard conversion from rectangular coordinates to polar coordinates is given by

$$r = \sqrt{x^2 + y^2}$$

and

$$\theta = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } x > 0 \\ \pi + \arctan\left(\frac{y}{x}\right) & \text{if } x < 0 \end{cases}$$

where  $(x, y)$  are rectangular coordinates and  $(r, \theta)$  are the corresponding polar coordinates (10).

First, we transform our image from a matrix of rectangular coordinate system values of the form  $(x, y)$  to a matrix of polar coordinate system values of the form  $(r, \theta)$ . This transformation is performed only if the reference and input images were not collected in a polar format. Next, we perform one iteration of the shift-invariant wavelet transform as described in Section 3.3.2. Linear shifts of the columns now correspond to rotations; thus, our shift-invariant wavelet transform is now rotation-invariant. We refer to this new transform as the rotation-invariant polar wavelet transform. See Figure 3.3 for an example of rectangular and polar plots for the cameraman image.

Unlike the standard wavelet transform, it is difficult to display the subbands on a single polar plot. Thus, we must provide a separate polar plot for each subband

(a) *Rectangular coordinates*



(b) *Polar coordinates*



Figure 3.3. *Cameraman image sampled on a rectangular grid and on a polar grid.*

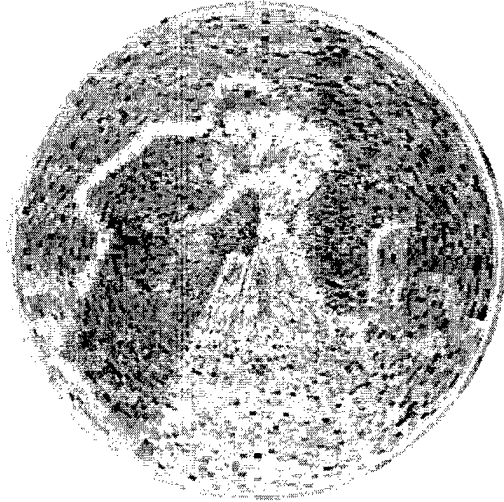
of the decomposition as shown in Figure 3.4. When we perform the shift-invariant wavelet transform, the information contained in each subband is highly directional: the LH subband preserves vertical edges, the HL subband retains horizontal edges, and the HH subband keeps the cross-hatch as described in Section 2.3.5. In the polar environment, these subbands preserve different image characteristics since rows and columns contain data of constant radius and constant angle, respectively.

The polar LL subband remains our coarse approximation since it is a blurred, smoother version of our original polar image created by lowpass filtering along the radii and angles. Since the polar LH subband is now formed by lowpass filtering the radii and highpass filtering the angles, it preserves edges as they are rotated around angles and smoothes edges along the radii. Similarly, the polar HL subband keeps information about edges along the radii, but blurs edges rotated around angles. The polar HH subband preserves a combination of edges that persist along radii and around angles. Thus, we again choose to use the bandpass subbands for our algorithm because they provide the necessary noise tolerance (they are lowpass filtered along either the radii or angles) and the required directionality for extracting polar features.

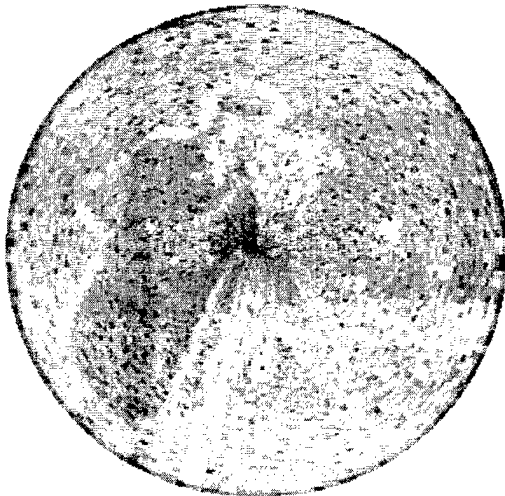
(a) Polar LL subband



(b) Polar LH subband



(c) Polar HL subband



(d) Polar HH subband

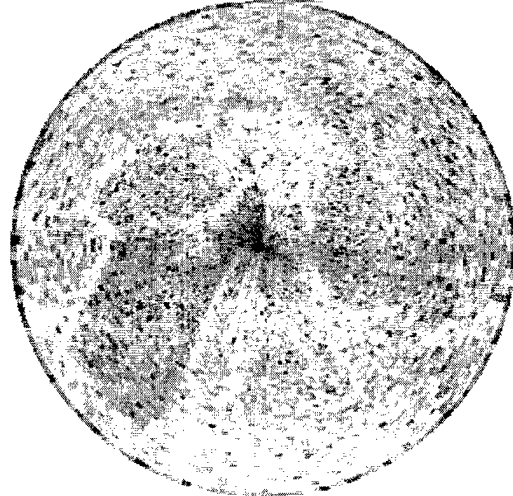


Figure 3.4. *Rotation-invariant wavelet transform (one iteration). It is not feasible to display polar wavelets in the same manner as standard wavelets because of the polar sampling grid. A separate plot is required for each subband.*

The wavelet recursion equations of Section 2.3.2 do not hold for our implementation of the polar redundant wavelet transform. These equations were derived for uniform integer shifts and dilations of the wavelet basis functions. Clearly, the polar grid is non-uniform and does not allow for this as shown in Figure 3.5. For simplicity we assume the polar data points are all uniformly spaced.

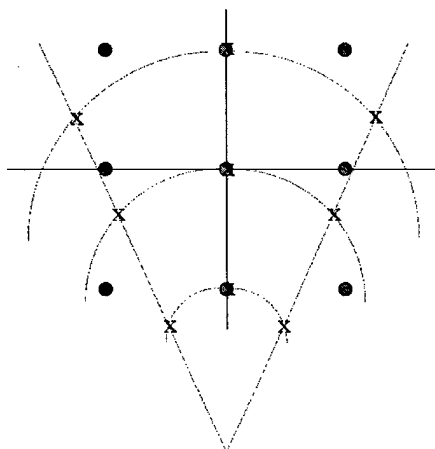


Figure 3.5. *Rectangular versus polar sampling grids. When convolving on the rectangular grid (the dots), all points are uniformly spaced. This is not true for the polar grid (the X's), where points along the radii are uniformly spaced, but not the points around the angles.*

To ensure this simplification is valid, we test our polar redundant wavelet transform to see if it is well behaved. First, we perform three iterations of the polar redundant wavelet transform. Next, we keep only the top 5% of the wavelet coefficients (zero all others) and invert. If our reconstructed image is of “good” quality (i.e., recognizable and free of significant artifacts), we know our transform is stable and acceptable to use in our registration algorithm. Figure 3.6 demonstrates that our new transform is well behaved and acceptable to use when estimating rotation.

After creating the polar LH and HL subbands for the reference and input images, we follow the exact same process as we did in translation estimation except that we are now calculating a rotation estimate.

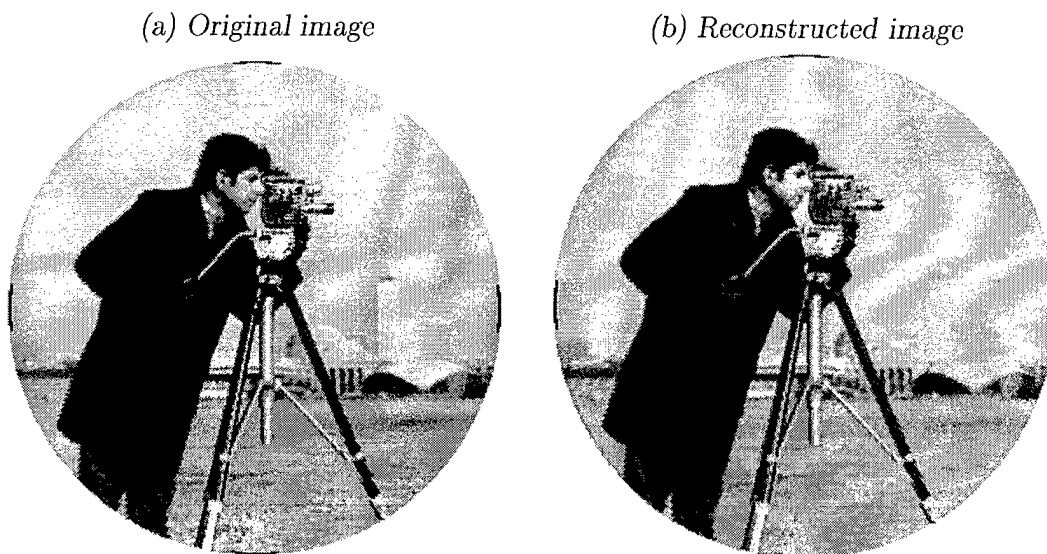


Figure 3.6. *Test for behavior of the polar redundant wavelet transform. We perform three iterations of the polar redundant wavelet transform on our original image, keep only the top 5% of the wavelet coefficients (zero all others), and then invert. The transform is well behaved because the reconstructed image is recognizable and free of significant artifacts.*

### 3.4 Summary

In this chapter, we discussed why the parsimony, locality, and clustering properties make wavelets suitable for feature extraction in registration algorithms. Next, we described the translation and rotation estimation components of our registration algorithm. Redundant wavelet transforms are used to provide the shift-invariance (for translation estimation) and rotation-invariance (for rotation estimation) necessary to extract consistent features for matching. The bandpass subbands are masked to increase computational efficiency and reduce the effects of noise. Initial estimates are refined to a single translation or rotation estimate for each bandpass subband by performing logical correlations between the shifted masked reference and the masked input. A final estimate is determined by performing a Pearson correlation in the spatial domain.

