ABSTRACT

Title of dissertation: ANISOTROPIC HARMONIC ANALYSIS AND INTEGRATION OF REMOTELY SENSED DATA

James M. Murphy, Doctor of Philosophy, 2015

Dissertation directed by: Dr. John J. Benedetto
Department of Mathematics

Dr. Wojciech Czaja
Department of Mathematics

This thesis develops the theory of discrete directional Gabor frames and several algorithms for the analysis of remotely sensed image data, based on constructions of harmonic analysis. The problems of image registration, image superresolution, and image fusion are separate but interconnected; a general approach using transform methods is the focus of this thesis. The methods of geometric multiresolution analysis are explored, particularly those related to the shearlet transform. Using shearlets, a novel method of image registration is developed that aligns images based on their shearlet features. Additionally, the anisotropic nature of the shearlet transform is deployed to smoothly superrsolve remotely-sensed image with edge features. Wavelet packets, a generalization of wavelets, are utilized for a flexible image fusion algorithm. The interplay between theoretical guarantees for these mathematical constructions, and their effectiveness for image processing is explored throughout.

ANISOTROPIC HARMONIC ANALYSIS AND INTEGRATION OF REMOTELY SENSED DATA

by

James Michael Murphy

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Dr. John J. Benedetto, Co-Chair
Dr. Wojciech Czaja, Co-Chair
Dr. Kasso Okoudjou
Dr. Radu Balan
Dr. David Mount
Dr. Jacqueline Le Moigne

Dedication

This thesis is dedicated to my parents, Michael and Lisa, and to my brother, William. Thank you for supporting me on this path.

Acknowledgments

I would like to start by thanking my thesis advisors, Dr. Wojciech Czaja and Dr. John J. Benedetto. I arrived at the University of Maryland with interests in geometry, but those quickly evaporated after taking the first year sequence in the subject. Real analysis, on the other hand, was wonderful. In particular, I enjoyed my professor, Wojtek Czaja, and we started meeting regularly to discuss mathematics. I was encouraged by Wojtek to take a topics class with John, and I soon found a home in the Norbert Wiener Center for Harmonic Analysis and Applications. My interactions with Wojtek and John opened up a beautiful world of pure mathematics, but perhaps their most significant contribution to my education and training was encouraging me to move beyond esoteric analysis into more down-to-earth applications. I refused at first, but eventually relented. Thankfully, they were more convincing than I was stubborn. As this thesis attests, I found great joy in developing applications with mathematics. For this, I am most grateful to Wojtek and John.

Dr. Kasso Okoudjou and Dr. Radu Balan have also contributed to my education in the Norbert Wiener Center. I developed a diverse mathematical palette, in no small part through the courses and seminars I have taken with them. From graph theory to sparse signal representations, I feel comfortable in the larger millieu of harmonic analysis, beyond the work emphasized in this thesis. I owe this breadth to what I have learned from them.

I have had the great fortune to be mentored outside the University of Mary-

land. Dr. Jacqueline Le Moigne and Dr. David Harding of NASA's Goddard Space Flight Center taught me much during summer 2014 and beyond. My knowledge of programming mathematical algorithms and the role of mathematics in earth science has expanded immeasurably through my interactions with them. Moreover, they were warm and welcoming to me at NASA, and helped me find projects suitable for my skills. I am very thankful for their influence on my graduate career.

I would also like to thank the staff in the mathematics department, for their help in navigating the bureaucratic nightmares of graduate school. In particular, thank you to Celeste Regalado for helping me through the degree requirements, especially my language exam, and to Linette Berry for correcting my errors in scheduling seminar talks.

During my time at Maryland, I have collaborated with several colleagues on research papers. I would like to thank Dr. Timothy Doster, Dr. Benjamin Manning, Kevin Stubbs, and Daniel Wienberg for the opportunity to work together, and for their camaraderie.

I have had many wonderful friends in the mathematics department. My current and former roommates Chae Clark, David Darmon, Stefan Doboszczak, Ryan Hunter, and Matthew Guay have suffered my various eccentricities and have been loyal friends to share this long experience with. I've had plenty of good times with Matt Begué, Rob Maschal, Tim Mercurse, Joe Paulson, Jake Ralston, Arijit Sehanobish, and Hisham Talukdar. I'd also like to acknowledge a few graduate student elders, Dr. Travis Andrews, Steve Balady, Gokhan Civan, Dr. Alex Cloninger, Karamatou Djima, Dr. Mark Lai, and Paul Koprowski, for their wisdom and friend-

ship.

My most frequent sounding board for serious decisions, and my most trusted confidant, is my brother William. He gave me invaluable insights into the big picture during these past four years, which are at least as important as advice on the minutae of research. He is a very wise younger brother.

Finally, I acknowledge that without the support and guidance of my parents, Lisa and Michael, I wouldn't have written this thesis. I wouldn't have made it past seventh grade algebra if they hadn't insisted on doing *every question correctly.* Their discipline has shaped me. Whenever I was nervous about decisions with regard to research or whether to see the program out, they listened and gave me the most honest feedback. They always had faith in my ability to choose the correct path for myself. This has meant a lot.

Table of Contents

List of Abbreviations

$\mathbb{R}$	the set of real numbers
$\mathbb{C}$	the set of complex numbers
$\mathbb{Z}$	the set of integers
S^d	the unit sphere in $\mathbb{R}^{d+1}$
$\mathbb{T}^d$	the torus in $\mathbb{R}^d$
$\hat{f}$	the Fourier transform of a function f
$f^\vee$	the inverse Fourier transform of a function f
$\widehat{\mathcal{G}}$	the dual group of the group $\mathcal{G}$
$[A, B] = AB - BA$	the commutator of two operators A, B
A^{-1}	the inverse of an operator A
$\mathrm{dom}(A)$	the domain of an operator A
$\mathcal{S}(\mathbb{R}^d)$	the space of Schwartz functions on $\mathbb{R}^d$
$i = \sqrt{-1}$	the imaginary unit
$\overline{z}$	the complex conjugate of $z \in \mathbb{C}$
$\langle f, g \rangle_{\mathcal{H}}$	the inner product of $f, g \in \mathcal{H}$, for a Hilbert space $\mathcal{H}$
$\| \cdot \|_{\mathcal{H}}$	the norm in a Hilbert space $\mathcal{H}$
$T_t(f)(x) = f(x - t)$	the translation operator acting on a function $f : \mathbb{R}^d \to \mathbb{C},\ t \in \mathbb{R}^d$
$M_m(f)(x) = e^{-2\pi i \langle m, x \rangle} f(x)$	the modulation operator acting on a function $f : \mathbb{R}^d \to \mathbb{X},\ m \in \mathbb{R}^d$
$D_a(f)(x) = a^{-\frac{1}{2}} f(a^{-1}x)$	the dilation operator acting on a function $f : \mathbb{R} \to \mathbb{C},\ a \in \mathbb{R} \setminus \{0\}$
$D_A(f)(x) = \lvert \det A \rvert^{-\frac{d}{2}} f(A^{-1}x)$	the dilation operator acting on a function $f : \mathbb{R}^d \to \mathbb{C},\ A \in GL_d(\mathbb{R})$
χ_E	the indicator function on the measurable set E
$\mathcal{C}(X)$	the space of continuous functions on the space X
$GL_d(\mathbb{R})$	the space of $d \times d$ real invertible matrices
NASA	National Aeronautics and Space Administration

Chapter 1: Summary of Results

This thesis presents new results in pure and applied harmonic analysis. Its focus is on anistropic generalizations of the Fourier transform and their application to algorithms for the analysis of remote sensing data. Anisotropic methods emphasize directionality, and the harmonic analysis detailed in this thesis focuses on representation systems that incorporate some amount of directional content. In Chapter 2, a survey of classical, isotropic harmonic analysis is presented, as well as some relevant material on shearlets, which have risen to preeminence amongst anisotropic transform methods. A survey of the theory of optimal dictionary representations is also presented, to give some indication of how anisotropic methods are theoretically optimal for certain signal classes. A brief survey of the theory of wavelet packets is also developed.

Chapter 3 charts the development of directional Gabor systems, as pioneered by Grafakos and Sansing. Their method is essentially continuous, and did not generate discrete frames. We present several results suggesting reasonable spaces of functions which such a discrete frame could represent, before stating sufficient conditions for the generation of such a frame. This is pursued by adapting classical methods of Hernandez, Labate and Weiss to the anisotropic setting. A construction

of a basic directional Gabor frame is given, which, along with a sufficient condition, provides a general method for constructing discrete directional Gabor frames. The theory presented in this chapter resolves some open questions regarding directional Gabor theory, and brings the subject closer to efficient numerical implementation, by giving specific examples of discrete directional Gabor frames.

After developing this theory for directional Gabor theory, we delve into the use of anisotropic systems for numerical analysis of data. Isotropic transform methods, such as Fourier techniques and wavelets, are classical and pervasive in signal and image processing. The power to decompose a complicated signal into simpler atoms, along with an inversion tool to reconstitute the original signal, has long been understood as necessary for efficient transmission and processing of data. In the particular case of image data, most signals of interest contain strong directional features, in the form of natural and man-made edges. The efficient representation of these features via anisotropic harmonic analysis allows for more robust and accurate algorithms, and more effective integration of disparate data sources. The remaining chapters of this thesis exploit this theory to develop several efficient algorithms for the integration of image data.

The field of remote sensing provides a useful class of image data to analyze. The variety of features displayed, along with the multitude of disparate data types collected, makes the class of remotely sensed images flexible and challenging for the development of novel algorithms. Chapters 4,5, and 6 develop algorithms pertaining to remotely sensed images. Image registration and superresolution using shearlet methods are detailed in Chapters 4 and 5, respectively, and an image fusion algo-

rithm using wavelet packets is the content of Chapter 6. In each chapter, heuristic arguments are deployed to justify the idea of using harmonic analytic methods, and competitive numerical results are shown to confirm their efficacy.

Chapter 2: Mathematical Preliminaries

2.1 The Classical Fourier Transform

The Fourier transform is a classical object of harmonic analysis. The original use of Fourier methods in the natural sciences were in Lagrange's study of partial differential equations modeling string vibration [10]. Since those early days, the Fourier transform has offered engineers and physicists a powerful computational tool, and mathematicians a rich and subtle subject of study.

The Fourier transform is naturally defined for functions defined on $L^1(\mathbb{R}^d)$. It produces a function defined on the dual group of $\mathbb{R}^d$, which is denoted $\widehat{\mathbb{R}^d}$. It can be shown that $\widehat{\mathbb{R}^d} = \widehat{\mathbb{R}}^d \cong \mathbb{R}^d$ [91].

Definition 2.1.1. *The Fourier transform of $f \in L^1(\mathbb{R}^d)$ is the function $\hat{f} \in L^\infty(\widehat{\mathbb{R}}^d) \cap \mathcal{C}(\widehat{\mathbb{R}}^d)$ given by the formula:*

$$\hat{f}(\gamma) := \int_{\mathbb{R}^d} f(x) e^{-2\pi i \langle x, \gamma \rangle} dx.$$

When considering the Fourier transform in a physical or engineering context, f is often understood to be a signal. We shall take this approach in later chapters, where f is understood to be an idealized image signal.

Intuitively, the value of $\hat{f}(\gamma)$ is the extent to which the frequency γ is found in

f. As a heuristic, consider the case $g(x) = e^{2\pi i\langle\gamma_0,x\rangle}$ for a fixed γ_0. Then extending the above definition in a *distributional sense* [4], we would find $\hat{g}(\gamma) = \delta_{\gamma_0}(\gamma)$ the Dirac delta centered at γ_0. This makes sense given our intuitive understanding, since $g(x)$ would be a "pure frequency", so its Fourier transform should be perfectly localized at this frequency.

Note that the Fourier transform may be extended to $f \in L^2(\mathbb{R}^d)$, in a classical manner that does not appeal to modern distribution theory:

Theorem 2.1.2. *(Plancherel) [4] There is a unique linear bijection* $\mathcal{F} : L^2(\mathbb{R}^d) \to L^2(\widehat{\mathbb{R}}^d)$ *such that:*

1. $\forall f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d),\ \forall \gamma \in \widehat{\mathbb{R}}^d, \hat{f}(\gamma) = \mathcal{F}f(\gamma)$.

2. $\forall f \in L^2(\mathbb{R}^d),\ \|f\|_{L^2(\mathbb{R}^d)} = \|\mathcal{F}f\|_{L^2(\widehat{\mathbb{R}^d})}$.

3. $\forall f \in L^2(\mathbb{R}^d),\ \exists\{f_n\}_{n=0}^{\infty} \subset L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ *for which*

$$\lim_{n\to\infty} \|f_n - f\|_{L^2(\mathbb{R}^d)} = 0,\ \lim_{n\to\infty} \|\hat{f}_n - \mathcal{F}f\|_{L^2(\widehat{\mathbb{R}}^d)} = 0.$$

The Fourier transform, while useful for a variety of problems, suffers from several deficiencies. Two of them are of particular interest, and we shall address them in this chapter:

1. The Fourier transform gives information about the frequency content of a signal $f \in L^1(\mathbb{R}^d)$ with respect to *stationary* signals. These stationary signals are the complex exponentials $e_\gamma(x) := e^{-2\pi i\langle\gamma,x\rangle}$.

2. The Fourier transform is *isotropic*, meaning it has no sensitivity to direction.

The issue of stationary signals can be understood intuitively by considering the behavior of the complex sinusoidal functions $e_\gamma(x) = e^{-2\pi i\langle\gamma,x\rangle}$; see Figure 2.1.