## IV. Results

### 4.1 Introduction

In this chapter, we analyze specific data illustrating the solid performance of our registration algorithm. First, we describe the design of our validation study by explaining the measures of performance, why the Daubechies (7,9) wavelet was selected over other wavelets for use in testing, and exactly how the data was collected. Next, we discuss the results obtained for translations first, and then rotations. We select the best choice for the number of significant coefficients,  $N$ , based on the simulation data. Finally, we assess the robustness of the algorithm by determining the error resulting when the origin of the input image is misaligned with respect to the reference image.

### 4.2 Design of the Algorithm Validation Study

The objective of our study is to validate the effectiveness of our registration algorithm in registering an input image relative to a reference image. To do this, we first determine if a redundant or critically sampled wavelet transform provides better registration accuracy since accuracy is most important objective. Next, we compare the performance of the LH, HL, and HH subbands to verify that the bandpass subbands are a sound choice. Finally, we empirically determine the best choice for the number of significant coefficients ( $N$ ) to use in our algorithm.

*4.2.1 Test Images.* We select the 256-by-256 8-bit grayscale Lenna (Figure 4.1 a) and cameraman (Figure 4.1 b) images on which to conduct our validation experiments. Besides that fact that these images are commonly used in image processing, they are suitable because each contains one significant object in the foreground to be registered and each has a balance of high frequency components (edges) and low frequency components (regions of similar pixel values) to test the flexibility

of our algorithm. For Lenna, her hat and the feather are the dominant high frequency components; the camera and the cameraman himself are the dominant high frequency components for that image.

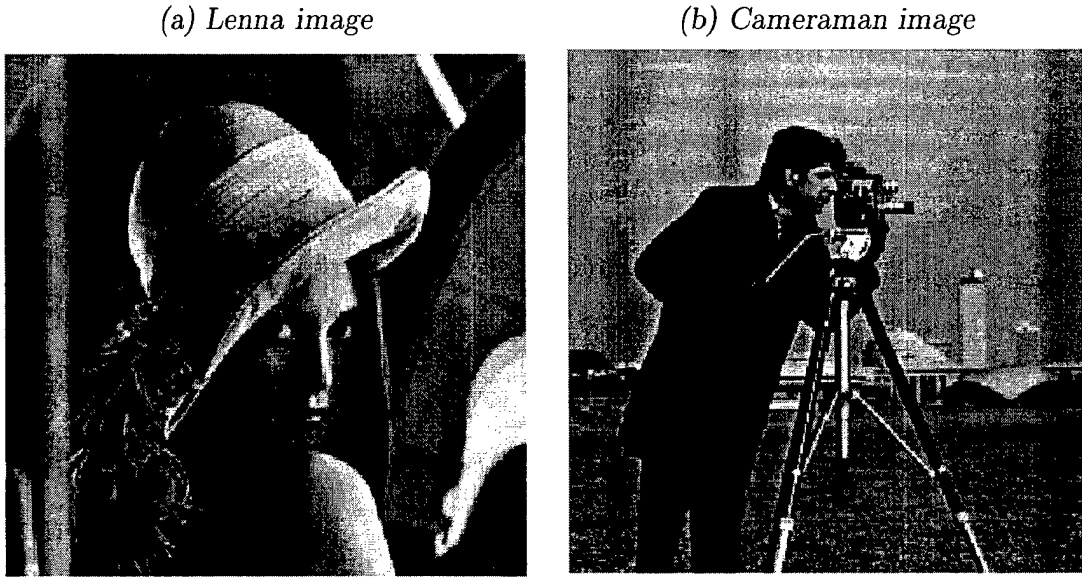


Figure 4.1. *Lenna and cameraman images. These images are used in validation testing of our algorithm.*

**4.2.2 Peak Signal-to-Noise Ratio.** Peak Signal-to-Noise Ratio (*PSNR*) is used in this study as a measure of how much an image is degraded by noise. *PSNR* is defined as

$$PSNR = 20 \log \left( \frac{\max |x_i|}{\sqrt{\sum_i \frac{(x_i - \hat{x}_i)^2}{N}}} \right) \quad (4.1)$$

where  $x_i$  are the pixel grayscale values of the original image,  $\hat{x}_i$  are the pixel grayscale values of the corrupted image, and  $N$  is the number of pixels in the image. A *PSNR* of 30 dB or higher corresponds to an image that is considered to be near perfect to the human eye. For any *PSNR* less than 30 dB, the image appears noisy. Thus, for our experiments, we select *PSNR* values of 30 dB and 22 dB because they represent high image quality (little noise) and poor image quality (significantly de-

graded by noise), respectively. See Figure 4.2 for examples of Lenna and cameraman at  $PSNR$  values of 30 dB and 22 dB.

*4.2.3 Selection of Wavelet.* The biorthogonal Daubechies (7,9) wavelet was chosen for use in our experiments over other wavelets for several reasons. First, the Daubechies (7,9) filters

$$h = [0.0378 \quad -0.0239 \quad -0.1106 \quad 0.3774 \quad 0.8527 \quad 0.3774 \quad -0.1106 \quad -0.0239 \quad 0.0378]$$
$$g = [0.0645 \quad -0.0407 \quad -0.4181 \quad 0.7885 \quad -0.4181 \quad -0.0407 \quad 0.0645]$$

are well behaved. Because the filters are symmetric, they exhibit the highly desirable linear phase property. Next, they possess excellent localization in space and spatial frequency, which results in wavelet coefficients that decay rapidly (parsimony). This is especially important when selecting features for matching as discussed in Section 3.2. See Figure 4.3 for an illustration of the parsimony of the Daubechies (7,9) wavelet. By reconstructing the Lenna image using a limited number of the most significant wavelet coefficients, we see that the Daubechies (7,9) wavelet provides the best  $PSNR$  when compared to the Haar, Daubechies (4), and Daubechies (8) orthogonal wavelets.

*4.2.4 The Validation Study.* All simulations of our registration algorithm were accomplished using MATLAB, version 5.3. For each  $PSNR$  value (30 dB and 22 dB), we ran 100 iterations to achieve robust statistical results. Since determining the best choice of  $N$  was a goal of this study, we let  $N$  range from 10 to 250 in steps of 20. A larger range is more desirable, but we selected this range because it was computationally manageable given the amount of iterations being performed.

The input image for translation testing was created by circularly shifting the reference image 7 pixels in the positive x-direction ( $T_x = 7$ ) and 4 pixels in the positive y-direction ( $T_y = 4$ ) of a right-handed rectangular coordinate system. Similarly,

(a) Lenna image with  $PSNR = 30$  dB



(b) Cameraman image with  $PSNR = 30$  dB



(c) Lenna image with  $PSNR = 22$  dB



(d) Cameraman image with  $PSNR = 22$  dB



Figure 4.2. Examples of different  $PSNR$  values for the Lenna and cameraman images.  $PSNR$  values of 30 dB and 22 dB represent images of high quality (little noise) and poor quality (significantly degraded by noise), respectively.

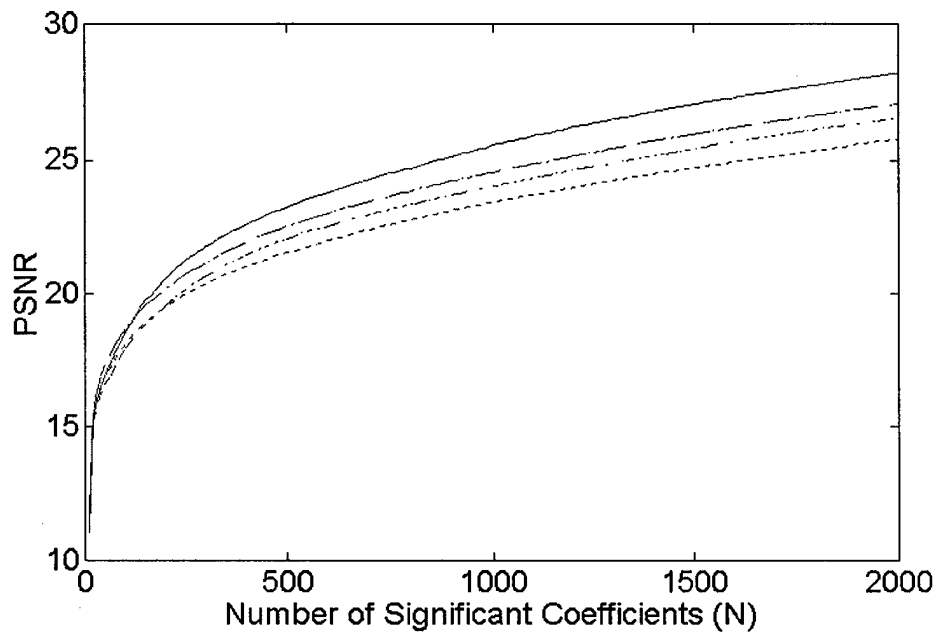


Figure 4.3. *PSNR versus the number of largest magnitude wavelet coefficients ( $N$ ) used in reconstruction for the Daubechies (7,9) (solid line), Daubechies (8) (dashed line), Daubechies (4) (dash dot line), and the Haar (dotted line) orthogonal wavelets. We see that the Daubechies (7,9) wavelet provides the most parsimonious signal representation (ideal for registration) since it provides the best PSNR for a given  $N$ .*

the input image for rotation testing was formed by circularly shifting the columns (which correspond to constant angles because we are in polar coordinates) of the polar reference image 13 pixels ( $R = 13$ ). For our implementation, a shift of 1 pixel corresponds to a shift of  $1^\circ$ . A different realization of additive white Gaussian noise was added to the reference and input images for each iteration in the translation and rotation cases. See Figures 4.4 and 4.5 for examples of translated and rotated versions of Lenna and cameraman, respectively.