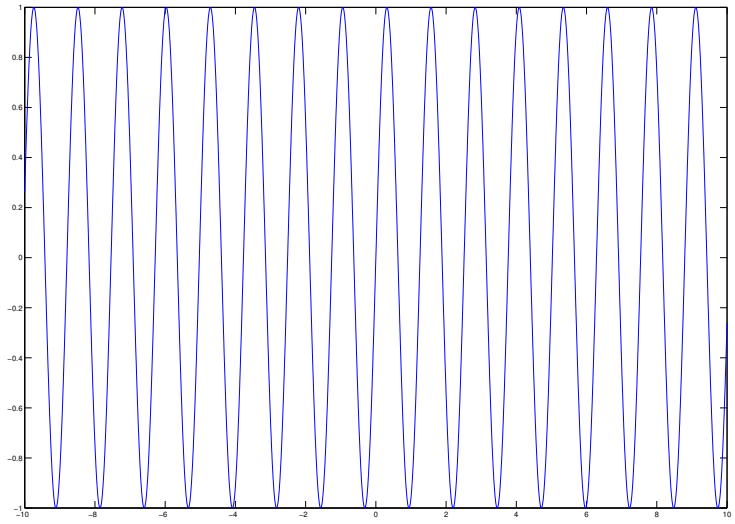


Figure 2.1: *Plot of the imaginary parts of the complex exponential* $f(x) = e^{5ix}, x \in \mathbb{R}$. *Notice the behavior of the signal is stationary; it does not change as one moves along the* x*-axis.*

These are 1-periodic, and are not localized in space at all. Indeed, their periodicity ensures there is no distinction between their local and global behavior. As such, local features of a function $f \in L^1(\mathbb{R}^d)$ needn't be captured by the coefficient $\langle f(x), e^{-2\pi i\langle\gamma,x\rangle}\rangle$. As mentioned above, the distributional Fourier transform of $e_\gamma(x) = e^{-2\pi i\langle\gamma,x\rangle}$ is $\delta_0(\gamma)$, which is perfectly localized as a distribution on $\widehat{\mathbb{R}}^d$. This dichotomy of awful localization in space and excellent localization in frequency is of great significance in the design of efficient waveforms for signal processing, and is a classical illustration of the Heisenberg uncertainty principle. This result is often stated in elementary physics courses in the context of quantum mechanics:

"The position and momentum of a particle cannot be known simultaneously"

This physical notion can be formalized by mathematically describing quantum mechanics; we shall not investigate this. We are interested in a mathematical formulation of the uncertainty principle: intuitively, a signal cannot be extremely well-localized in both the time domain and the frequency domain. Indeed, there is a precise limitation on how localized it can be in both.

Theorem 2.1.3. *(Heisenberg Uncertainty Principle) Let $f \in L^2(\mathbb{R})$, $a, b \in \mathbb{R}$. Then:*

$$\left(\int_{\mathbb{R}} (x-a)^2 |f(x)|^2 dx\right)^{\frac{1}{2}} \left(\int_{\widehat{\mathbb{R}}} (\gamma-b)^2 |\hat{f}(\gamma)|^2 d\gamma\right)^{\frac{1}{2}} \geq \frac{1}{4\pi} \|f\|_2^2.$$

Equality holds if and only if f is a scalar multiple of $e^{2\pi i b(x-a)} \cdot e^{\frac{-\pi(x-a)^2}{c}}$, for some $a, b \in \mathbb{R}$, some $c > 0$.

In the above theorem, the first factor is a measure of localization in space around the point a; the second factor is a measure of localization in frequency about the point b. Note that this result generalizes not only to functions in $L^2(\mathbb{R}^d)$ for $d > 1$, but to general Hilbert spaces [49]:

Theorem 2.1.4. *Let A, B be self-adjoint operators on a Hilbert space $\mathcal{H}$, not necessarily bounded. Then for all $f \in dom(AB) \cap dom(BA)$ and for all $a, b \in \mathbb{R}$, we have the following:*

$$\|(A-a)f\|_{\mathcal{H}} \cdot \|(B-b)f\|_{\mathcal{H}} \geq \frac{1}{2} \left|\langle [A,B]f, f\rangle_{\mathcal{H}}\right|.$$

Equality holds if and only if $(A-a)f = ic(B-b)f$, for some $c \in \mathbb{R}$.

As Theorem 2.1.3 illustrates, the Gaussian has optimal time-frequency localization. It is worth nothing that a classic calculation [4] shows that the Gaussian is invariant under the Fourier transform.

A useful abstraction in Fourier analysis is the *time-frequency plane*, namely the Cartesian plane with the time/space variable on the x-axis and the frequency variable on the y-axis. With this in mind, the Gaussian has, in some sense, *minimal support in the time-frequency plane*, as evidenced by its minimization of the Heisenberg uncertainty principle. In the quest to move beyond stationary Fourier bases, it is natural to ask whether functions that have good localization in *both time and frequency* can generate representation systems. A first question in this regard is whether translates and modulates of the Gaussian span $L^2(\mathbb{R}^d)$; this question was first addressed by Nobel Laureate Dennis Gabor [45], and led to the study of Gabor theory and time-frequency analysis [49].

2.2 Gabor Systems

With this analysis in terms of uncertainty principles in mind, we shall replace the *stationary harmonics* $e^{-2\pi i\langle\gamma,x\rangle}$ with *windowed harmonics* $g(x)e^{-2\pi i\langle\gamma,x\rangle}$ for some suitable window $g(x)$.

Dennis Gabor, 1971 Nobel laureate in Physics, was interested in the case where the window function g is assumed to be Gaussian, but we shall consider a general $g \in L^2(\mathbb{R}^d)$. Noting the family of functions

$$\{e^{-2\pi i\langle m,x\rangle}\}_{m\in\mathbb{Z}^d} \tag{2.1}$$

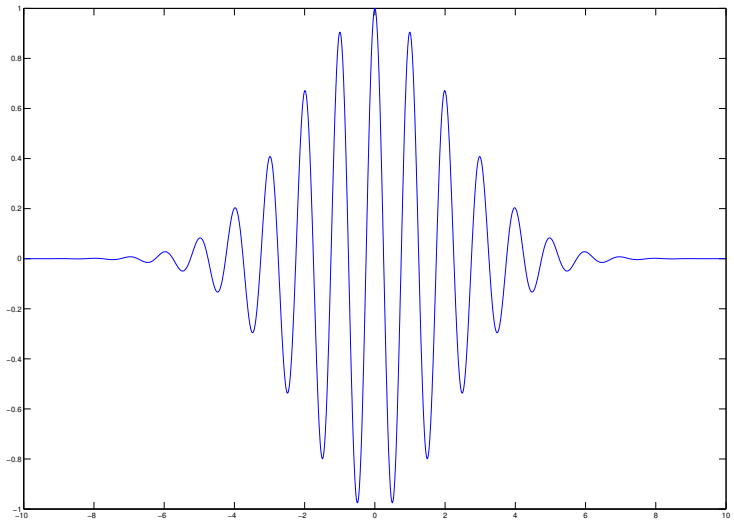


Figure 2.2: *Plot of the real part of the windowed harmonic function* $f(x) = e^{2\pi i x} e^{-\frac{x^2}{10}}$. *Notice that the signal is non-stationary; its behavior changes as one moves along the* x*-axis.*

forms an orthonormal basis for $L^2(\mathbb{T}^d)$ (via Fourier series for $\mathbb{T}^d$), it is of interest to know when $g \in L^2(\mathbb{R})$ is such that

$$\{g(n-x)e^{-2\pi i\langle m,x\rangle}\}_{m,n\in\mathbb{Z}^d} \tag{2.2}$$

is a basis for $L^2(\mathbb{R}^d)$. In order to address this question satisfactorily, we shall consider a broader class families generated from g, and also a weaker notion of representing family than orthonormal basis.

Definition 2.2.1. *Let* $g \in L^2(\mathbb{R}^d)$. *Let* $\alpha, \beta > 0$. *The family of functions*

$$\mathcal{G}(g,\alpha,\beta) := \{g(x-n\alpha)e^{-2\beta\pi i\langle m,x\rangle}\}_{m,n\in\mathbb{Z}^d} \tag{2.3}$$

is a regular Gabor system. *Let* $\Lambda \subset \mathbb{R}^{2d}$ *be discrete. The family of functions*

$$\mathcal{G}(g, \Lambda) := \{g(x - \lambda)e^{-2\pi i\langle \gamma, x\rangle}\}_{(\gamma,\lambda)\in\Lambda} \tag{2.4}$$

is an (irregular) Gabor system.

In the study of regular Gabor systems, the parameters α, β play a crucial role, as we shall see. Note that regular Gabor systems (2.3) are a special case of irregular Gabor systems (2.4), corresponding to $\Lambda = \alpha\mathbb{Z}^d \times \beta\mathbb{Z}^d$.

The notion of orthonormal basis is too restrictive for many applications, including transmission of signals. Consider the problem of signal transmission in the following paradigm: a signal f is to be encoded into coefficients $\{c(f)_n\}_{n\in\mathbb{Z}}$. This countable set of coefficients is transmitted to a receiver, where they are used to recover the signal f.

In the case that $c(f)_n = \langle f, \phi_n\rangle$ for some orthonormal basis $\{\phi_n\}_{n\in\mathbb{Z}}$, the corruption of even a single coefficient by noise can be problematic. This is because none of the other coefficients can replace the information lost in the noisy corruption. The consequent reconstruction of f could be terrible. Thus, we consider a more flexible construction, one that allows for a degree of *redundancy* in signal representation. This notion is due originally to Duffin and Schaeffer, who introduced it in 1952 in the course of their study of non-harmonic Fourier series [37].

Definition 2.2.2. *Let* $\mathcal{H}$ *be a Hilbert space. A discrete set* $\{\phi_i\}_{i\in I} \subset \mathcal{H}$ *is called a* (discrete) frame for $\mathcal{H}$ *if there exist constants* $0 < A \leq B < \infty$ *such that:*

$$\forall f \in \mathcal{H},\ A\|f\|_{\mathcal{H}}^2 \leq \sum_{i\in I} |\langle f, \phi_i\rangle_{\mathcal{H}}|^2 \leq B\|f\|_{\mathcal{H}}^2.$$

The optimal choices of A, B *are the* frame bounds. *A frame is said to be* tight *if* $A = B$. *It is* Parseval *if* $A = B = 1$.

Of course, all orthonormal bases are Parseval frames. A frame for $\mathcal{H}$ offers a system for representing elements of $\mathcal{H}$ that is weaker than an orthonormal basis. More precisely, frames represent $f \in \mathcal{H}$ *redundantly*, which makes frames useful for many signal processing applications. The theory of frames is rich and well-studied [18], [59], [15] but we shall not make use of much of the general theory. We shall consider the case $\mathcal{H} = L^2(\mathbb{R}^d)$.

Determining when a Gabor system forms a frame is a difficult problem; a survey of the history of this problem may be found in [58]. We consider a few classical results in the field, to give a flavor for what about g, α, β are relevant to the problem. A well-known result provides a complete characterization of regular, one-dimensional Gabor systems for certain window functions g [31].

Theorem 2.2.3. *Let* $g \in L^\infty(\mathbb{R}^d)$ *be supported on* $[0, L]^d,\ L > 0$. *Then* $\mathcal{G}(g, \alpha, \beta)$ *is a frame with frame bounds* $\beta^{-d}a,\ \beta^{-d}b$ *if and only if*

$$a \leq \sum_{k \in \mathbb{Z}^d} |g(x - \alpha k)|^2 \leq b \text{ almost everywhere.}$$

A more precise result requires some additional notation.

Definition 2.2.4. *Let* $g \in L^\infty(\mathbb{R}^d)$. *We say* g *is an element of the* Wiener space, *denoted* $W = W(\mathbb{R}^d)$, *if*

$$\|g\|_W := \sum_{n \in \mathbb{Z}^d} \text{ess sup}_{x \in [0,1]^d} |g(x + n)| < \infty.$$

Naturally enough, the Wiener space was introduced by the great Norbert Wiener, in the context of his study of *Tauberian theorems* [106]. Intuitively, W may be thought of as the space of functions that are *locally bounded* but *globally* ℓ^1. The Wiener space is a useful class of functions to consider as windows for Gabor systems. To give conditions on which $g \in W$ produce regular Gabor systems for certain parameters α, β, we must consider *correlation functions associated with* g.

Definition 2.2.5. *Let g, α, β be as above. The n^{th}* correlation function *corresponding to these choices of g, α, β is*

$$G_n(x) = G_n^{(\alpha,\beta)} := \sum_{k\in\mathbb{Z}^d} \overline{g\left(x - \frac{n}{\beta} - \alpha k\right)} g(x - \alpha k).$$

These correlation functions yield a sufficient condition for $\mathcal{G}(g, \alpha, \beta)$ to be a frame [103]:

Theorem 2.2.6. *Suppose $g \in W(\mathbb{R}^d)$, and that $\alpha > 0$ is such that there exist constants $0 < a \leq b < \infty$ such that:*

$$a \leq \sum_{k\in\mathbb{Z}^d} |g(x - \alpha k)|^2 \leq b < \infty \text{ almost everywhere.}$$

Then there exists a number $\beta_0 = \beta_0(\alpha) > 0$ such that $\mathcal{G}(g, \alpha, \beta)$ is a frame for all $\beta \leq \beta_0$. More precisely, if $\beta_0 > 0$ is chosen such that:

$$\sum_{n\in\mathbb{Z}^d\setminus\{0\}} \|G_n^{(\alpha,\beta_0)}\|_\infty < \text{ess inf}_{x\in\mathbb{R}^d}|G_0(x)| = \text{ess inf}_{x\in\mathbb{R}^d} \sum_{k\in\mathbb{Z}^d} |g(x - \alpha k)|^2,$$

then $\mathcal{G}(g, \alpha, \beta)$ is a frame for all $\beta \leq \beta_0$, with frame bounds

$$A = \beta^{-d}\left(a - \sum_{n\in\mathbb{Z}^d\setminus\{0\}} \|G_n^{(\alpha,\beta)}\|_\infty\right),$$

$$B = \beta^{-d} \sum_{n\in\mathbb{Z}^d} \|G_n^{(\alpha,\beta)}\|_\infty.$$

The theory of Gabor systems extends well beyond this, including a theory of *irregular Gabor systems.* This theory is quite subtle, and notions of density are of paramount importance; see [58] for an overview. Multiple window functions $g_1, g_2, ..., g_N$ may also be considered [113]. There is also a continuous theory of Gabor systems, which review briefly.