When using the discrete wavelet transform with our algorithm, a shift of 1 pixel in the wavelet domain corresponds to a shift of 2 pixels in the spatial domain because of the effects of downsampling. The discrete wavelet transform is only accurate for one of four possible shifts of the images because we downsample both the rows and columns by two. Because our algorithm is precise only to 1 pixel in the translation case and  $1^\circ$  in the rotation case, it is impossible to estimate any odd number translation or rotation when using the discrete wavelet transform since we must double the estimate of our algorithm. Thus, we loosen the criteria of registration accuracy. A correct estimate for translation is  $T_x = 6, 7, \text{ or } 8$  pixels and  $T_y = 3, 4, \text{ or } 5$  pixels and a correct estimate for rotation is  $R = 12, 13, \text{ or } 14^\circ$ .

### 4.3 Translation Performance

We first examine the performance of our algorithm for estimating translation.

*4.3.1 Redundant versus Standard Wavelet Transforms.* Both the redundant and discrete wavelet transforms meet the last three criteria of Li and Zhou as described in Section 3.2. The difference is in the first requirement, consistency of features. The shift-invariant wavelet transform maintains consistency of features as images are translated. The discrete wavelet transform does not. For this reason, we expect the shift-invariant wavelet transform to provide superior registration accuracy over the discrete wavelet transform.

(a) Lenna



(b) Translated Lenna



(c) Lenna



(d) Rotated Lenna



Figure 4.4. Translated and rotated versions of Lenna. Top row: (a) Lenna (in rectangular coordinates) and (b) Lenna translated by  $T_x = 7$  and  $T_y = 4$  pixels (circular shifts). Second row: (c) Lenna (in polar coordinates) and (d) Lenna rotated by  $R = 13^\circ$ .

(a) Cameraman



(b) Translated cameraman



(c) Cameraman



(d) Rotated cameraman



Figure 4.5. Translated and rotated versions of cameraman. Top row: (a) Cameraman (in rectangular coordinates) and (b) Cameraman translated by  $T_x = 7$  and  $T_y = 4$  pixels (circular shifts). Second row: (c) Cameraman (in polar coordinates) and (d) Cameraman rotated by  $R = 13^\circ$ .

Figure 4.6 confirms this. The shift-invariant wavelet transform provides higher registration accuracy for every value of  $N$  for both values of  $PSNR$ . All the curves follow a general upward trend as  $N$  increases, supporting our expectation that as we use more significant coefficients in matching, we achieve a higher level of accuracy. The best explanation for the drop in performance of both transforms in Figure 4.6 d for  $N = 110$  and greater is that this is the point of diminishing returns for  $N$  for this image. Beginning with this value of  $N$ , enough noise creeps in and replaces legitimate significant wavelet coefficients as features for matching. Since we do not observe this effect for the shift-invariant wavelet transform in the plot for the Lenna image, we hypothesize that this point of diminishing returns occurs at a value of  $N$  greater than 250. Hence, the point at which noise adversely affects our registration accuracy is image dependent.

*4.3.2 Subband Comparison.* Since we have determined from our simulation that the shift-invariant wavelet transform provides better registration accuracy than the discrete wavelet transform, we now compare the performance of the detail subbands of the shift-invariant wavelet transform only.

The Lenna image contains several vertical edges: her hat, the feather in her hat, her right arm, and the right side of her face. The cameraman image is also heavily populated by vertical edges: the tripod, the body of the cameraman, and the buildings in the background. Although horizontal edges exist in both images, they are fewer and less persistent than the vertical edges.

Stone *et al* showed that the HH subband is the most sensitive subband to translation in the presence of white noise (28). Since, the HH subband is created by highpass filtering the rows and columns of the original image, we expect it to also be the most sensitive to noise since noise is comprised mainly of high frequency components. For these reasons, we anticipate the best registration accuracy from the LH subband since it preserves vertical edges better than the other subbands.

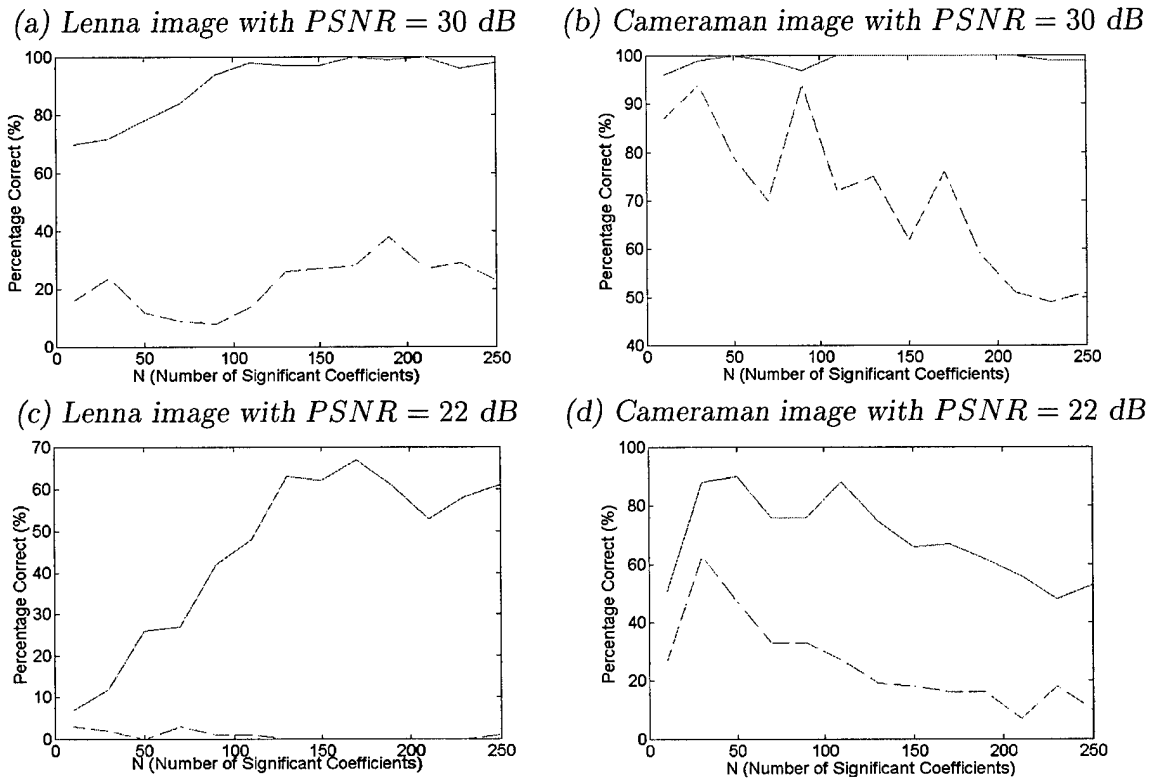


Figure 4.6. Registration accuracy of the shift-invariant (solid line) and discrete wavelet transforms (dashed line) versus the number of significant coefficients ( $N$ ) used for feature matching. For a translation of  $T_x = 7$  pixels and  $T_y = 4$  pixels, the shift-invariant wavelet transform provides better registration accuracy because it maintains consistent features.

Figure 4.7 supports this thought. For both Lenna plots and the cameraman plot for  $PSNR = 22$  dB, the LH subband is significantly more accurate than the other two subbands over the entire range of  $N$  because it extracts the best features, a direct result of the Lenna and cameraman images being dominated by vertical edges. The cameraman plot for  $PSNR = 30$  dB does have an anomaly at  $N = 70$  for which we have no immediate explanation. We leave this for future research. In Figures 4.7 a and b, where there is little noise, the HH subband provides consistently better registration than the HL subband. This is attributed to the fact that the HH subband preserves  $45^\circ$  edges and the hat and the tripod are dominant edges oriented at roughly  $45^\circ$  in the Lenna and cameraman images, respectively. Notice in Figures 4.7 c and d, when noise is significant, the HH subband provides little accuracy over the entire range of  $N$ , confirming our earlier thought. Recall from the previous section our discussion of the diminishing returns of  $N$  as noise becomes a factor. This is profoundly evident in Figure 4.7 d starting at  $N = 110$ .

#### 4.4 Rotation Performance

When estimating translation, we observed that the shift-invariant wavelet transform provided the best registration accuracy because it maintained consistency of features. The LH subband was the primary subband driving registration accuracy because our test images are highly populated with vertical edges, the information contained in the LH subband. We now examine the performance of our algorithm for estimating rotation to see if similar results occur.

*4.4.1 Redundant versus Standard Wavelet Transforms.* Recall from Section 3.2 the need for consistent features for matching. We found in Section 4.3.1 that a shift-invariant transform provided the consistency required for accurate registration results. We anticipate the same result when using our rotation-invariant polar wavelet transform.