Definition 2.2.7. *Let $g \in L^2(\mathbb{R}^d)$. The* short-time Fourier transform *of a function $f : \mathbb{R}^d \to \mathbb{C}$ with respect the window function g is the function $V_g(f) : \mathbb{R}^{2d} \to \mathbb{C}$ given by:*

$$V_g(f)(t, m) := \langle f, M_{-m}T_t g\rangle = \int_{\mathbb{R}^d} f(x)\overline{g(x-t)}e^{-2\pi i\langle m,x\rangle}dx,\ t, m \in \mathbb{R}^d.$$

The short-time Fourier transform is also called a "sliding window Fourier transform," [49] or "voice transform," for its connections to speech processing [48]. The window function g may be chosen with desired regularity and localization properties.

In the case of discrete Gabor frames, we are interested in recovering the original signal f from its Gabor coefficients. In the case of the short-time Fourier transform, the original function, f, is recovered from $V_g(f)$ by integrating against translations and modulations of any $\psi \in L^2(\mathbb{R}^d)$ such that $\langle \psi, g\rangle \neq 0$:

$$f = \frac{1}{\langle \psi, g\rangle}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} V_g(f)(t, m)\psi^{m,t}dm\ dt,\ m, t \in \mathbb{R}^d, \tag{2.5}$$

where $\psi^{m,t}(x) := e^{2\pi i\langle m,(x-t)\rangle}\psi(x-t) = T_t M_{-m}\psi(x)$ and convergence is pointwise almost everywhere [49].

The short-time Fourier transform is a continuous frame:

Definition 2.2.8. *Let $\mathcal{H}$ be a separable Hilbert space, X a locally compact Hausdorff space equipped with a positive Radon measure μ such that supp(μ) $= X$. A family*

$F = \{\psi_x\}_{x \in X}$ *is a* continuous frame with respect to μ *for* $\mathcal{H}$ *if there exist constants* $0 < A \leq B < \infty$ *such that:*

$$\forall f \in \mathcal{H},\ A\|f\|_{\mathcal{H}}^2 \leq \int_X |\langle f, \psi_x \rangle_{\mathcal{H}}|^2 d\mu(x) \leq B\|f\|_{\mathcal{H}}^2.$$

Discrete frames are contained in the broad class of continuous frames, by taking μ to be a counting measure. This definition generalizes Definition 2.2.2 by allowing for a continuous indexing of the frame elements. In the study of continuous frames, the *discretization problem arises naturally:* when can a continuous frame be discretized to acquire a discrete frame? That is, when can a continuous indexing set be replaced by a discrete indexing set, and the continuous measure with a counting measure? The most prominent abstract approach to the discretization problem involves the coorbit space theory of Feichtinger and Gröchenig [41], [42]. This method considers the representations induced by the group structure of the indexing set X. Thus, coorbit space theory is most effective when the continuous frame is parametrized by a group. This is the case for shearlet systems, as we shall see in Section 2.4.

Beyond being of mathematical interest, Gabor systems have proven useful for a variety of applications, including denoising [79], speech recognition [77] and image analysis [75]. However, a Gabor system is fundamentally isotropic. The family is generated from a window function g, through modulations and translations, without emphasis on directional content. The problem of incorporating directionality into discrete Gabor frames is studied in Chapter 3.

2.3 Wavelets

Another approach to localized representation systems is to consider systems generated by *dilations* and *translations* of a base function. The study of such functions and the induced bases and frames is the content of *wavelet theory*. Wavelets revolutionized harmonic analysis and image processing. Their rich theory opened new avenues for constructing bases and frames, and their numerical implementation generated efficient methods for image compression [16], segmentation [99], fusion [88], and registration [108].

Definition 2.3.1. *Let* $\psi \in L^2(\mathbb{R})$. *The* (one-dimensional) discrete wavelet system *generated by* ψ *is the family of functions*

$$\{\psi_{j,m}(x) := D_{2^{-j}} T_m \psi(x) = 2^{\frac{j}{2}} \psi(2^j x - m)\}_{j,m \in \mathbb{Z}}.$$

The associated Discrete Wavelet Transform *is the mapping*

$$f \mapsto \{\langle f, \psi_{j,m} \rangle\}_{j,m \in \mathbb{Z}}, \ f \in L^2(\mathbb{R}).$$

It is possible to choose ψ appropriately so that the associated discrete wavelet system is an orthonormal basis or tight frame. This can be done through the multiresolution analysis (MRA) technique, pioneered by Yves Meyer and Stéphane Mallat [82], or through direct constructions [31]. Wavelet systems have a greater degree of localization than simple Fourier series. In particular, ψ may be chosen to be compactly supported and to have a high degree of smoothness.

For the problem of image analysis, the one-dimensional wavelet systems defined above are insufficient; image data is generally two-dimensional, but is sometimes of

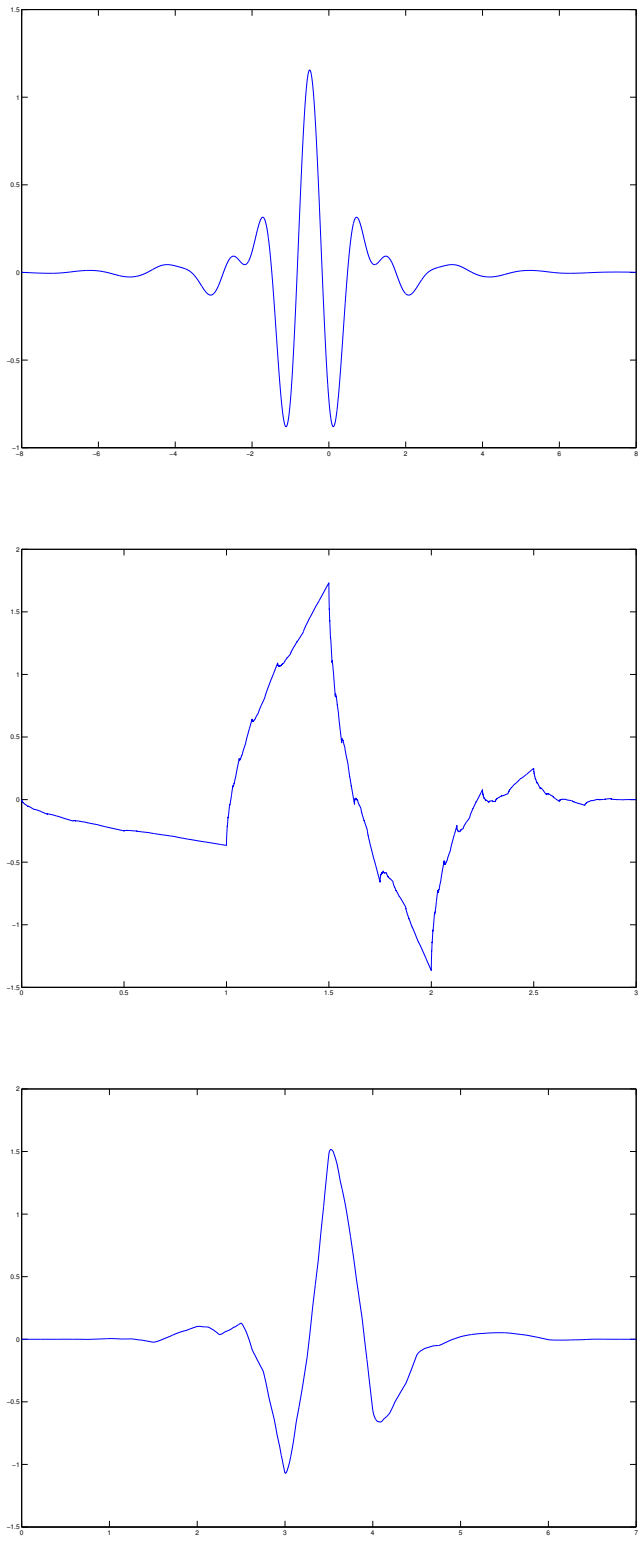


Figure 2.3: *Plots of some well-known wavelet functions ψ: Meyer, Daubechies 4, Symmetric 4 (top to bottom).*

even higher dimension, in the case of lidar data, for example. A generalization to higher dimensional spaces is required. To construct wavelet systems in $\mathbb{R}^d,\ d > 1$, certain algebraic notions are required.

We shall begin with a survey of the continuous theory of wavelet systems, before developing a theory of discrete wavelet orthonormal bases. Although our interests ultimately lie in discrete wavelet systems, the continuous theory provides some helpful motivation.

2.3.1 Continuous Wavelet Systems

Definition 2.3.2. *Let $\psi \in L^2(\mathbb{R}^d)$, $G \subset GL_d(\mathbb{R})$. The* continuous affine system *associated with G, ψ is the family of functions*

$$\{\psi_{M,t}(x) := T_t D_{M^{-1}} \psi(x) = |\det M|^{\frac{1}{2}} \psi(M(x-t))\}_{(M,t)\in G\times\mathbb{R}^d}.$$

This system is denoted $\mathcal{A}_d = \mathcal{A}_d(G,\psi)$.

We wish to understand conditions on ψ such that the induced continuous affine system is a continuous frame, or if an arbitrary $f \in L^2(\mathbb{R}^d)$ can be recovered from the coefficients $\{\langle f, \psi_{M,t}\rangle\}_{(M,t)\in G\times\mathbb{R}^d}$. To this end, we endow the indexing set of this system with a group structure:

Definition 2.3.3. *The* affine group on $\mathbb{R}^d$, *denoted $\mathcal{A}_d$, is the group structure on $GL_d(\mathbb{R}) \times \mathbb{R}^d$ with multiplication given by:*

$$(M_1, t_1) \cdot_{\mathcal{A}_d} (M_2, t_2) := (M_1 M_2, t_1 + M_1 t_2).$$

Reconstruction from the continuous affine system is possible, so long as a certain condition on ψ is satisfied. The notion of an *admissibility condition* for ψ is pervasive in wavelet theory, and will be important in shearlet theory as well.

Definition 2.3.4. *Let $\psi \in L^2(\mathbb{R}^d)$, and let $d\mu$ be a left-invariant Haar measure on $G \subset GL_d(\mathbb{R})$. We say ψ is an* admissible wavelet *if:*

$$\forall \gamma \in \mathbb{R}^d, \ \int_G |\hat{\psi}(M^T\gamma)|^2 |\det M| d\mu(M) = 1. \tag{2.6}$$

Theorem 2.3.5. *Let $G \subset GL_d(\mathbb{R})$. Let $d\mu$ be a left-invariant Haar measure on G, and $d\lambda$ a left-invariant Haar measure on $\mathcal{A}_d$. Furthermore, suppose that $\psi \in L^2(\mathbb{R}^d)$ is admissible. Then for all $f \in L^2(\mathbb{R}^d)$, we have the following weak convergence representation of f in terms of its coefficients under the continuous affine system:*

$$f = \int_{\mathcal{A}_d} \langle f, \psi_{M,t} \rangle \psi_{M,t} d\lambda(M,t).$$

Taking the special case $G := \{aI_d\}_{a>0}$ where I_d is the $d \times d$ identity matrix, corresponding to the case of *isotropic dilations*, the admissibility condition for ψ simplifies to:

$$\int_0^\infty |\hat{\psi}(a\gamma)|^2 \frac{da}{a} = 1.$$

The associated affine system can, in the case $d = 1$, be discretized to produce discrete wavelet systems. The same notion is valid for dimensions $d > 1$, to produce discrete wavelet systems in any dimension. We now examine such discrete wavelet systems.

2.3.2 Discrete Wavelet Systems

For simplicity, we consider in this subsection only the one-dimensional case. With some work, the theory may be generalized to arbitrary Euclidean space [107].

The theory in Section 2.3.1 is continuous, and thus ill-suited to applications. We seek instead a discrete theory for wavelets. The problem at the heart of this theory is the construction of discrete wavelet orthonormal bases and frames, i.e. families

$$\{\psi_{m,n}(x) = 2^{\frac{m}{2}}\psi(2^m x - n)\}_{m,n\in\mathbb{Z}},$$

which form orthonormal bases or frames, for some *wavelet* function $\psi \in L^2(\mathbb{R})$; this set-up is easily generalized to high dimensions [107].

One of the key results related to discrete wavelet systems is the MRA theorem of Meyer and Mallat [82]. We shall state it shortly, after introducing the crucial machinery.

Definition 2.3.6. *A* multiresolution analysis *on* $\mathbb{R}$ *is a sequence of subspaces* $\{V_j\}_{j\in\mathbb{Z}}$ *together with a* scaling function $\phi \in L^2(\mathbb{R})$ *with the following properties:*

1. $V_j \subset V_{j+1},\ \forall j \in \mathbb{Z}$.
2. $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$.
3. $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$.
4. $f(x) \in V_0$ *if and only if* $D_{2^j} f(x) = 2^{\frac{j}{2}} f(2^j x) \in V_j$.
5. *The set* $\{T_n\phi\}_{n\in\mathbb{Z}}$ *is an orthonormal basis for* V_0.

The following lemma states the *two-scale property* of ϕ, which is crucial for the constructions that follow.

Lemma 2.3.7. *Suppose ϕ is a scaling function for an MRA. Then there exists a sequence of coefficients $\{h_0[n]\}_{n\in\mathbb{Z}}$ such that:*

$$\phi(x) = \sum_{n\in\mathbb{Z}} h_0[n]\phi(2x-n).$$

A multiresolution analysis is often abbreviated MRA. Theorem 2.3.8 states how a wavelet orthonormal basis is generated from an MRA.

Theorem 2.3.8. *Let $\{V_j\}_{j\in\mathbb{Z}}$ be an MRA with scaling function ϕ. Define a filter $h_1[n] = (-1)^n\overline{h_0[1-n]}$, and a function $\psi \in L^2(\mathbb{R})$ by the formula*

$$\psi(x) = \sum_{n\in\mathbb{Z}} h_1[n]\phi(2x-n).$$

Then the family

$$\{\psi_{m,n}(x) = \psi(2^m x - n)\}_{m,n\in\mathbb{Z}}$$

is an orthonormal wavelet basis for $L^2(\mathbb{R})$, i.e. a discrete wavelet systems that is also an orthonormal basis. Such a ψ is called a wavelet *function.*

The function ψ is thus a wavelet function. While there are other methods for constructing discrete wavelet systems, the MRA method has the benefit of generalizing from Euclidean space to general locally compact abelian groups and beyond [5]. Moreover, it provides a setting to define *wavelet packets*; see Section 2.6.

2.3.3 Numerical Wavelet Implementations

Suppose $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is an orthonormal wavelet basis. Then any $f \in L^2(\mathbb{R})$ may be written as:

$$f = \sum_{m,n\in\mathbb{Z}} \langle f, \psi_{m,n}\rangle \psi_{m,n}. \tag{2.7}$$

Several schemes have been proposed to discretize (2.7). In this thesis, Simoncelli filters and spline wavelet decompositions are used in Chapter 4. These have the benefit of translation invariance and fast implementation in the C programming language. Additional details will be given in Chapter 4.

2.4 Shearlets

Wavelets have a rich mathematical theory, and have numerous practical uses. One weakness they possess, however, is their *isotropy*. More precisely, wavelets are not able to represent functions with curve-like discontinuities optimally. This inhibits their performance in image processing, since many images possess curve-like discontinuities, which are better known as edges. A detailed discussion of the inefficiency of wavelets for representing anisotropic signals is in Section 2.5.

The idea of generalizing wavelets to be anisotropic has yielded several representation systems with rich theory, for example the contourlets of Do and Vetterli [33], the curvelets of Donoho and Candès [13], and shearlets, pioneered by Labate, Kutyniok, King, and their collaborators [70], [39], [51], [65], [26]. Of these methods, shearlet systems have the benefit of requiring only a single generating function, and

of admitting fast numerical implementations. For these reasons, they are well-suited for applications, which are discussed in Chapters 4 and 5 of this thesis. We note that the shearlet construction to be presented is a specific case of the more general theory of composite wavelets, developed by Weiss, Wilson, Manning, and their collaborators [52], [83].

Shearlet systems generalize wavelets by introducing a directionally sensitive dilation term as follows.

Definition 2.4.1. *Let A_a, S_s be the matrices given by*

$$A_a := \begin{pmatrix} a & 0 \\ 0 & a^{\frac{1}{2}} \end{pmatrix}, \; S_s := \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}.$$

We shall first consider continuous shearlet systems, before developing a discrete theory. We consider shearlets in two dimensions, though they generalize to three [71].

2.4.1 Continuous Shearlet Systems

For $\psi \in L^2(\mathbb{R}^2)$, the *continuous shearlet system SH(ψ)* is the family of functions:

$$SH(\psi) := \{\psi_{a,s,t}(x) := T_t D_{A_a} D_{S_s} \psi(x) = a^{\frac{3}{4}} \psi(A_a^{-1} S_s^{-1} x - t)\}_{a>0, s\in\mathbb{R}, t\in\mathbb{R}^2}.$$

The matrix D_a is no longer isotropic; it emphasizes a particular direction. The *shearing matrix* S_s selects that direction. As $a \to 0$, the anisotropy of A_a will force a particular direction to be overemphasized to such a degree that only local information in that direction is captured by the corresponding shearlet coefficient.

By selecting s appropriately, one can tune a to capture precise information in a specified direction. This emphasis on directional information will prove useful in the image registration and superresolution algorithms presented in Chapters 4 and 5.

As in the case of wavelets, which were associated with the affine group, shearlets are associated with an algebraic object:

Definition 2.4.2. *The* shearlet group*, denoted* $\mathbb{S}$*, is the semi-direct product* $(\mathbb{R}^+ \times \mathbb{R}) \rtimes \mathbb{R}^2$ *with multiplication given by the rule*

$$(a_1, s_1, t_1) \cdot_{\mathbb{S}} (a_2, s_2, t_2) := (a_1 a_2, s_1 + s_2\sqrt{a_1}, t_1 + S_{s_1} A_{a_1} t_2).$$

We shall now write a continuous shearlet system in terms of a representation of the shearlet group. Let $\mathcal{U}(L^2(\mathbb{R}^2))$ be the group of unitary operators on $L^2(\mathbb{R}^2)$. Noting that $\frac{da}{a^3} dsdt$ is a left-invariant Haar measure on the shearlet group, we may define a unitary representation $\sigma : \mathbb{S} \to \mathcal{U}(L^2(\mathbb{R}^2))$ by:

$$\sigma(a, s, t)\psi := \psi_{a,s,t} = T_t D_{A_a} D_{S_s} \psi.$$

This representation may be made irreducible by extending it by symmetry to $(\mathbb{R}^* \times \mathbb{R}) \rtimes \mathbb{R}^2$. For the remainder of this chapter, we consider this irreducible representation.

With this notation, we see that:

$$SH(\psi) = \{\sigma(a, s, t)\psi\}_{(a,s,t) \in \mathbb{R}^* \times \mathbb{R} \times \mathbb{R}^2}.$$

This algebrically-induced structure of shearlet systems lends itself to the coorbit theory of Feichtinger and Gröchening [41], [29]. Indeed, we shall see shortly that $SH(\psi)$ forms a continuous frame. This gives one approach to discretizing continuous shearlet systems to acquire discrete frames [29].

The mapping that sends a signal to its shearlet coefficients is of great significance. Indeed, these coefficients allow the reconstruction of a signal from its shearlet information.

Definition 2.4.3. *For $\psi \in L^2(\mathbb{R}^2)$, the* continuous shearlet transform *of $f \in L^2(\mathbb{R}^2)$ is the mapping*

$$f \mapsto \{\langle f, \sigma(a,s,t)\psi\rangle\}_{(a,s,t)\in\mathbb{S}}.$$

As in the case of wavelets, a shearlet basis function must satisfy a certain admissibility condition in order for the corresponding continuous shearlet system to have a reproducing formula.

Definition 2.4.4. *A function $\psi \in L^2(\mathbb{R}^2)$ is* an admissible shearlet *if*

$$\int_{\widehat{\mathbb{R}}^2} \frac{|\hat{\psi}(\gamma_1,\gamma_2)|^2}{\gamma_1^2} d\gamma_2 d\gamma_1 < \infty. \tag{2.8}$$

Condition (2.8) should be compared to the corresponding wavelet admissibility condition (2.6). A canonical example of an admissible shearlet is a so-called *classical shearlet function*:

Definition 2.4.5. *Let $\psi \in L^2(\mathbb{R}^2)$ be defined by the formula*

$$\hat{\psi}(\gamma) = \hat{\psi}(\gamma_1,\gamma_2) := \hat{\psi}_1(\gamma_1)\hat{\psi}_2\left(\frac{\gamma_2}{\gamma_1}\right),$$

where ψ_1, ψ_2 satisfy the following conditions:

- $\psi_1 \in \mathcal{C}^\infty(\mathbb{R})$ *satisfies the* discrete Calderón condition*:*

$$\sum_{j\in\mathbb{Z}} |\hat{\psi}_1(2^{-j}\gamma)|^2 = 1, \text{ for almost every } \gamma \in \mathbb{R}.$$

- $supp(\hat{\psi}_1) \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$.

- $\psi_2 \in \mathcal{C}^\infty(\mathbb{R})$ *is a bump function, i.e. satisfies*

$$\sum_{k=-1}^{1} |\hat{\psi}_2(\gamma + k)|^2 = 1, \text{ for almost every } \gamma \in [-1, 1].$$

- $\text{supp}(\hat{\psi}_2) \subset [-1, 1]$.

Such a ψ *is called a* classical shearlet function.

As in our discussion of continuous frames, we are interested in when a general $f \in L^2(\mathbb{R}^2)$ can be reconstructed from its shearlet coefficients $\{\langle f, \psi_{a,s,t}\rangle\}_{(a,s,t)\in\mathbb{S}}$. A sufficient condition for ψ to generate a system from which f may be recovered from its shearlet coefficients is when the coefficient mapping

$$f \mapsto \{\langle f, \sigma(a, s, t)\psi\rangle\}_{(a,s,t)\in\mathbb{S}} \tag{2.9}$$

is an *isometry*, i.e.

$$\|f\|_2^2 = \sum_{(a,s,t)\in\mathbb{S}} |\langle f, \sigma(a, s, t)\psi\rangle|^2.$$

Necessary conditions for (2.9) to be an isometry may be formulated in terms of certain integrals [28]:

Theorem 2.4.6. *Let* $\psi \in L^2(\mathbb{R}^2)$ *be an admissible shearlet. Define*

$$C_\psi^+ = \int_0^\infty \int_{\mathbb{R}} \frac{|\hat{\psi}(\gamma_1, \gamma_2)|^2}{\gamma_1^2} d\gamma_2 d\gamma_1, \quad C_\psi^- = \int_{-\infty}^0 \int_{\mathbb{R}} \frac{|\hat{\psi}(\gamma_1, \gamma_2)|^2}{\gamma_1^2} d\gamma_2 d\gamma_1.$$

If $C_\psi^+ = C_\psi^- = 1$, *then (2.9) is an isometry.*

Thus, in certain cases, a continuous shearlet system forms a continuous frame.

2.4.2 Discrete Shearlet Systems

For applications purposes, continuous shearlets are insufficient; a discrete system is required. As mentioned, one approach to the discretization of continuous shearlet systems is coorbit theory. This is possible because shearlet systems form a square-integrable irreducible representation of $\mathbb{S}$. We proceed directly, however, to produce discrete shearlet systems [68].

Definition 2.4.7. *Let $\psi \in L^2(\mathbb{R}^2)$ and $\Lambda \subset \mathbb{S}$. A* regular discrete shearlet system *associated with ψ, denoted $SH(\psi)$, is the family of functions*

$$SH(\psi) := \{\psi_{j,k,m}(x) = 2^{\frac{3j}{4}}\psi(S_k A_{2^j} x - m)\}_{j,k\in\mathbb{Z},m\in\mathbb{Z}^2}.$$

An irregular discrete shearlet system *associated with ψ and Λ, denoted $SH(\psi,\Lambda)$, is the family of functions*

$$SH(\psi,\Lambda) := \{\psi_{a,s,t}(x) = a^{-\frac{3}{4}}\psi(A_a^{-1}S_s^{-1}(x-t))\}_{(a,s,t)\in\Lambda}.$$

Of course, regular discrete shearlet systems are a special case of irregular discrete shearlet systems. Indeed, taking

$$\Lambda = \{(2^{-j}, -k, S_{-k}A_{2^{-j}}m)\}_{j,k\in\mathbb{Z},m\in\mathbb{Z}^2}$$

yields a regular discrete shearlet system.

The value of classical shearlets is clear from the following result:

Theorem 2.4.8. *Let $\psi \in L^2(\mathbb{R}^2)$ be a classical shearlet. Then the regular discrete shearlet system associated with ψ, $SH(\psi)$, is a Parseval frame for $L^2(\mathbb{R}^2)$.*

Note that since a classical shearlet ψ is constructed to be well-localized, we conclude there exist well-localized discrete shearlet frames. Moreover, ψ may taken to be compactly supported [69]; this requires development different from classical shearlets, since classical shearlet functions are compactly supported in the frequency domain, and thus cannot be compactly supported in the time domain. The existence of highly regular shearlet frames provides a powerful tool for signal analysis, as signals can be decomposed into directionally sensitive shearlet coefficients, which can then be manipulated.

2.4.3 Numerical Implementations of Shearlets

In the past few years, several numerical shearlet implementations have been released. The algorithms used in Chapters 4 and 5 of this thesis make use of the Fast Shearlet Transform MATLAB library [54], which is available at `http://www.mathematik.uni-kl.de/imagepro/software/ffst/`.

2.5 Sparsity in Dictionaries

One of the greatest strengths of harmonic analysis for data processing is in its ability to convert data from one domain to another. In the new domain, the data may take a particularly convenient representation. In particular, data is often *sparse* in the transformed domain, meaning it is represented with respect to a basis or collection of frame elements with coefficients that are mostly zero. Computation and analysis can then be performed in the transformed domain, where the data is

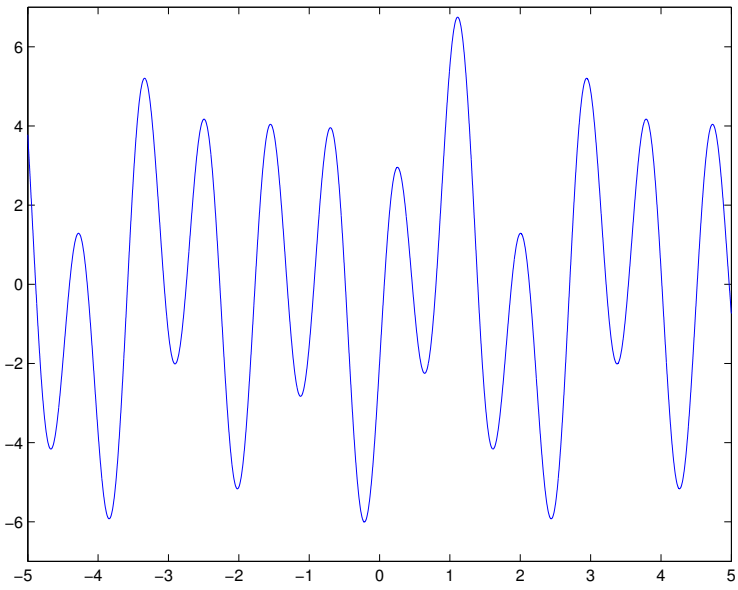


Figure 2.4: *The plot of the function* $f(t) = \sin(2t) - 2\cos(3t) + 4\sin(7t)$. *Being the linear sum of sinusoids,* f *is sparse in the Fourier domain.*

simple and more readily processed. We shall examine objects which are sparsely represented by some of the systems we have encountered: Fourier series, discrete wavelet systems, and discrete shearlet systems. For example, Figure 2.4 depicts a signal that is sparse in the Fourier domain.