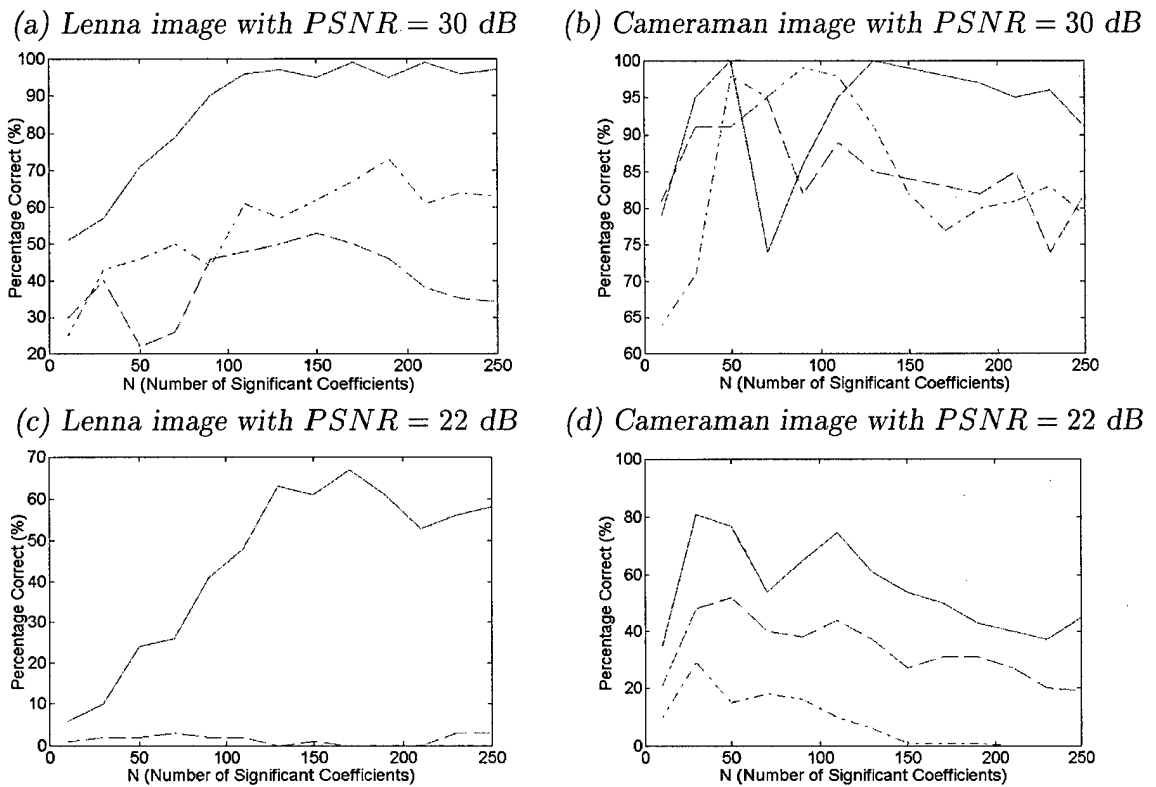


Figure 4.7. Registration accuracy of the LH (solid line), HL (dashed line), and HH (dash dot line) subbands versus the number of significant coefficients ( $N$ ) used for feature matching. For a translation of  $T_x = 7$  pixels and  $T_y = 4$  pixels, the LH subband provides the best registration accuracy since it preserves vertical edges.

The results for both the polar discrete wavelet transform and the rotation-invariant polar wavelet transform are outstanding for both images and for both values of  $PSNR$ . In all plots, for  $N = 50$  or greater, registration accuracy is consistently 95% accurate or higher. This level of accuracy is never achieved when estimating translation, even in the case where there is little noise present.

The rotation-invariant polar wavelet transform does provide better performance than the polar discrete wavelet transform. Recall from Section 4.2.4, we determine a correct rotation to be  $R = 12, 13, \text{ or } 14^\circ$  because the polar discrete wavelet transform cannot estimate odd valued rotations. Thus, if we only consider  $R = 13^\circ$  as a correct rotation estimate, the polar discrete wavelet transform will never estimate the true rotation. This further strengthens the idea that invariance is crucial in accurate registration.

*4.4.2 Subband Comparison.* As discussed in Section 4.3.2, both the Lenna and cameraman images are heavily dominated by vertical edges. For this reason, we expected the LH subband to provide the most accurate registration results since it preserves vertical edges. Because they are created from polar coordinates, the polar detail subbands provide different information than the standard detail subbands. The polar LH subband preserves edges around angles and the polar HL subband preserves edges along radii. Examining the polar versions of Lenna and cameraman (Figures 4.4 c and 4.5 c), we observe that both have edges that persist along the radii. For Lenna, these edges are her hat and feather; for cameraman, these edges are the tripod and his body. Hence, we anticipate the polar HL subband to provide the most accurate registration results.

Figure 4.9 illustrates what we expect. The polar HL subband consistently provides the best results, particularly when there is significant noise. When there is little noise present ( $PSNR = 30$  dB), the polar HH subband performs well. In the presence of significant noise, we observe that accuracy of the polar HH subband

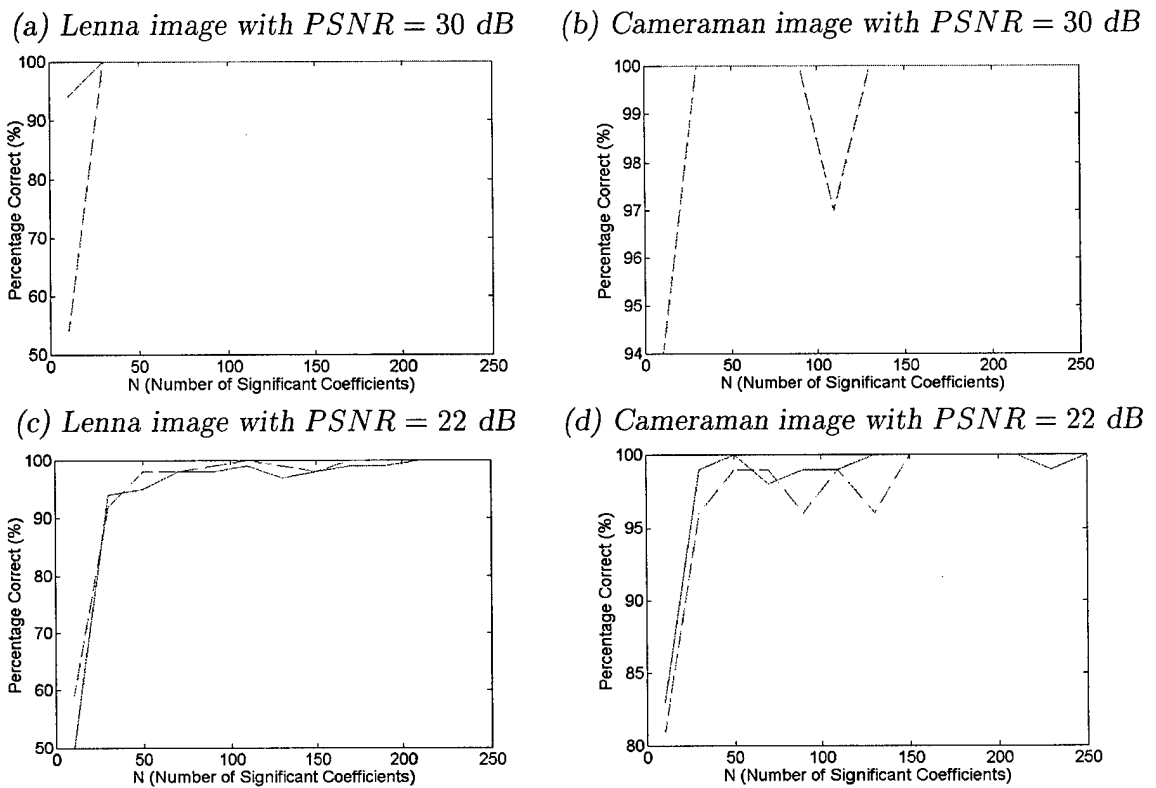


Figure 4.8. Registration accuracy of the rotation-invariant (solid line) and polar discrete wavelet transforms (dashed line) versus the number of significant coefficients ( $N$ ) used for feature matching. For a rotation of  $R = 13^\circ$ , the rotation-invariant polar wavelet transform provides better registration accuracy because it maintains consistent features.

declines considerably, although there is an upward trend as  $N$  increases. This decline is expected; even though we are now working on a polar sampling grid, noise still corresponds to high frequency components and the polar HH subband preserves these components along the radii and around the angles.

Having determined from our simulations that the bandpass subbands of redundant wavelet transforms provide the highest registration accuracy in the presence of noise, we move on to the final objective of our study, the empirical determination of the best selection for  $N$ .

#### 4.5 Empirical Determination of $N$

Since  $N$  is the number of the largest magnitude wavelet coefficients we keep for matching, we expect better algorithm performance as  $N$  increases (more features for matching intuitively equates to better accuracy). Examining the curves of Figures 4.6 and 4.8, this is the case (to a point). The costs of using larger values of  $N$  are a longer computation time and more significant coefficients are adversely affected by noise when matching features. Thus, our goal is to estimate the lowest  $N$  while maintaining the best possible registration accuracy.