In order to relate these systems to the sparse representation of signals, we introduce some notation. We consider data modeled as a function $f : \Omega \to M$ where Ω is some domain, and M is a manifold. This data may be given explicitly, or implicitly as $F(f) = 0$ for some operator F. The implicit case may be understood as modeling an operator equation or inverse problem. A *dictionary* is a set of functions Φ with which a collection of signals C may be decomposed.

Our ambition is to find a dictionary $\{\phi_i\}_{i \in I}$ of functions such that each signal

f in a given signal class C may be represented as a *sparse* linear combination of dictionary elements:

$$f \sim \sum_{i \in J} c_i \phi_i, \ |J| = N << \infty. \tag{2.10}$$

In the case that f is not given explicitly, but rather implicitly via an inverse problem $F(f) = 0$, our dictionary ought to diagonalize the operator F. A notion of such diagonalization can be seen in the works of Meyer and Coifman [86].

A canonical example of such a sparse dictionary in practice is the JPEG 2000 algorithm, which represents an image signal sparsely in a wavelet basis [105].

It is important to have a mathematical description of what types of signals we shall represent with our dictionary. One standard class of signals is the class of *cartoon-like images*:

Definition 2.5.1. *The set of* cartoon-like images *in* $\mathbb{R}^2$ *is*

$$\mathcal{E} := \{f \mid f = f_0 + \chi_B f_1, \ f_i \in \mathcal{C}^2([0,1]^2), \|f_i\|_{\mathcal{C}^2} \leq 1, \ B \subset [0,1]^2, \ \partial B \in \mathcal{C}^2([0,1])\}.$$

The space of cartoon-like images is a quantitative definition of signals that represent images. That is, although images are discrete, if we are to consider only continuous signals, then $\mathcal{E}$ represents the class of signals corresponding to images.

We shall present the theory of dictionary representations generally, and consider arbitrary Euclidean spaces. In order to develop a notion of dictionary quality, we fix ideas: $C \subset L^2(\mathbb{R}^d)$ is a signal class for some integer $d \geq 1$ an integer. $\Phi = \{\phi_i\}_{i \in I}$ is a dictionary.

Definition 2.5.2. *For a fixed positive integer* N*, the* set of N-term approximations

with Φ *is*

$$\sum_N(\Phi) := \left\{ \sum_{i\in J} c_i\phi_i \;\middle|\; J \subset I,\ |J| = N \right\}.$$

For a fixed $f \in C$, the N-term error of f in the dictionary Φ *is*

$$\sigma_N(f,\Phi) := \inf_{g\in\sum_N(\Phi)} \|f-g\|_2.$$

Definition 2.5.3. *For a dictionary Φ and signal class C, we set*

$$\mathcal{A}^s(\Phi) := \{f \in L^2(\mathbb{R}^d) \mid \sigma_N(f,\Phi) \le KN^{-s}\},$$

$$s^*(C,\Phi) := \sup_{s>0}\{C \subset \mathcal{A}^s(\Phi)\},$$

where K is fixed constant independent of N, s.

The set $\mathcal{A}^s(\Phi)$ is the set of L^2 functions f whose N-term approximation in the dictionary Φ converges to f at a rate of order $\frac{1}{N^s}$. The number $s^*(C,\Phi)$ measures the worst-case approximation rate for the set of functions C with respect to the dictionary Φ.

The larger $s^*(C,\Phi)$ is, the better the dictionary Φ is at representing the signal class C. To illustrate, suppose Φ is a dense subset of $L^2(\mathbb{R}^d)$, for example the space of Schwartz functions $\mathcal{S}(\mathbb{R}^d)$. Then given a precision ϵ and function $f \in C$, one can always find a single element $g \in \Phi$ such that $\|f-g\|_2 < \epsilon$. Thus, $s^*(C,\Phi) = +\infty$. This is good, in the sense that Φ represents C very well; this is of course in agreement with the intuition of a dense set.

However, the utility of $\Phi = \mathcal{S}(\mathbb{R}^d)$ as a computational dictionary is essentially nil. Indeed, there is in general no way of efficiently computing which element of

Φ is a good approximation of $f \in C$. Moreover, even if such an approximation could be known easily, there is not a simple way of *storing* it. The space $\mathcal{S}(\mathbb{R}^d)$ does not admit a finite dimensional parametrization. One would need to know the approximation values on a set of infinite cardinality to store an approximation with respect to this gigantic dictionary, which is of course computationally infeasible.

This brings us to a clarification of our goal. We want to find dictionaries Φ such that $s^*(C, \Phi)$ is *large enough*, while the N-term approximations in this dictionary can be computed and stored efficiently. More precisely, we take as our indexing set $I = \mathbb{N}$, and search for the best N-term approximation

$$f_N = \sum_{k=1}^{N} c_{\sigma(k,f)} \phi_{\sigma(k,f)},$$

where $\sigma(k, f) \leq \pi(k)$ for a polynomial π. In other words, we want to consider only approximations that can be determined in *polynomial time*. This re-formulated notion of dictionary quality shall be denoted $s^*_{\text{poly}}(C, \Phi)$.

Definition 2.5.4. *Set*

$$s^*(C) := \sup_{\Phi \text{ a dictionary}} \{s^*_{\text{poly}}(C, \Phi)\}.$$

A dictionary Φ is said to be optimal *for a signal class C if $s^*(C, \Phi) = s^*(C)$.*

In order to begin an analysis of rates of approximation in dictionaries, we review some pioneering work of Donoho and his collaborators [35]. Consider the following definitions.

Definition 2.5.5. *A class of functions $C \subset L^2(\mathbb{R}^d)$* contains an embedded orthogonal hypercube of dimension m and side lengths δ *if there exists $f_0 \in C$ and an*

orthonormal family $\{\psi_i\}_{i=1}^m$ *such that the collection of hypercube vertices*

$$\mathcal{H}(m; f_0; \{\psi_i\}) := \left\{ h = f_0 + \sum_{i=1}^{m} \epsilon_i \psi_i \;\middle|\; \epsilon_i \in \{0,1\} \right\}$$

is contained in C*.*

Definition 2.5.6. *A class of functions* C contains a copy of $\ell_0^p, p > 0$, *if there exists a sequence of orthogonal hypercubes* $\{\mathcal{H}_k\}_{k\in\mathbb{N}}$ *of dimensions* $\{m_k\}_{k\in\mathbb{N}}$*, embedded in* C*, of dimensions* m_k *and side lengths* δ_k*, such that* $\delta_k \to 0$ *and for some* $K > 0$*,*

$$m_k \geq K\delta_k^{-p}, \ \forall k \in \mathbb{N}.$$

These notions allow us to derive an upper bound on how well an arbitrary dictionary can represent a signal space:

Theorem 2.5.7. *Suppose* $C \subset L^2(\mathbb{R}^d)$ *is uniformly* L^2*-bounded and contains a copy of* ℓ_0^p*. Then*

$$s^*(C) \leq \frac{2-p}{p}.$$

In order to better understand optimal dictionaries for $\mathcal{E}$, we consider certain spaces of smooth functions.

Definition 2.5.8. *For* $\alpha = (\alpha_1, ..., \alpha_d)$*,* $\alpha_i \in \mathbb{Z}, \alpha_i \geq 0$*, the space*

$$\mathcal{C}^\alpha := \{ f \in \mathcal{C}^\alpha([0,1]^d) \mid \ \|f\|_{\mathcal{C}^\alpha} \leq 1 \}$$

is the closed unit ball in $\mathcal{C}^\alpha([0,1]^d)$*, where*

$$\|f\|_{\mathcal{C}^\alpha} := \sum_{0 \leq \beta \leq \alpha} \|D^\beta g\|_{L^\infty([0,1]^d)},$$

for multi-indices β*.*

Theorem 2.5.9. *Suppose $\mathcal{C}^\alpha$ contains a copy of ℓ_0^p with $p = \frac{2}{\frac{2\alpha}{d}+1}$. Then $s^*(\mathcal{C}^\alpha) \leq \frac{\alpha}{d}$.*

It can be shown that in dimension 2, $\mathcal{C}^\alpha$ contains a copy of $\ell_0^{\frac{2}{3}}$. Thus, in the case $d = 2$, $\mathcal{C}^2 \subset \mathcal{E}$ and $s^*(\mathcal{C}^2) \leq 1$ by above. In particular, $s^*(\mathcal{E}) \leq 1$. Thus, the best N-term approximation rate we can achieve for $\mathcal{E}$ is of order N^{-1}. Whether this theoretically optimal rate can be realized is a separate matter entirely. To do so, we shall examine how several dictionaries represent $\mathcal{E}$.

Definition 2.5.10. *For a sequence of complex numbers $\{c_i\}_{i\in\mathbb{N}}$, we define the* weak ℓ_p-norm *to be:*

$$\|\{c_i\}\|_{w^{\ell_p}} := \inf_{K>0}\{|c_n^*| \leq Kn^{-\frac{1}{p}}\},$$

where $\{c_n^*\}$ is a non-decreasing re-arrangement of $\{c_i\}$ and K is a constant independent of n.

It is worth noting that $\|\{c_i\}\|_{w^{\ell_p}} \leq \|\{c_i\}\|_{\ell_p}$, which implies the weak ℓ_p norm is indeed weaker than the ℓ_p norm.

The next result requires the notion of a *dual frame*, which is classical in frame theory.

Definition 2.5.11. *Let $\{\phi_i\}_{i\in I}$ be a discrete frame for the Hilbert space $\mathcal{H}$. A* dual frame for $\{\phi_i\}_{i\in I}$ *is a set of vectors $\{\tilde{\phi}_i\}_{i\in I} \subset \mathcal{H}$ such that:*

$$\forall f \in \mathcal{H},\ f = \sum_{i\in I}\langle f, \tilde{\phi}_i\rangle\phi_i.$$

Intuitively, a dual frame is a set of functions which can be used along with the original frame coefficients to precisely reconstruct our signal. With this in mind, we

can relate the weak ℓ_p convergence of frame coefficients to the approximation rate in that dictionary [50] :

Theorem 2.5.12. *Let $\Phi = \{\phi_i\}_{i\in I}$ be a frame for $L^2(\mathbb{R}^d)$ with dual frame $\tilde{\Phi} = \{\tilde{\phi}_i\}_{i\in I}$. Suppose that*

$$\|\{\langle f, \tilde{\phi}_i\rangle\}\|_{w^{\ell_p}} < \infty.$$

Then, with $s = \frac{1}{p} - \frac{1}{2}$, we have $f \in \mathcal{A}^s(\Phi)$ and

$$\left\| f - \sum_{i\in I_N} \langle f, \tilde{\phi}_i\rangle \phi_i \right\|_2 \leq CN^{-s}.$$

2.5.1 Point Singularities: Fourier Series and Wavelets

We shall now begin our analysis of convergence rates for certain dictionary representations of important classes of functions. The classes of functions are intuitively related to classes of images, but are described in a continuous, rather than discrete, language. We are ultimately interested in $\mathcal{E}$, but we shall first consider slightly more basic signal classes.

We begin by examining how one-dimensional singularities can be resolved by certain dictionaries.

Definition 2.5.13. *For $\alpha \geq 0$, define*

$$\mathcal{D}^\alpha := \{f \in L^2([0,1]) \mid f = f_1\chi_{[0,a]} + f_2\chi_{(a,1]},\ \|f_i\|_{\mathcal{C}^\alpha} \leq 1,\ a \in [0,1]\}$$

to be the set of $\mathcal{C}^\alpha$ functions with a point singularity*.*

The Gibbs phenomenon [4] can be used to show that Fourier series poorly represent the signal class $\mathcal{D}^\alpha$:

Theorem 2.5.14. *Let $\mathcal{F} = \{e^{2\pi inx}\}_{n\in\mathbb{Z}}$ be the dictionary of Fourier basis elements. Then $s^*(\mathcal{D}^\alpha, \mathcal{F}) = \frac{1}{2}$, for all α.*

This theorem must be understood in light of the theoretical optimum $s^*(\mathcal{D}^\alpha, \Phi) = \alpha$ for Φ a dictionary in the class we consider, guaranteed by Theorem 2.5.9. For $\alpha = 1$ in particular, the theoretical optimal is $s^*(D^1, \Phi) = 1$, whereas $s^*(D^1, \mathcal{F}) = \frac{1}{2}$. An intuitive explanation for this sub-optimality is that elements of the Fourier basis are globally supported, making them ill-suited for resolving singularities, which are local.

In order to show that the theoretical optimum is actually attainable, we consider wavelet systems. Using a slightly different convention than in Section 2.3, consider a fixed $\alpha \in \mathbb{R}$, scaling function $\phi(x)$, and mother wavelet $\psi(x)$. The corresponding wavelet system is:

$$\mathcal{W}(\phi, \psi, \alpha) := \{\phi(x - \alpha k)\}_{k\in\mathbb{Z}} \cup \{2^{\frac{j}{2}}\psi(2^j x - \alpha k)\}_{j,k\in\mathbb{Z}}.$$

For simplicity's sake, consider the case $\alpha = 1$. Daubechies et al. showed that ϕ, ψ may be chosen to be compactly supported and very smooth [31]. Let

$$\psi_{j,k}(x) := 2^{\frac{j}{2}}\psi(2^j x - k);$$

intuitively, $\langle f, \psi_{j,k}\rangle$ extracts frequency information about f on an interval of length proportional to 2^{-j}.

In order to state the next theorem [44], we consider a particular orthonormal wavelet basis:

Definition 2.5.15. *The* orthonormal basis of Meyer wavelets *is the collection of*

functions $\{\phi(x-k)\}_{k\in\mathbb{Z}}\cup\{2^{\frac{j}{2}}\psi(2^j x-k)\}_{j,k\in\mathbb{Z}}$*, where* ϕ,ψ *are defined in the frequency domain as:*

$$\hat{\psi}(\gamma)=\begin{cases}\frac{1}{\sqrt{2\pi}}\sin\left(\frac{\pi}{2}\nu\left(\frac{3|\gamma|}{2\pi}-1\right)\right)e^{\frac{i\gamma}{2}},\ |\gamma|\in(\frac{2\pi}{3},\frac{4\pi}{3}).\\ \frac{1}{\sqrt{2\pi}}\cos\left(\frac{\pi}{2}\nu\left(\frac{3|\gamma|}{4\pi}-1\right)\right)e^{\frac{i\gamma}{2}},\ |\gamma|\in(\frac{4\pi}{3},\frac{8\pi}{3}).\\ 0,\quad \text{else.}\end{cases}$$

$$\hat{\phi}(\gamma)=\begin{cases}\frac{1}{\sqrt{2\pi}},\ |\gamma|\in[0,\frac{2\pi}{3}).\\ \frac{1}{\sqrt{2\pi}}\cos\left(\frac{\pi}{2}\nu\left(\frac{3|\gamma|}{2\pi}-1\right)\right),\ |\gamma|\in(\frac{2\pi}{3},\frac{4\pi}{3}).\\ 0,\ \text{else.}\end{cases}$$

and where the auxiliary function ν *is defined as:*

$$\nu(x)=\begin{cases}0,\ x\in(\infty,0].\\ x,\ x\in(0,1).\\ 1,\ x\in[1,\infty).\end{cases}$$

Theorem 2.5.16. *Suppose* Φ *is the orthonormal basis of Meyer wavelets* $p=(1+\alpha)^{-1}$*. Then:*

$$\sup_{f\in\mathcal{D}^\alpha}\|\{\langle f,\psi_{j,k}\rangle\}\|_{w^{\ell_p}}<\infty.$$

In particular,

$$s^*(\mathcal{W}(\phi,\psi,1),\mathcal{D}^\alpha)=s^*(\mathcal{C}^\alpha)=\alpha.$$

In other words, wavelets provide optimal approximation of functions with point singularities. Indeed, the approximation given by the Meyer wavelets is *as efficient as if the singularities were not present.*

However, the efficiency of wavelets in representing singularities does not extend to higher dimensions [44]. Indeed, let α, ϕ, ψ be as above and let $\mathcal{W}^{2D}(\phi, \psi, \alpha)$ denote a $2D$-wavelet system generated in the natural way by tensor products of ϕ, ψ, i.e. using translations and dilations of the functions $\{\psi_i(x, y)\}_{i=1}^4$ given by:

$$\psi_1(x, y) = \phi(x)\phi(y),$$

$$\psi_2(x, y) = \phi(x)\psi(y),$$

$$\psi_3(x, y) = \psi(x)\phi(y),$$

$$\psi_4(x, y) = \psi(x)\psi(y).$$

Theorem 2.5.17. $s^*(\mathcal{E}, \mathcal{W}^{2D}(\phi, \psi, 1)) = \frac{1}{2}$.

From a computational standpoint, this a dreadful result. Indeed, if we wish to approximate $f \in \mathcal{E}$ with $2D$ tensor wavelets to an error of 10^{-5}, i.e. if we wish to find an N-term approximation of f in the $2D$ tensor wavelet dictionary, call it f_N, such that

$$\|f - f_N\| \leq 10^{-5},$$

We must take $N = 10^{10}$, which is infeasible. If the optimal approximation rate of $\frac{1}{N}$ is achieved, we would only need $N = 10^5$. This would significantly ease the computational burden of using these dictionaries for data processing. Note that this sub-optimality also holds for more elaborate $2D$ wavelet schemes, not just tensor wavelets.

2.5.2 Geometric Frames: Curvelets and Shearlets

Our goal is to construct a tight frame for $L^2(\mathbb{R}^2)$ that has theoretically optimal representation properties for $\mathcal{E}$, i.e. for all $f \in \mathcal{E}$,

$$\|f - f_N\|_2 \leq CN^{-1},$$

where f_N is the reconstruction from the N-largest frame coefficients and C is a constant independent of N.

The problem with wavelets in the case of resolving singularities in $\mathbb{R}^2$ is that wavelets are supported in isotropic quadrilaterals of width $\sim 2^j$; too many such quadrilaterals are needed to cover a curve singularity. This argument is explicitly detailed in [44].

To rectify this, we consider dictionary elements supported on *anisotropic rectangles* with width $\sim 2^{\frac{j}{2}}$ and length $\sim 2^j$. This is the basic idea behind the curvelet construction of Donoho and Candès [13]. This was the first frame construction to exploit the *parabolic scaling* regime in order to acquire near-optimal representation systems for $\mathcal{E}$. We give a rough sketch of their construction; details are in [13], [12].

We seek to construct systems of the form

$$\{\phi_{j,\ell,k}(x) := 2^{\frac{3j}{k}} \psi(D_{2^j} R_{\theta_{j,\ell}} x - k)\}_{j,\ell \in \mathbb{Z}, k \in \mathbb{Z}^2},$$

where

$$D_a := \begin{pmatrix} a & 0 \\ 0 & a^{\frac{1}{2}} \end{pmatrix}, \; R_\theta := \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}, \; \theta_{j,\ell} \sim 2^{\frac{-j}{2}} \ell 2\pi.$$

To perform this construction, we partition frequency space into *parabolic wedges* of aspect ratio length $\sim$ width^2. We associate to this a partition of unity $\{V_{j,\ell}\}_{j\in\mathbb{N},\ell\in\left\{-2^{\frac{j}{2}},\ldots,2^{\frac{j}{2}}\right\}}$. We build a dictionary in frequency space by modulating the partition functions:

$$\hat{\phi}_{j,\ell,k}(\gamma) := 2^{\frac{-3j}{4}} e^{2\pi i R_{\theta_{j,\ell}}^{-1} D_{2^{-j}} k\gamma} V_{j,\ell}(\gamma), j \in \mathbb{N}, \ell \in \{-2^{\frac{j}{2}}, ..., 2^{\frac{j}{2}}\}.$$

Collecting all the indices into the set Λ, we acquire a dictionary $\{\phi_\lambda\}_{\lambda\in\Lambda}$, which is actually a Parseval frame [13]:

Theorem 2.5.18. *The dictionary $\{\phi_\lambda\}_{\lambda\in\Lambda}$ is a Parseval frame for $L^2(\mathbb{R}^2)$, i.e.*

$$\|f\|_2^2 = \sum_{\lambda\in\Lambda} |\langle f, \phi_\lambda\rangle|^2,$$

and

$$f = \sum_{\lambda\in\Lambda} \langle f, \phi_\lambda\rangle \phi_\lambda.$$

Theorem 2.5.19. *Curvelets are optimal for cartoon-like images:*

$$s^*(\mathcal{E}, \{\phi_\lambda\}_{\lambda\in\Lambda}) = s^*(\mathcal{E}) = 1.$$

While curvelets are theoretically optimal [13], from a computational standpoint, they are insufficient; they don't preserve the digital grid. While there have been numerical implementations of curvelets [12], they rely on countably many generating functions and are slower than desirable for many scenarios. Shearlets were developed after curvelets, partially in response to this criticism [70]. As we have seen in Section 2.4, shearlets are more simply generated than curvelets, and preserve the digital grid, in the sense that the shearing matrix preserves the integer lattice, making them more suitable for applications.

Following the literature of dictionary representations, we give a slightly different shearlet construction than the one in Section 2.4. For this construction, we consider the index set

$$\Lambda^\sigma := \{(\epsilon, j, \ell, k) \in \mathbb{Z}_2 \times \mathbb{Z}^4 \mid \epsilon \in \{0,1\},\ j \geq 0,\ \ell = -2^{\lfloor \frac{j}{2} \rfloor}, ..., 2^{\lfloor \frac{j}{2} \rfloor},\ k \in \mathbb{Z}^2\}.$$

The associated *shearlet system* generated by $\phi, \psi^0, \psi^1 \in L^2(\mathbb{R}^2)$ is

$$\Sigma := \{\sigma_\lambda \mid \lambda \in \Lambda^\sigma\},$$

where

$$\sigma_{(\epsilon,0,0,k)}(x) := \phi(x-k),\ \sigma_{(\epsilon,j,\ell,k)}(x) := 2^{\frac{3j}{4}} \psi^\epsilon(D^\epsilon_{2^j} S^\epsilon_{\ell,j} x - k).$$

Here,

$$D^0_a = D_a,\ D^1_a = \begin{pmatrix} a^{\frac{1}{2}} & 0 \\ 0 & a \end{pmatrix},\ S^0_{\ell,j} := \begin{pmatrix} 1 & \ell 2^{-\lfloor \frac{j}{2} \rfloor} \\ 0 & 1 \end{pmatrix},\ S^1_{\ell,j} = (S^0_{\ell,j})^T.$$

Theorem 2.5.20. *Shearlets are optimal for cartoon-like images: for a band-limited shearlet frame Σ (i.e. $\hat{\phi}, \hat{\psi}$ are compactly supported), we have*

$$s^*(\mathcal{E}, \Sigma) = s^*(\mathcal{E}) = 1.$$

Thus, shearlets achieve the optimal rate of approximation for the space of cartoon-like images. This suggests their use for image processing. Shearlet sparsity and optimality for anisotropic images shall be exploited for image registration and superresolution in Chapters 4 and 5 of this thesis.

2.6 Wavelet Packets

2.6.1 Wavelet Packets: Completing the Tree

This chapter provides the theoretical background, now classical, necessary for the material in Chapter 4. We construct wavelet packets in one dimension, a process which easily generalizes to higher dimensions. Our work in Chapter 4 analyzes remotely sensed data with wavelet packets, which requires a two-dimensional generalization.

Consider a multiresolution analysis $\{V_j\}_{j\in\mathbb{Z}},\ \phi$, with corresponding wavelet function ψ as guaranteed by the Meyer-Mallat algorithm; see Theorem 2.3.8. Then $\{T_n\phi\}_{n\in\mathbb{Z}}$ is an orthonormal basis for V_0 and $\{T_n\psi\}_{n\in\mathbb{Z}}$ is an orthonormal basis for W_0, where W_0 is the unique closed subspace of V_1 such that $V_0 \oplus W_0 = V_1$. In general, W_k is the unique closed subspace of V_{k+1} such that $V_k \oplus W_k = V_{k+1}$.

Besides $\{T_n\phi\}_{n\in\mathbb{Z}}$, the set $\{\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}-n\right), \frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}-n\right)\}_{n\in\mathbb{Z}}$ is an orthonormal basis for $V_0 = V_{-1} \oplus W_{-1}$. This procedure allows us to construct *many* orthonormal bases for V_0. Indeed,

$$V_0 = V_{-m} \oplus W_{-m} \oplus W_{-m+1} \oplus ... \oplus W_{-1}.$$

So, orthonormal bases for $W_k,\ k \leq -1$, may be incorporated into a basis for V_0. More precisely, the set

$$\{\phi_{-m,n},\ \psi_{-m,n},\ \psi_{-m+1,n},\ ...,\ \psi_{-1,n}\}_{n\in\mathbb{Z}}$$

is an orthonormal basis for V_0 for all $m \geq 1$. As usual, $f_{m,n}(t) = 2^{\frac{m}{2}} f(2^m t - n)$.

Considering the associated tree of subspaces, we see that there is an incompleteness. The left-most branch of the tree is deep, but as one moves further right, branches cut off prematurely. The tree for $m = 5$ is depicted in Figure 2.5.

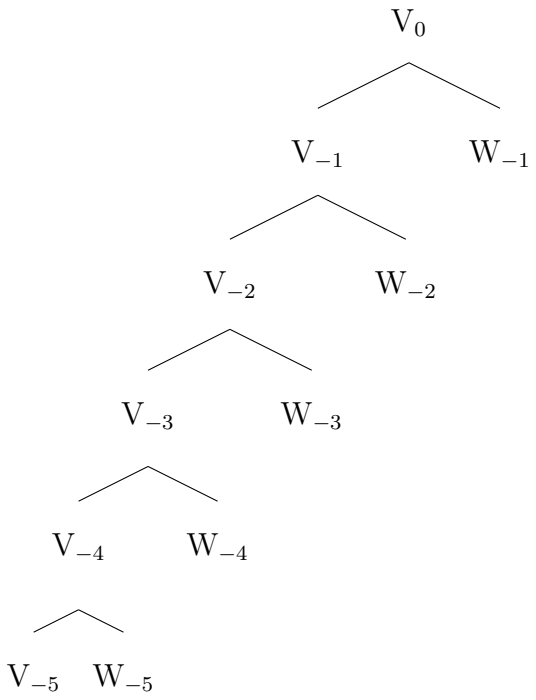


Figure 2.5: *The "incomplete" tree produced by the Meyer-Mallat algorithm, iterated to 5 levels. The leftmost path has maximal depth.*

One way of understanding the notion of *wavelet packets* is as an attempt to *complete this tree* by having all branches be of equal depth. We now pursue this idea rigorously.