The selections below are by no means optimal and the true choice of  $N$  is highly dependent on the image. We pick the smallest value of  $N$  which is 95% accurate since it is unrealistic to expect to achieve 100% accuracy in the presence of noise. If there is no  $N$  which produces that level of accuracy, we select the  $N$  which provides accuracy nearest to 95%.

When estimating translation, choosing  $N = 100$  gives a minimum registration accuracy of 95% when there is little noise present ( $PSNR = 30$  dB). For the case of significant noise ( $PSNR = 22$  dB), choosing  $N = 170$  gives a minimum registration accuracy of only 67%. Better results may be achieved with values of  $N$  beyond 250, but this is not guaranteed since there is a point of diminishing returns. For larger

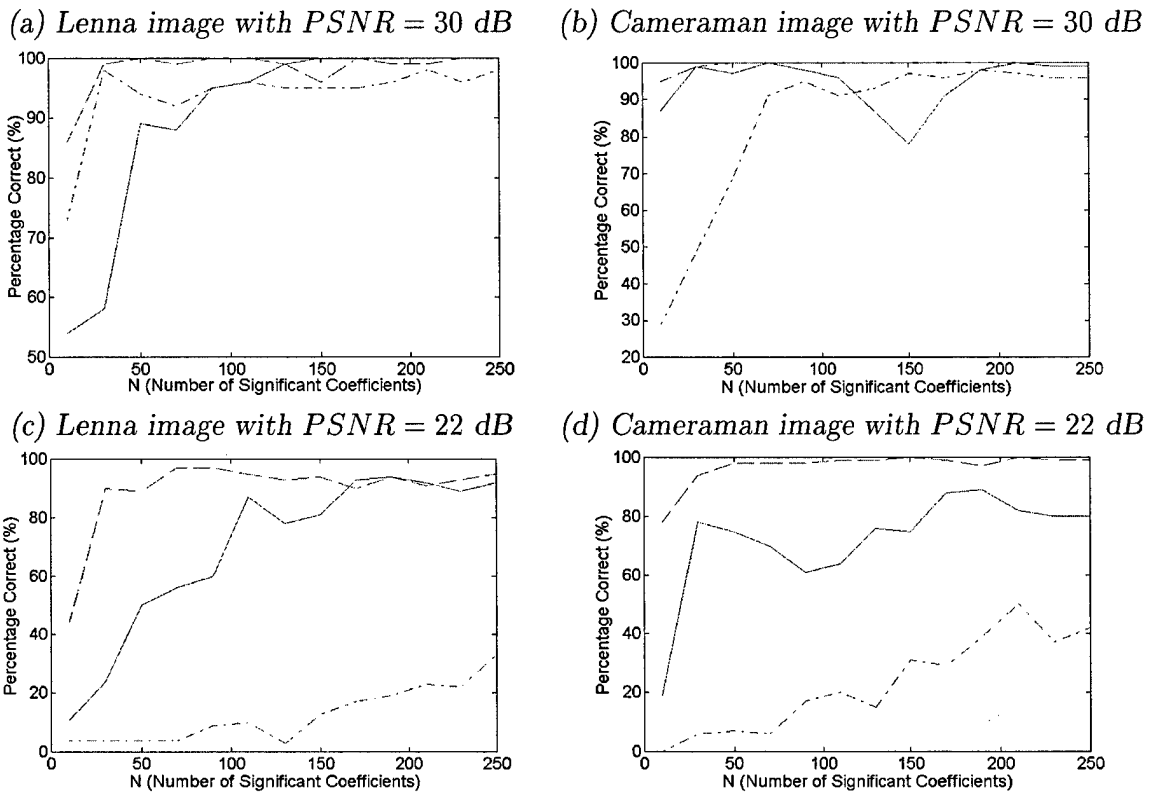


Figure 4.9. Registration accuracy of the polar LH (solid line), HL (dashed line), and HH (dash dot line) subbands versus the number of significant coefficients ( $N$ ) used for feature matching. For a rotation of  $R = 13^\circ$ , the polar HL subband provides the best registration accuracy since it preserves edges along the radii.

values of  $N$ , the significant coefficients are more adversely affected by noise. Hence, we leave finding the optimum choice of  $N$  as future research.

When estimating rotation, choosing  $N = 50$  gives a minimum registration accuracy of 95% regardless of the amount of noise present. This small value of  $N$  allows for fast rotation estimation.

#### 4.6 *Robustness*

We have examined the performance of our algorithm for a specific translation and a specific rotation independently. We now attempt to determine empirically the amount of translation we can tolerate and still correctly estimate rotation when noise is not present. Prior to the conversion to polar coordinates, we shifted our reference image (thus misaligning the origin). We then ran our registration routine to determine the effect of this misalignment on our algorithm performance.

From Table 4.1, we see that for the Lenna image, our algorithm is highly shift dependent as to whether we correctly register the input image for a rotation of  $R = 13^\circ$  (using our criteria of Section 4.2.4). For the cameraman image (Table 4.2), the opposite is true. The error is minimal and slowly grows with more pixel translations (as we go down and right across the table). Thus, we see that our algorithm robustness is highly image dependent, but we can still determine the correct rotation for small levels of misalignment.

#### 4.7 *Summary*

After describing the design of our validation study to include the measures of performance, why the Daubechies (7,9) wavelet was selected over other wavelets for use in testing, and exactly how the data was collected, we analyzed the results obtained from our simulations for translation first and rotation second. We observed that the redundant wavelet transforms provided higher registration accuracy than the discrete wavelet transforms because of their invariance, which allowed for the

**Magnitude of Error for Lenna (in degrees)**

	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
<b>0</b>	0	0	42	44	2	0	0	0	0	0	0
<b>1</b>	90	0	41	1	0	2	0	3	3	3	3
<b>2</b>	1	173	160	172	1	171	2	175	13	179	12
<b>3</b>	159	159	0	173	191	1	171	14	179	13	12
<b>4</b>	180	1	104	0	0	1	15	180	14	180	3
<b>5</b>	2	158	1	103	0	0	1	15	14	3	3
<b>6</b>	0	0	103	1	170	16	102	15	0	0	1
<b>7</b>	103	0	0	102	102	101	16	0	0	168	3
<b>8</b>	156	102	0	0	101	101	101	16	100	100	168
<b>9</b>	0	102	101	101	0	101	100	100	100	0	0
<b>10</b>	0	0	101	101	100	0	100	100	99	99	0

Table 4.1. *Algorithm robustness for Lenna. The first row (in boldface) is the amount the input image pixels have been shifted in the positive x-direction. The first column (in boldface) is the amount the input image pixels have been shifted in the positive y-direction. The values of the table are the magnitude of error in degrees between the actual rotation value ( $R = 13^\circ$ ) and the estimated rotation value for the noiseless case when  $N = 10$ . The more pixel translation the algorithm can tolerate (i.e., still correctly estimate rotation), the higher the level of robustness. Note that the shift is applied prior to the conversion to polar coordinates to simulate misalignment of the origin.*

Magnitude of Error for Cameraman (in degrees)

	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
<b>0</b>	0	0	0	2	2	2	2	1	3	3	3
<b>1</b>	0	0	1	1	1	0	0	0	2	2	2
<b>2</b>	0	0	1	0	44	0	0	1	1	1	0
<b>3</b>	2	4	0	0	0	0	0	0	1	46	0
<b>4</b>	1	1	1	1	2	1	2	0	0	0	0
<b>5</b>	2	2	2	3	3	1	1	1	1	1	1
<b>6</b>	3	2	2	2	2	2	2	2	2	2	47
<b>7</b>	3	3	3	3	3	3	3	3	2	2	2
<b>8</b>	4	3	6	4	4	4	3	3	3	3	3
<b>9</b>	5	5	5	4	7	47	5	4	4	4	3
<b>10</b>	6	5	6	5	5	5	5	5	5	4	4

Table 4.2. *Algorithm robustness for cameraman. The first row (in boldface) is the amount the input image pixels have been shifted in the positive x-direction. The first column (in boldface) is the amount the input image pixels have been shifted in the positive y-direction. The values of the table are the magnitude of error in degrees between the actual rotation value ( $R = 13^\circ$ ) and the estimated rotation value for the noiseless case when  $N = 10$ . The more pixel translation the algorithm can tolerate (i.e., still correctly estimate rotation), the higher the level of robustness. Note that the shift is applied prior to the conversion to polar coordinates to simulate misalignment of the origin.*

extraction of consistent features. The bandpass subbands were observed to provide solid registration results in the presence of little and significant noise, confirming our choice to use them in parameter estimation. We empirically determined the best choices of  $N$  depending on the amount of noise present observing that we achieve much better results for much smaller values of  $N$  when estimating rotation. Finally, we tested the robustness of the algorithm when estimating an image that had been translated and rotated. For this case, we saw that our algorithm robustness was highly image dependent, but we were able to determine the correct rotation estimate for small levels of misalignment.

## V. Discussion and Future Work

### 5.1 Contributions of this Thesis

In this thesis, we developed a new algorithm for registering a strongly translated or rotated input image relative to a reference image in the presence of additive white Gaussian noise. Redundant wavelet transforms outperformed the discrete wavelet transform because of their invariance. We verified that although the high-pass subband was too sensitive to noise to use in registration, the bandpass subbands were well suited because of the characteristics of the original image they preserved. We showed that registration accuracy was dependent on the number of significant coefficients ( $N$ ) kept for a given image  $PSNR$  and that our algorithm was robust enough to handle small levels of misalignment when estimating rotation.

In accomplishing the goal of accurate registration, we also successfully developed and utilized a new computationally efficient masking procedure to suppress the adverse effects of noise and a rotation-invariant wavelet transform, which proved especially effective for estimating rotation.

### 5.2 Potential for Future Research

There are several areas of potential research that may be explored in future theses. We present the most interesting.

*5.2.1 Develop a More Robust Polar Wavelet Transform.* Currently the polar redundant wavelet transform is computed simply by converting rectangular image coordinates to polar coordinates and then performing the shift-invariant wavelet transform. Although shown to be effective in the image registration problem, this method of creating the wavelet transform violates the wavelet recursion equations. A more mathematically accurate transform must be developed. Additionally, ap-

plication of this new transform to problems other than image registration must be explored.

*5.2.2 Leverage Multiscale Properties.* Our registration algorithm is designed for accuracy, not speed. For use in real world applications, the algorithm must be made more computationally efficient to better handle large sets of imagery. A possible way to speed up the algorithm is to leverage the multiscale properties of wavelets and use a coarse-to-fine iteration strategy between the same and different subbands. There is particular promise for cultivating this idea since wavelet coefficients at higher scales appear to be more correlated than those at lower ones (30). Additionally, different (and perhaps better) information may be provided at the lower scales.

*5.2.3 Calculate Translation and Rotation Simultaneously.* Our algorithm performs exceptionally well in registering strongly translated or rotated images; however, it does not support translation and rotation simultaneously. Obviously real world data collected by aircraft or satellites has components of both as well as some scaling and perhaps shearing. A true challenge is to modify our algorithm to simultaneously process translations and rotations, and then to extend to more difficult data sets where affine and polynomial transformations are required. This is an exceptionally challenging problem since estimating translation or rotation alone is a significant task; joint estimation is many times more difficult.

## *Appendix A. Additional Information on Filter Banks*

We provide additional information regarding signal reconstruction using filter banks. For more information, refer to (5, 32, 33).

### *A.1 Polyphase Representation*

The polyphase representation provides an easier way to analyze filter banks (5, 32). To best explain the polyphase representation, we will divide the filter  $h(n)$  into its even and odd polyphase components. First, we begin with the definition of the z-transform (32) of  $h(n)$

$$H(z) = \sum_{n=-\infty}^{\infty} h(n)z^{-n}.$$

Then we break that sequence into a sequence of the even components of  $h(n)$  and a sequence of the odd components of  $h(n)$ ,

$$H(z) = \sum_{n=-\infty}^{\infty} h(2n)z^{-2n} + \sum_{n=-\infty}^{\infty} h(2n+1)z^{-(2n+1)}.$$

Simplifying the  $z$  exponent in the last term above since we are integrating over all possible values of  $n$ , we see

$$H(z) = \sum_{n=-\infty}^{\infty} h(2n)z^{-2n} + \sum_{n=-\infty}^{\infty} h(2n+1)z^{-2n}.$$

Finally, we write  $H(z)$  in terms of new filters

$$\begin{aligned} E_0(z) &= \sum_{n=-\infty}^{\infty} h(2n)z^{-n} \\ E_1(z) &= \sum_{n=-\infty}^{\infty} h(2n+1)z^{-n}, \end{aligned}$$

which are the even and odd polyphase components, respectively. This yields

$$H(z) = E_0(z^2) + z^{-1}E_1(z^2).$$

Figure A.1 illustrates how polyphase components replace the original filter  $H(z)$  in a filter bank. First, we start with the original system. Next, we break  $H(z)$  into its even and odd polyphase components and then apply the same decimator to both channels. Finally, we invoke the Noble Identity for Decimators (Equation 2.4) and bring  $E_0(z^2)$  and  $E_1(z^2)$  through the decimators to achieve the polyphase representation of the original system.

In general, the polyphase representation of a filter  $H(z)$  decimated by decimation ratio  $M$  is

$$\begin{aligned} H(z) &= E_0(z^M) + z^{-1}E_1(z^M) + \dots + z^{-(M-1)}E_{M-1}(z^M) \\ &= \sum_{i=0}^{M-1} z^{-i}E_i(z^M) \end{aligned}$$

The polyphase representation is derived in the same manner for expanders with the exception that the Noble Identity for Expanders (Equation 2.5) is applied instead of the Noble Identity for Decimators.

## A.2 Perfect Reconstruction

Perfect reconstruction is achieved when the output of a filter bank is a delayed (and maybe scaled) version of the input (5, 32). Given a set of analysis bank filters, perfect reconstruction may be achieved through careful selection of the synthesis bank filters.

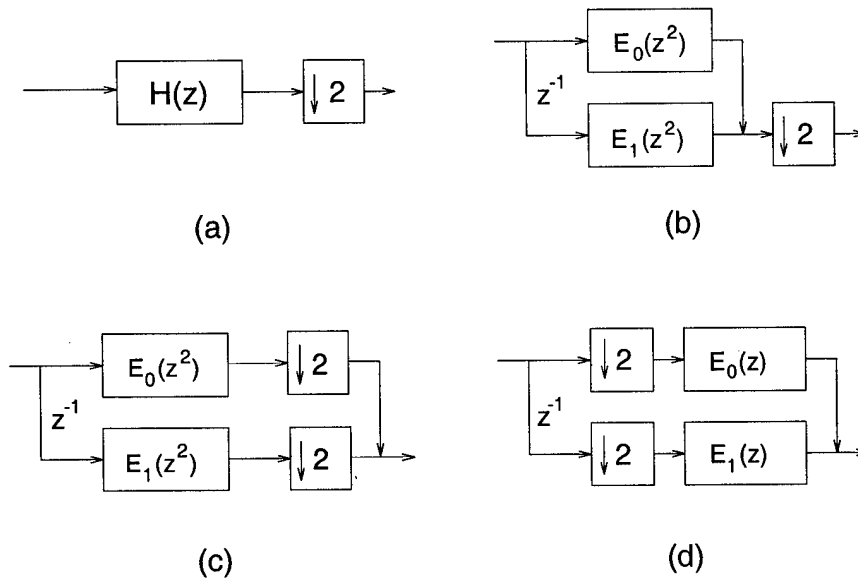


Figure A.1. *Polyphase representation. (a) Original system. (b) Substitution of polyphase components for  $H(z)$ . (c) Splitting the decimator. (d) Application of Noble Identity for Decimators.*

Each of the filters  $H_i(z)$  and  $F_j(z)$  are first written in terms of their polyphase components,  $E_{ij}(z)$  and  $R_{ij}(z)$ , as follows

$$H_i(z) = \sum_{j=0}^{M-1} z^{-j} E_{i,j}(z^M)$$

$$F_j(z) = \sum_{i=0}^{M-1} z^{-(M-1-i)} R_{i,j}(z^M).$$

For convenience in analysis, the polyphase components of the analysis filters  $H_i(z)$  may be placed in the matrix  $E(z^M)$ , where the  $i^{\text{th}}$  row of  $E(z^M)$  contains the polyphase components of  $H_i(z)$ . Similarly, the polyphase components of the synthesis filters  $F_j(z)$  comprise the  $j^{\text{th}}$  columns of the  $R(z^M)$  polyphase matrix. Thus,

we have

$$E(z^M) = \begin{bmatrix} E_{0,0}(z^M) & E_{0,1}(z^M) & \cdots & E_{0,M-1}(z^M) \\ E_{1,0}(z^M) & E_{1,1}(z^M) & \cdots & E_{1,M-1}(z^M) \\ \vdots & \vdots & \ddots & \vdots \\ E_{M-1,0}(z^M) & E_{M-1,1}(z^M) & \cdots & E_{M-1,M-1}(z^M) \end{bmatrix}$$

$$R(z^M) = \begin{bmatrix} R_{0,0}(z^M) & R_{0,1}(z^M) & \cdots & R_{0,M-1}(z^M) \\ R_{1,0}(z^M) & R_{1,1}(z^M) & \cdots & R_{1,M-1}(z^M) \\ \vdots & \vdots & \ddots & \vdots \\ R_{M-1,0}(z^M) & R_{M-1,1}(z^M) & \cdots & R_{M-1,M-1}(z^M) \end{bmatrix}$$

Applying the Noble Identities to these matrices, we can redraw our  $M$  channel filter bank into the polyphase equivalent shown in Figure A.2.

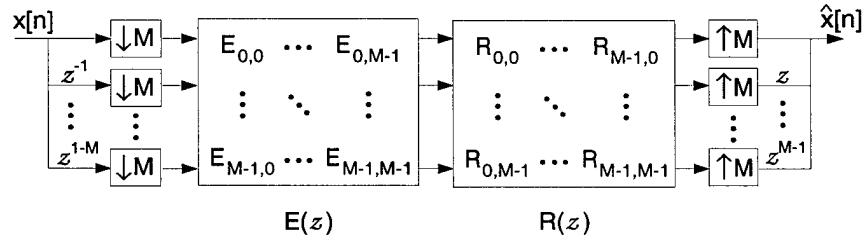


Figure A.2. Polyphase representation of an  $M$ -channel, maximally decimated uniform filter bank. Polyphase matrices  $E(z)$  and  $R(z)$  are both  $M \times M$ .