2.6.2 The Main Construction

Letting ψ and ϕ with associated filters H_0, H_1 be our wavelet and scaling function, it is clear from the Meyer-Mallat algorithm that:

$$\begin{aligned}
&\left\{\phi_{-1,n}(t) = \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2} - n\right)\right\}_{n\in\mathbb{Z}} \text{ is an orthonormal basis for } V_{-1},\\
&\left\{\psi_{-1,n}(t) = \frac{1}{\sqrt{2}}\psi\left(\frac{t}{2} - n\right)\right\}_{n\in\mathbb{Z}} \text{ is an orthonormal basis for } W_{-1}.
\end{aligned} \tag{2.11}$$

More precisely, we introduce the notation

$$\begin{aligned}
\phi &= \sum_{n\in\mathbb{Z}} h_0[n]\phi_{1,n},\ \phi_{1,n} \in V_1,\\
\psi &= \sum_{n\in\mathbb{Z}} h_1[n]\phi_{1,n}, \phi_{1,n} \in V_1,
\end{aligned}$$

which implies

$$\begin{aligned}
\hat{\phi}(\gamma) &= \frac{1}{\sqrt{2}}\hat{\phi}\left(\frac{\gamma}{2}\right) H_0\left(\frac{\gamma}{2}\right),\\
\hat{\psi}(\gamma) &= \frac{1}{\sqrt{2}}\hat{\phi}\left(\frac{\gamma}{2}\right) H_1\left(\frac{\gamma}{2}\right).
\end{aligned}$$

Applying the Fourier transform to (2.11), we have that:

$$\begin{aligned}
&\{\hat{\phi}_{-1,n}(\gamma) = \sqrt{2}e^{2\pi i n}\hat{\phi}\,(2\gamma)\}_{n\in\mathbb{Z}} \text{ is an orthonormal basis for } \widehat{V_{-1}},\\
&\{\hat{\psi}_{-1,n}(\gamma) = \sqrt{2}e^{2\pi i n}\hat{\psi}\,(2\gamma)\}_{n\in\mathbb{Z}} \text{ is an orthonormal basis for } \widehat{W_{-1}}.
\end{aligned}$$

Recall that since $V_{-1} = V_{-2} \oplus W_{-2}$, we have that

$$\left\{\frac{1}{2}\phi\left(\frac{t}{4} - n\right), \frac{1}{2}\psi\left(\frac{t}{4} - n\right)\right\}_{n\in\mathbb{Z}}$$

is an orthonormal basis for V_{-1}.

Setting $\phi_{-2,n}(t) = \frac{1}{2}\phi\left(\frac{t}{4} - n\right)$ and taking the Fourier transform, $\hat{\phi}_{-2,n}(\gamma) = 2e^{4\pi i n}\phi(4\gamma)$. Recall by the scaling equation of the Meyer-Mallat algorithm,

$$2\hat{\phi}(4\gamma) = \sqrt{2}\hat{\phi}(2\gamma)H_0(2\gamma) = \hat{\phi}(\gamma)H_0(\gamma)H_0(2\gamma).$$

Similarly,

$$2\hat{\psi}(4\gamma) = \hat{\phi}(\gamma)H_0(\gamma)H_1(2\gamma).$$

This decomposition of $\hat{\phi}$ and $\hat{\psi}$ in terms of the filters H_0, H_1 suggests the following definition:

Definition 2.6.1. *Let $\{V_j\}_{j\in\mathbb{Z}}, \phi$ be an MRA with wavelet ψ and filters H_0, H_1. The* wavelet packet filters *associated with this MRA is the sequence of products*

$$\prod_{j=1}^{r} H_{\epsilon_j}(2^{j-1}\gamma),\ r \geq 1,\ \epsilon_j \in \{0,1\}.$$

This idea can be understood in terms of *bit reversal* permutations. This notion is classical [63], and can be defined inductively. Indeed, let $N = 2^r$. At level $r = m$, suppose the set $\{0, 1, ..., 2^m - 1\}$ has bit reversal ordering $(b_0, b_1, ..., b_M)$, where $M = 2^m - 1$. At level $r = m+1$, the bit reversal ordering of the set $\{0, 1, ..., 2^{m+1} - 1\}$ is the set $\{2b_0, 2b_1, 2b_2, ..., 2b_{M-1}, 2b_0 + 1, 2b_1 + 1, 2b_2 + 1, ..., 2b_{M-1} + 1\}$. Setting the bit reversal of $\{0\}$ to be $\{0\}$, all other reversals are well-defined by induction.

For example, the bit reversal of some small sets are:

$$\{0, 1, 2, 3\} \mapsto \{0, 2, 1, 3\},$$

$$\{0, 1, 2, 3, 4, 5, 6, 7\} \mapsto \{0, 4, 2, 6, 1, 5, 3, 7\}.$$

To relate bit reversal to wavelet packets, we shall consider a tree of subspaces $\{X_n^r\}_{n=0}^N$, where r is the depth of the tree and $N = 2^r$. To define these subspaces, we recall that all integers $1 \leq n \leq 2^r - 1$ may be written uniquely as:

$$n = \sum_{j=1}^{r} \epsilon_j 2^{j-1}, \ \epsilon_j \in \{0, 1\}.$$

We may thus identify the indices n of the X_n^r with binary r-tuples. For example, if $r = 4$:

$$3 \mapsto (1, 1, 0, 0),$$

$$11 \mapsto (1, 1, 0, 1).$$

Consider Figure 2.6

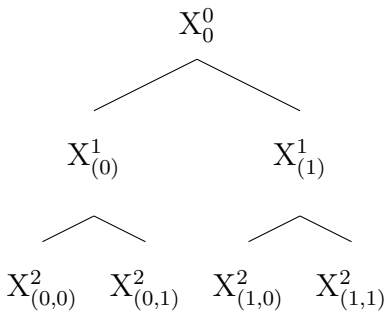


Figure 2.6: *An example of a subspace tree of depth 2, with subspaces at the bottom level indexed by binary pairs.*

At level $r-1$, the space $X^{r-1}_{(\epsilon_1,\epsilon_2,\ldots,\epsilon_{r-1})}$ is the single parent of $X^r_{(\epsilon_1,\epsilon_2,\ldots,\epsilon_{r-1},0)}$ and $X^r_{(\epsilon_0,\epsilon_1,\ldots\epsilon_{r-1},1)}$. Making the identification $X_0^0 = V_0$, it follows that at level 1,

$$X_0^1 = V_{-1},$$

$$X_1^1 = W_{-1},$$

and at level 2,

$$X^2_{(0,0)} = V_{-2},$$

$$X^2_{(0,1)} = W_{-2}.$$

However, a creative choice must be made to determine $X^2_{(1,0)}$ and $X^2_{(1,1)}$.

In terms of our filters,

$$X^2_{(0,0)} = V_{-2} \leftrightarrow H_0(\gamma)H_0(2\gamma),$$

$$X^2_{(0,1)} = W_{-2} \leftrightarrow H_0(\gamma)H_1(2\gamma).$$

It is thus natural to set our unknown spaces in correspondence with these filters based on their indices:

$$X^2_{(1,0)} \leftrightarrow H_1(\gamma)H_0(2\gamma),$$

$$X^2_{(1,1)} \leftrightarrow H_1(\gamma)H_1(2\gamma).$$

Therefore, we define $\theta_{1,0}$ and $\theta_{1,1}$ as:

$$2\hat{\theta}_{1,0}(4\gamma) = \hat{\phi}(\gamma)H_1(\gamma)H_0(2\gamma),$$

$$2\hat{\theta}_{1,1}(4\gamma) = \hat{\phi}(\gamma)H_1(\gamma)H_1(2\gamma).$$

Hence,

$$X^2_{(1,0)} = \overline{\text{span}}\left\{\frac{1}{2}\theta_{1,0}\left(\frac{t}{4} - n\right)\right\}_{n\in\mathbb{Z}},$$
$$X^2_{(1,1)} = \overline{\text{span}}\left\{\frac{1}{2}\theta_{1,1}\left(\frac{t}{4} - n\right)\right\}_{n\in\mathbb{Z}}.$$

To summarize the construction, given an MRA $\{V_j\}_{j\in\mathbb{Z}}, \phi$ with Meyer-Mallat wavelet ψ, filters H_0, H_1 from the Meyer-Mallat theorem, we construct a perfect binary tree of subspaces $\{X^r_m\}_{r\geq 1,\ m=0,1,\ldots,2^r-1}$ and $X^0_0 = V_0$, with the properties:

- Each X_n^r is a closed linear subspace of $L^2(\mathbb{R})$.

- $X_m^r,\ r \geq 1$, is determined by the function

$$\hat{\theta}_{\epsilon_1,\ldots,\epsilon_r}(2^r\gamma) = 2^{-\frac{r}{2}}\hat{\phi}(\gamma)H_{\epsilon_1}(\gamma)H_{\epsilon_2}(2\gamma)...H_{\epsilon_r}(2^{r-1}\gamma), \tag{2.12}$$

  where $m = \sum_{j=1}^{r} \epsilon_j 2^{j-1},\ \epsilon_j \in \{0,1\}$. By determined, we mean the set

$$\left\{2^{-\frac{r}{2}}\theta_{\epsilon_1,\ldots,\epsilon_r}\left(\frac{t}{2^r} - n\right)\right\}_{n\in\mathbb{Z}}$$

  is an orthonormal basis for X_m^r.

- For any fixed integer $r \geq 0$, $V_0 = \bigoplus_{m=0}^{2^r-1} X_m^r$.

We can define a more general class of wavelet packets based on this principle:

Definition 2.6.2. *A sequence $\{w_i\}_{i=0}^{\infty} \subset L^2(\mathbb{R})$ is a sequence of* wavelet packets *corresponding to filters H_0, H_1 if $\forall i \geq 0$,*

$$w_{2i}(t) = \sum_{n\in\mathbb{Z}} h_0[n]\sqrt{2}w_i(2t-n),$$

$$w_{2i+1}(t) = \sum_{n\in\mathbb{Z}} h_1[n]\sqrt{2}w_i(2t-n).$$

The sequence $\{w_{i,j,k}\}_{i\geq 0,\ j,k\in\mathbb{Z}}$, defined by

$$w_{i,j,k}(t) := 2^{\frac{j}{2}}w_i(2^j t - k),$$

is a sequence of generalized wavelet packets corresponding to H_0, H_1.

Generalized wavelet packets modify a wavelet packet with additional dilations and translations.

Let us consider this construction in the specific context of the Haar MRA. In this case, the scaling function is $\phi = \chi_{[0,1]}$, $h_0[0] = h_0[1] = \frac{1}{\sqrt{2}}$, $h_0[k] = 0,\ k \neq 0, 1$. This forces $H_0(\gamma) = \frac{1}{\sqrt{2}} e^{-\pi i \gamma} \cos(\pi\gamma)$. Since $h_1[k] = (-1)^k \overline{h_0[1-k]}$, we compute $H_1(\gamma) = \frac{1}{\sqrt{2}} e^{-\pi i \gamma} \sin(\pi\gamma)$.

Thus, by the Meyer-Mallat algorithm, the wavelet function for the Haar MRA is $\psi = \chi_{[0,\frac{1}{2})} - \chi_{[\frac{1}{2},1)}$. In order to find the wavelet packets corresponding to H_0, H_1, we must solve the following equations:

$$\begin{aligned} w_0(t) =& w_0(2t) + w_0(2t-1), \\ w_1(t) =& w_0(2t) - w_0(2t-1). \end{aligned} \tag{2.13}$$

Notice that $w_0 = \phi, w_1 = \psi$ solve these equations. To show these are unique in some sense, we apply the Fourier transform to (2.13):

$$\hat{w}_0(\gamma) = \frac{1}{2}\left(1 + e^{-\pi i \gamma}\right) \hat{w}_0\left(\frac{\gamma}{2}\right).$$

Thus,

$$\begin{aligned} \hat{w}_0(\gamma) =& \frac{1}{\sqrt{2}} H_0\left(\frac{\gamma}{2}\right) \hat{w}_0\left(\frac{\gamma}{2}\right) \\ =& \frac{1}{\sqrt{2}} H_0\left(\frac{\gamma}{2}\right) \left(\frac{1}{\sqrt{2}} H_0\left(\frac{\gamma}{4}\right)\right) \hat{w}_0\left(\frac{\gamma}{4}\right) \\ =& \prod_{j=1}^{\infty} \left(\frac{H_0\left(\frac{\gamma}{2^j}\right)}{\sqrt{2}}\right) \hat{w}_0(0) \\ =& \prod_{j=1}^{\infty} e^{\frac{-\pi i \gamma}{2^j}} \cos\left(\frac{\pi\gamma}{2^j}\right) \\ =& e^{-\pi i \gamma \sum_{j=1}^{\infty} 2^{-j}} \prod_{j=1}^{\infty} \cos\left(\frac{\pi\gamma}{2^j}\right) \\ =& e^{-\pi i \gamma} \cdot \frac{\sin(\pi\gamma)}{\pi\gamma} \end{aligned}$$

$$= \hat{\phi}(\gamma).$$

The penultimate line follows from Viète's formula.