We achieve perfect reconstruction if and only if

$$P(z) = R(z)E(z) = cz^{-n_0} \begin{bmatrix} 0 & I_{M-r} \\ z^{-1}I_r & 0 \end{bmatrix} \quad (\text{A.1})$$

where  $c$  is a scalar,  $n_0$  is a nonnegative integer, and  $r \in \{0, 1, \dots, M-1\}$ . When we let  $r = 0$ , we can simplify the equation above to  $P(z) = R(z)E(z) = cz^{-n_0}I_M$ . Clearly

we must have  $R(z) = cz^{-n_0} E^{-1}(z)$  for this equation to hold true for a given  $E(z)$ . In general,  $E^{-1}(z)$  does not exist. If  $E^{-1}(z)$  does exist, it may only be realizable with infinite impulse response (IIR) filters even if  $E(z)$  is composed entirely of finite impulse response (FIR) filters. To avoid this problem, we must choose  $E(z)$  such that the  $\det[E(z)]$  is a delay (32).

A special useful case is when we select  $E(z)$  to be paraunitary.  $E(z)$  is paraunitary when  $E^{-1}(z) = \tilde{E}(z)$ , where  $\tilde{E}(z) \equiv E^H(z^{-1})$ . In other words, we find  $E^H(z^{-1})$  by replacing all  $z$ 's in  $E(z)$  with  $z^{-1}$ , transposing, and taking the complex conjugate. To satisfy Equation A.1 and achieve perfect reconstruction, we let  $R(z) = \tilde{E}(z)$  to ensure  $P(z) = R(z)E(z) = I$ . An additional benefit of choosing  $R(z) = \tilde{E}(z)$  is that if  $E(z)$  is comprised entirely of FIR filters, then  $R(z)$  is guaranteed to also be comprised entirely of FIR filters. Thus, paraunitary filters greatly simplify the design of FIR perfect reconstruction filter banks (32, 33).

## *Appendix B. Additional Information on Wavelets*

We provide additional information regarding reconstruction of orthogonal and biorthogonal wavelets and selection of filters that result in perfect reconstruction for the filter bank implementation of the discrete wavelet transform.

### *B.1 Reconstruction*

In Sections 2.3.1, 2.3.2, and 2.3.3, we showed how to decompose the signal  $f$  into its orthogonal components. Now, we will show how to reconstruct  $f$  from these components.

Starting at scale  $m = 1$ , we know that  $f$  exists in  $V_0$  and can be represented as

$$f(t) = \sum_k c_0(k) \phi_{0,k}(t).$$

We must be given the coarse and detail coefficients,  $c_1(k)$  and  $d_1(k)$ . Recall that these coefficients represent the projection of  $f$  into  $V_1$  and  $W_1$ , respectively. Since  $V_1 \oplus W_1$  and  $V_1 \cap W_1 = \phi$ ,

$$f(t) = \sum_k c_1(k) \phi_{1,k}(t) + \sum_k d_1(k) \psi_{1,k}(t). \quad (\text{B.1})$$

Next, we take the inner product of both sides of the equation with  $\phi_{0,n}(t)$ . Since  $\{\phi_{0,n}(t)\}$  are orthonormal,  $\langle f, \phi_{0,n} \rangle = c_0(n)$ . Recall from Section 2.3.2

$$\begin{aligned} h(n-2k) &= \langle \phi_{1,k}, \phi_{0,n} \rangle \\ g(n-2k) &= \langle \psi_{1,k}, \phi_{0,n} \rangle. \end{aligned}$$

Substituting into Equation B.1

$$c_0(n) = \sum_k c_1(k)h(n-2k) + \sum_k d_1(k)g(n-2k). \quad (\text{B.2})$$

where  $c_0(n)$  are the coefficients of  $f \in V_0$ . This provides insight into signal reconstruction when using the filter bank implementation of the inverse discrete wavelet transform. Note that for the orthogonal discrete wavelet transform described, the digital filters  $h$  and  $g$  must be the same for the forward and inverse transforms to achieve perfect reconstruction of the signal (3).

Having presented orthogonal reconstruction, we present a more detailed look at biorthogonal wavelets to include the biorthogonal wavelet recursion equations, construction of the discrete wavelet transform using biorthogonal wavelets, and signal reconstruction.

## B.2 Biorthogonal Wavelet Recursion Equations

Like orthogonal wavelets, we must have wavelet recursion relations

$$\begin{aligned} \phi_{m,n}(t) &= \sum_k h_n(k)\phi_{m-1,k}(t) \\ \psi_{m,n}(t) &= \sum_k g_n(k)\phi_{m-1,k}(t). \end{aligned}$$

To find  $h$ , we must take the inner product of  $\phi_{m,n}$  with its dual function  $\tilde{\phi}_{m-1,k}$

$$\begin{aligned} \langle \phi_{m,n}, \tilde{\phi}_{m-1,k} \rangle &= \sum_n h_n(n) \langle \phi_{m-1,n}, \tilde{\phi}_{m-1,k} \rangle \\ &= \sum_n h_n(n) \delta(k-n) \\ &= h_n(k). \end{aligned}$$

Similarly,  $g_n(k) = \langle \psi_{m,n}, \tilde{\phi}_{m-1,k} \rangle$ . We have equivalent recursion relations with  $\tilde{h}$  and  $\tilde{g}$  for the dual basis.

### B.3 Construction of the Biorthogonal Discrete Wavelet Transform

Given the dual basis, we will decompose a signal  $f = \sum_k c_{m-1}(k)\tilde{\phi}_{m-1,k}(t)$ , which exists entirely in  $V_{m-1}$ , into

$$f(t) = \sum_k c_m(k)\tilde{\phi}_{m,k}(t) + \sum_k d_m(k)\tilde{\psi}_{m,k}(t)$$

To find  $c_m(k)$ , we take the inner product of  $f$  with  $\phi_{m,k}$  due to the biorthogonality condition. This yields

$$\begin{aligned} c_m(k) &= \langle f, \phi_{m,k} \rangle \\ &= \sum_n c_{m-1}(n) \langle \phi_{m,k}, \tilde{\phi}_{m-1,n} \rangle \\ &= \sum_n c_{m-1}(n) h_k(n). \end{aligned} \tag{B.3}$$

Likewise, we find  $d_m(k)$  by taking the inner product of  $f$  with  $\psi_{m,k}$

$$d_m(k) = \sum_n c_{m-1}(n) g_k(n). \tag{B.4}$$

The digital filters  $h$  and  $g$  are independent of the level of decomposition. Equations B.3 and B.4 perform the decomposition and are equivalent to the orthogonal wavelet transform equations derived in Section 2.3.3.

### B.4 Reconstruction with the Biorthogonal Discrete Wavelet Transform

Given the biorthogonal wavelet decomposition for scale  $m$ ,

$$f(t) = \sum_k c_m(k)\tilde{\phi}_{m,k}(t) + \sum_k d_m(k)\tilde{\psi}_{m,k}(t),$$

we will reconstruct our signal  $f$ .

We know  $f$  exists entirely in  $V_{m-1}$  and may be represented as

$$f(t) = \sum_k c_{m-1}(k) \tilde{\phi}_{m-1,k}(t).$$

Taking the inner product of  $f$  and  $\phi_{m-1,n}(t)$ , we get

$$\begin{aligned} \langle f, \phi_{m-1,n} \rangle &= \sum_k c_{m-1}(k) \langle \tilde{\phi}_{m-1,k}, \phi_{m-1,n} \rangle \\ &= \sum_k c_{m-1}(k) \delta(k-n) \\ &= c_{m-1}(n) \\ &= \sum_k c_m(k) \langle \tilde{\phi}_{m,k}, \phi_{m-1,n} \rangle + \sum_k d_m(k) \langle \tilde{\psi}_{m,k}, \phi_{m-1,n} \rangle \\ &= \sum_k c_m(k) \tilde{h}_k(n) + \sum_k d_m(k) \tilde{g}_k(n). \end{aligned}$$

We see that this is exactly the same as in the orthogonal transform case except that we use biorthogonal filters  $\tilde{h}$  and  $\tilde{g}$  which are different from our  $h$  and  $g$  filters used for decomposition.

### B.5 Selection of Wavelet Filters

In section A.2, we stated that careful selection of the synthesis bank filters leads to perfect reconstruction. We now discuss how to select these filters.

Given lowpass filter  $h(n)$ , the following choices

$$\begin{aligned} G(z) &= H(-z) \\ \tilde{H}(z) &= H(z) \\ \tilde{G}(z) &= -G(z) = -H(-z) \end{aligned}$$

lead to alias cancellation. However, the only orthogonal wavelet with these properties that may be implemented with FIR filters which yield perfect reconstruction are the

Haar filters (32):

$$\begin{aligned}h &= [.7071 \ .7071] \\g &= [.7071 \ - .7071]\end{aligned}$$

Thus, we must move to other FIR filters that yield perfect reconstruction. The best set to use are the paraunitary filters because they greatly simplify the design of perfect reconstruction FIR filters. In this case, we have

$$\begin{aligned}G(z) &= H(-z^{-1}) \\ \widetilde{H}(z) &= H(z^{-1}) \\ \widetilde{G}(z) &= G(z^{-1})\end{aligned}$$

which corresponds to orthogonal wavelets (32).

For biorthogonal wavelets, we have the similar relationship between  $H(z)$  and  $\widetilde{G}(z)$ , and between  $G(z)$  and  $\widetilde{H}(z)$ .  $H(z)$  and  $G(z)$  are not related in the orthogonal way.

## Bibliography

1. Alliney, S. and C. Morandi. "Digital Image Registration Using Projections," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(2):222-233 (1986).
2. Brown, L. G. "A Survey of image Registration Techniques," *ACM Computing Surveys*, 24(4):325-376 (December 1992).
3. Burrus, C. Sidney, et al. *Introduction to Wavelets and Wavelet Transforms: A Primer*. Englewood Cliffs, NJ: Prentice-Hall, 1997.
4. Castro, E. De and C. Morandi. "Registration of Translated and Rotated Images Using Finite Fourier Transforms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(5):700-703 (1987).
5. Claypoole, Roger L. *Multipoint Multirate Signal Processing*. MS thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, December 1994.
6. Claypoole, Roger L. *Adaptive Wavelet Transforms Via Lifting*. PhD dissertation, Rice University, Houston, TX, October 1999.
7. Corvi, M. and G. Nicchiotti. "Multiresolution Image Registration." *Proceedings of IEEE International Conference on Image Processing*. 224-227. October 23-26 1995.
8. Daubechies, I. *Ten Lectures on Wavelets*. CBMS-NSF Regional Conf. Series in Appl. Math., Vol. 61, Philadelphia, PA: SIAM, 1992.
9. Djamjdi, J. P., et al. "Geometrical Registration of Images. The Multiresolution Approach," *Photogrammetric Engineering and Remote Sensing Journal*, 59(5) (May 1993).
10. Edwards, C. H. and David E. Penney. *Calculus and Analytic Geometry*. Englewood Cliffs, New Jersey: Prentice Hall, 1990.
11. Kingsbury, N. "Shift-Invariant Properties of the Dual-Tree Complex Wavelet Transform." *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 3. 1221 - 1224. 1999.
12. Kuglin, C. D. and D. C. Hines. "The Phase Correlation Image Alignment Method." *Proceedings of IEEE International Conference on Cybernetics and Society*. 163-165. September 1975.
13. LeMoigne, J. "Parallel Registration of Multi-Sensor Remotely Sensed Imagery Using Wavelet Coefficients." *OE/Aerospace Sensing, Wavelet Applications Conference*, 2242. 432-443. April 5-8 1994.

14. LeMoigne, J. "Towards a Parallel Registration of Multiple Resolution Remote Sensing Data." *International Geoscience and Remote Sensing Symposium*. 1011-1013. July 10-14 1995.
15. LeMoigne, J., et al., "An Automated Parallel Image Registration Technique of Multiple Source Remote Sensing Data." *CESDIS TR-96-182* - submitted to *IEEE Transactions in Geoscience and Remote Sensing*.
16. LeMoigne, J. and R. F. Crompt. "The Use of Wavelets for Remote Sensing Image Registration and Fusion." *Wavelet Applications*, 2762. 535-544. 1996.
17. LeMoigne, J. and I. Zavorin. "Use of Wavelets for Image Registration." *Wavelet Applications VII 4056*, edited by H. Szu, et al. 99-108. 2000.
18. Li, H. H. and Y. Zhou. "A Wavelet-Based Feature Extractor for Multi-Sensor Image Registration." *Proceedings of the SPIE*, 2762. 524-534. 1996.
19. Lim, J. S. *Two-Dimensional Signal and Image Processing*. New Jersey: Prentice Hall, 1990.
20. Mallat, S. and W. Hwang. "Singularity Detection and Processing with Wavelets," *IEEE Transactions on Information Theory*, 38(2):617-643 (1992).
21. Mallat, S. and S. Zhong. "Characterization of Signals from Multiscale Edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:710-732 (July 1992).
22. Mallat, S. G. "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674-693 (1989).
23. Manjunath, B. S., et al. "A New Approach to Image Feature Detection with Applications," *Elsevier Science Pattern Recognition*, 29(4):627-640 (1996).
24. Reddy, B. S. and B. N. Chatterji. "An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration," *IEEE Transactions on Signal Processing*, 3(8):1266-1270 (August 1996).
25. Sharman, R., et al. "Compensating for Wavelet Sensitivity to Translation in Image Registration Applications." *Wavelet Applications VI*, 3723. 466-468. 1999.
26. Sharman, R., et al. "A Fast and Accurate Way to Register Medical Images Using Wavelet Modulus Maxima," *Elsevier Science Pattern Recognition Letters*, 21:447-462 (2000).
27. Simoncelli, E. P., et al. "Shiftable Multiscale Transforms," *IEEE Transactions on Information Theory*, 38(2):587-607 (March 1992).

28. Stone, H. S., et al. "The Translation Sensitivity of Wavelet-Based Registration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(10):1074-1081 (October 1999).
29. Strang, G. and T. Nguyen. *Wavelets and Filter Banks*. Wellesley, Cambridge, 1996.
30. Tashakkori, R., et al. "Prediction of Medical Images Using Wavelets." *Wavelet Applications VII, 4056*, edited by H. Szu, et al. 332-340. 2000.
31. Unser, M., et al. "Registration and Statistical Analysis of PET Images Using the Wavelet Transform," *IEEE Engineering in Medicine and Biology*, 603-611 (September/October 1995).
32. Vaidyanathan, P. P. *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice Hall, 1992.
33. Vetterli, Martin. "A Theory of Multirate Filter Banks," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(3):356-372 (March 1987).
34. Zheng, Q. and R. Chellappa. "A Computational Vision Approach to Image Registration." *IEEE Transactions on Signal Processing*, 2. 311-326. December 1993.

## *Vita*

First Lieutenant Richard K. Brown, Jr., graduated salutatorian from Rockland High School in 1993. Accepting an appointment to the United States Naval Academy, he graduated with distinction in 1997, earning a Bachelor of Science in Systems Engineering. Upon accepting a commission in the United States Air Force, he was stationed at Hickam Air Force Base, HI, where he served as Director of Telephone Operations and later as Chief of the Network Control Center. Since coming to the Air Force Institute of Technology in 1999, he has pursued a Masters of Science in Electrical Engineering. He is a member of the Tau Beta Pi and Eta Kappa Nu engineering honor societies.

Permanent address: 2950 P Street, Bldg 640  
Wright-Patterson AFB, OH 45433

<b>REPORT DOCUMENTATION PAGE</b>	Form Approved OMB No. 074-0188
----------------------------------	-----------------------------------

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of the collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

**PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD-MM-YYYY)</b> 07-03-2001	<b>2. REPORT TYPE</b> Master's Thesis	<b>3. DATES COVERED (From - To)</b> Jan 2000 - Mar 2001
--	--	--

<b>4. TITLE AND SUBTITLE</b>  IMAGE REGISTRATION USING REDUNDANT WAVELET TRANSFORMS	<b>5a. CONTRACT NUMBER</b>
	<b>5b. GRANT NUMBER</b>
	<b>5c. PROGRAM ELEMENT NUMBER</b>

<b>6. AUTHOR(S)</b>  Brown, Richard K., Jr., First Lieutenant, USAF	<b>5d. PROJECT NUMBER</b>
	<b>5e. TASK NUMBER</b>
	<b>5f. WORK UNIT NUMBER</b>

<b>7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)</b>  Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 P Street, Building 640 WPAFB OH 45433-7765	<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>  AFIT/GE/ENG/01M-21
---	---

<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>  AFRL/IFED Attn: Maj John Keller 32 Brooks Rd. Rome, NY 13441-4114  Commercial: (315) 330-3944	<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>
	<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>

<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b>  APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.
--

<b>13. SUPPLEMENTARY NOTES</b>
--------------------------------

**14. ABSTRACT**  
Imagery is collected much faster and in significantly greater quantities today compared to a few years ago. Accurate registration of this imagery is vital for comparing the similarities and differences between multiple images. Since human analysis is tedious and error prone for large data sets, we require an automatic, efficient, robust, and accurate method to register images. Wavelet transforms have proven useful for a variety of signal and image processing tasks, including image registration. In our research, we present a fundamentally new wavelet-based registration algorithm utilizing redundant transforms and a masking process to suppress the adverse effects of noise and improve processing efficiency. The shift-invariant wavelet transform is applied in translation estimation and a new rotation-invariant polar wavelet transform is effectively utilized in rotation estimation. We demonstrate the robustness of these redundant wavelet transforms for the registration of two images (i.e., translating or rotating an input image to a reference image), but extensions to larger data sets are certainly feasible. We compare the registration accuracy of our redundant wavelet transforms to the "critically sampled" discrete wavelet transform using the Daubechies (7,9) wavelet to illustrate the power of our algorithm in the presence of significant additive white Gaussian noise and strongly translated or rotated images.

**15. SUBJECT TERMS**  
image registration, wavelets, shift-invariance, rotation-invariance, masking, feature extraction

<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b>
<b>a. REPORT</b>	<b>b. ABSTRACT</b>	<b>c. THIS PAGE</b>			Maj Roger L. Claypoole, Jr., ENG
U	U	U	UU	90	<b>19b. TELEPHONE NUMBER (Include area code)</b> (937) 255-3636, ext 4625

Standard Form 298 (Rev. 8-98)  
Prescribed by ANSI Std. Z39-18