Thus, the unique L^1 solution of (2.13) such that $\int_{\mathbb{R}} w_0 = 1$ is $w_0 = \phi$. In particular, we can compute the wavelet packets

$$w_2(t) = w_1(2t) + w_1(2t-1),$$

$$w_3(t) = w_1(2t) - w_1(2t-1).$$

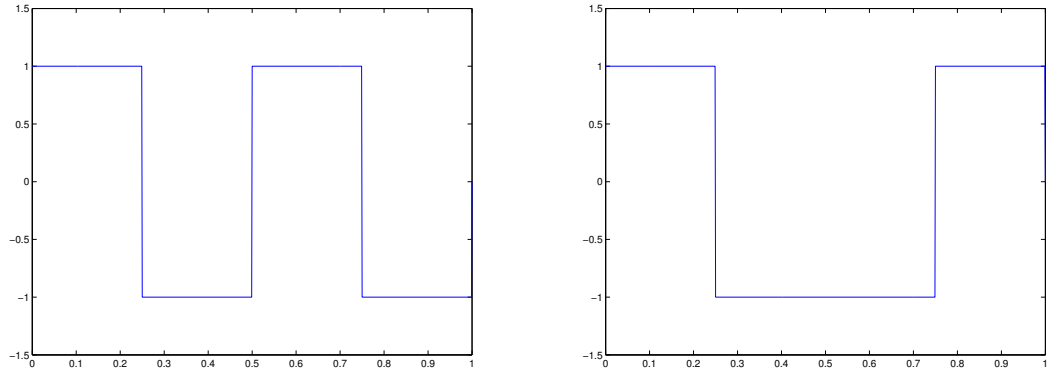


Figure 2.7: *Plots of w_2, w_3, respectively.*

Theorem 2.6.3. *For a given level $r \geq 1$, the function $\theta_{\epsilon_1,\ldots\epsilon_r}$ as defined in (2.12) is a wavelet packet w_k, where for $m = \sum_{j=1}^{r} \epsilon_j 2^{j-1}$, $\epsilon_j \in \{0,1\}$ k, m are related by the bit-reversal map*

$$(0, 1, ..., k, ..., 2^r - 1) \mapsto (b_0, ..., b_{2^r-1}),$$

$$k \mapsto b_k = m.$$

2.6.3 Relationship to Walsh Functions

The theory of wavelet packets can be traced back to Haar [53].

Definition 2.6.4. *The sequence* $\{R_n\}_{n=1}^{\infty}$ *of* Rademacher functions *on* $\mathbb{T}$ *is defined by:*

$$\forall n \geq 1,\ R_n(t) := \operatorname{sgn}(\sin(2^n \pi t)).$$

The n^{th} Rademacher function has 2^{n-1} cycles, each of length $\frac{1}{2^{n-1}}$ on $[0,1]$. The value of R on the first half of each cycle is 1; the value is -1 on the second half.

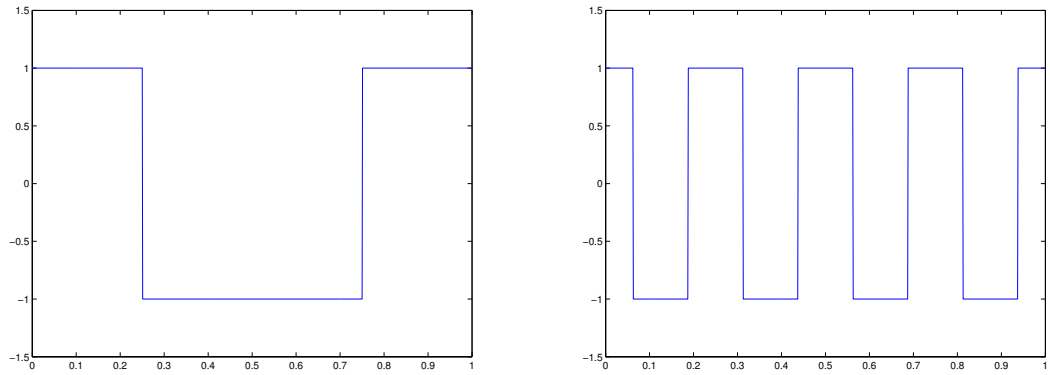


Figure 2.8: *Plots of* R_n *for* $n = 1, 3$ *respectively.*

Note that the family $\{R_n\}_{n=1}^{\infty}$ is orthonormal, but not an orthonormal basis. For example, the function $\cos(2\pi x)$ cannot be represented by this family, since

$$\int_{\mathbb{R}} \cos(2\pi x) R_n(x) dx = 0,\ \forall n.$$

The following notion is due to Paley, though was popularized by Walsh, and hence bears his name:

Definition 2.6.5. *The sequence* $\{W_n\}_{n=0}^{\infty}$ *of* Walsh functions *on* $\mathbb{T}$ *is defined by* $W_0 \equiv 1$, *and*

$$W_n(t) := \prod_{j \text{ such that } \epsilon_j = 1} R_j(t),\ \text{where } 1 \leq n = \sum_{j=1}^{\infty} \epsilon_j 2^{j-1},\ \epsilon_j \in \{0,1\}.$$

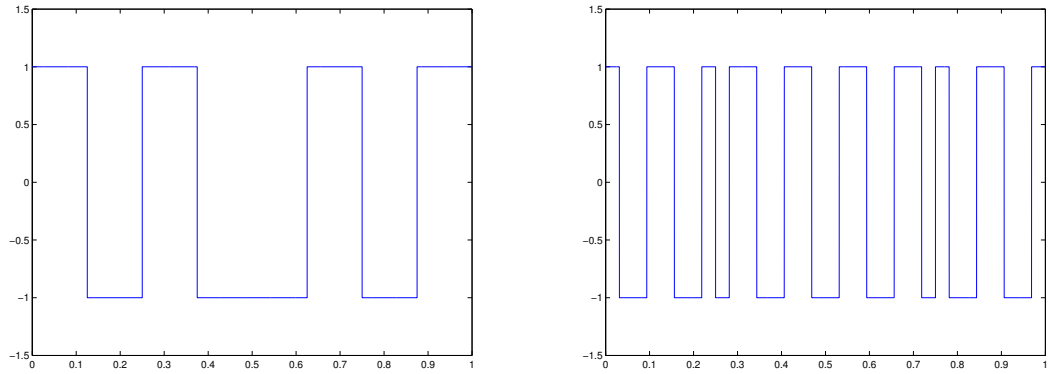


Figure 2.9: *Plots of W_n for $n = 3, 9$ respectively.*

Note that $W_{2^k} = R_k$.

Theorem 2.6.6. *Given the Haar MRA $\{V_j\}_{j\in\mathbb{Z}}, \phi$ of $L^2(\mathbb{R})$, where $\phi = \chi_{[0,1)}$, and*

$$\forall n \in \mathbb{Z},\ h_1[n] = (-1)^n \overline{h_0[1-n]},$$

the wavelet packets $\{w_j\}_{j=0}^{\infty}$ defined in terms of H_0, H_1 correspond to the Walsh functions in a manner that can be made precise, by re-ordering the Walsh functions.

Chapter 3: Directional Gabor Systems

3.1 Background on Gabor Ridge Systems

An anisotropic generalization of Gabor systems was developed by Grafakos and Sansing in [47]. They introduced directional sensitivity by considering inner products against elements of the unit sphere S^{d-1}. This method is related to the ridgelet constructions of Do and Vetterli [34]. Unlike the theory of Gabor frames articulated above, which is based on finding a discrete representation for $L^2(\mathbb{R}^d)$, their theory is continuous. That is, instead of considering *countable sums* of coefficients, they consider *integrals over uncountably many coefficients*. The theory of this work has been applied abstractly to problems in tomographic image analysis [43], [19], as well as problems in classical analysis [20], [7]. However, for use in concrete applications such as signal processing, a discrete theory is required. We shall give some results in this direction at the end of this section. We begin with some relevant notions from continuous Gabor theory, as well as from the theory of continuous frames.

We shall consider adding a directional component to Gabor systems in a manner parametrized by S^{d-1}. It is natural that the *Radon transform*, in particular the Fourier slice theorem, shall play a crucial role in what is to come.

Definition 3.1.1. *Let $f \in \mathcal{S}(\mathbb{R}^d)$. The* Radon transform *of f is a function $R(f) : S^{d-1} \times \mathbb{R} \to \mathbb{C}$ given by the formula*

$$R(f)(u,s) := \int_{\langle u,x\rangle = s} f(x)dx,\ u \in S^{d-1},\ s \in \mathbb{R}.$$

The operator $R : f \mapsto R(f)$ can be extended beyond the Schwarz class to be a continuous operator that maps $L^1(\mathbb{R}^d)$ into $L^1(\mathbb{R})$ uniformly in $u \in S^{d-1}$ [60]. It is thus natural to consider the Fourier transform's relation to the Radon transform. The Fourier slice theorem describes a basic property of this relation. The notation $R_u(f) = R(f)(u,\cdot)$ shall be used in what follows.

Theorem 3.1.2. *(Fourier slice theorem) Let $f \in L^1(\mathbb{R}^d)$. Then the Fourier transform of f and $R_u(f)$ are related in the following way:*

$$\widehat{R_u(f)}(\gamma) = \hat{f}(\gamma u).$$

One of the tools used in the construction of Grafakos and Sansing is the *weighting* of a base function in the frequency domain:

Definition 3.1.3. *For $\alpha > 0$, define the* differential operator D_α of order α *for functions $h : \mathbb{R} \to \mathbb{C}$ to be:*

$$D_\alpha(h) := (\hat{h}(\gamma)|\gamma|^\alpha)^\vee.$$

Here, we generalize the classical interplay between polynomial multiplication in the frequency domain and differentiation in the time domain. This definition naturally extends to real, non-negative multi-indices $\alpha = (\alpha_1, ..., \alpha_d)$ for functions $h : \mathbb{R}^d \to \mathbb{C}$.

Recall that for $g \in L^2(\mathbb{R})$, $g^{m,t} \in L^2(\mathbb{R})$ is the function defined by:

$$g^{m,t}(s) = e^{2\pi i m(s-t)} g(s-t) = T_t M_{-m} g(s),\ m, t \in \mathbb{R}.$$

Definition 3.1.4. *Let $g \in \mathcal{S}(\mathbb{R})$ be a $\mathbb{R}$-valued, non-zero window function. For $d \geq 1$ an integer and $m, t \in \mathbb{R}$, we define:*

$$G^{m,t}(s) := D_{\frac{d-1}{2}}(g^{m,t})(s) = (\widehat{g^{m,t}}(\gamma)|\gamma|^{\frac{d-1}{2}})^{\vee}(s),\ s \in \mathbb{R}.$$

The weighted Gabor ridge functions *are*

$$G_{m,t,u}(x) := G^{m,t}(\langle u, x\rangle) = (\widehat{g^{m,t}}(\gamma)|\gamma|^{\frac{d-1}{2}})^{\vee}(\langle u, x\rangle),\ x \in \mathbb{R}^d.$$

The convention of using capital letters to refer to the weighted Gabor ridge function shall be maintained throughout this section [47].

It is clear that a weighted Gabor ridge function is constant along any hyperplane $\{x \in \mathbb{R}^d \mid \langle u, x\rangle = C\}$ for u, C constant. Along the direction u, these functions modulate like a one dimensional Gabor function. The weighting in the Fourier domain allows us to reconstruct a signal f from the coefficients $\{\langle f, G_{m,t,u}\rangle\}_{m,t\in\mathbb{R},u\in S^{d-1}}$. Weighted Gabor ridge functions yield a continuous reproducing formula that is analogous to (2.5), with an additional integral to account for the directional character of the ridge function construction.

Theorem 3.1.5. *Let $g, \psi \in \mathcal{S}(\mathbb{R})$ be two window functions such that $\langle g, \psi\rangle \neq 0$. Suppose $f \in L^1(\mathbb{R}^d)$ and $\hat{f} \in L^1(\widehat{\mathbb{R}}^d)$. Then:*

$$f = \frac{1}{2\langle g, \psi\rangle} \int_{S^{d-1}} \int_{\mathbb{R}} \int_{\mathbb{R}} \langle f, G_{m,t,u}\rangle \Psi_{m,t,u} dm\ dt\ du,$$

with convergence in the pointwise almost everywhere sense.

Here, we are recovering our original signal f from the coefficients $\{\langle f, G_{m,t,u}\rangle\}_{m,t\in\mathbb{R},u\in S^{d-1}}$; this is a continuous reproducing formula for weighted Gabor ridge system.

Note that without the weights in the Fourier domain, the reconstruction formula does not hold [47] . Indeed, let $g_{m,t,u}(x) = g^{m,t}(\langle u, x\rangle)$, and R^* the adjoint of the Radon transform:

$$R^*g(x) = \int_{S^{d-1}} g(\langle u, x\rangle, u)du.$$

Theorem 3.1.6. *Let $f \in L^1(\mathbb{R}^d), \hat{f} \in L^p(\mathbb{R}^d)$ for some $1 < p < d$. Then for $g, \psi \in \mathcal{S}(\mathbb{R})$ with $\langle g, \psi\rangle \neq 0$, we have the following identity:*

$$R^*(R(f)) = \frac{1}{\langle g, \psi\rangle} \int_{S^{d-1}} \int_{\mathbb{R}} \int_{\mathbb{R}} \langle f, g_{m,t,u}\rangle \psi_{m,t,u} dm\ dt\ du.$$

Thus, instead of reconstructing f itself, we reconstruct the *weighted back-projection* $R^*(R(f))$. In general, $R^*(R(f)) \neq f$. This is the root of many issues in tomographic applications based on Radon transforms. There is also a Parseval-type formula for Gabor ridge functions:

Theorem 3.1.7. *Let f, g be as above. There exists a constant C_g, depending only on g, such that:*

$$\|f\|^2_{L^2(\mathbb{R}^d)} = C_g \int_{S^{d-1}} \int_{\mathbb{R}} \int_{\mathbb{R}} |\langle f, G_{m,t,u}\rangle|^2 dm\ dt\ du.$$

We note that Theorem 3.1.7 expresses the fact that the Gabor ridge functions $\{G_{m,t,u}\}_{(m,t,u)\in\mathbb{R}\times\mathbb{R}\times S^{d-1}}$ form a *continuous frame* for $f \in L^1(\mathbb{R}^d) \cap \widehat{L^1(\mathbb{R}^d)}$, with the usual inner product.

In the case of d-dimensional Gabor ridge systems, the continuous frame is indexed by the set $S^{d-1} \times \mathbb{R} \times \mathbb{R}$. In general, this set does not admit a non-trivial

group structure. Indeed, the only spheres which admit a non-trivial group structure are S^0, S^1 and S^3 [62]. This means the co-orbit theory of Feichtinger and Gröchenig [41] does not apply easily. We shall prove some results regarding the discretization problem for Gabor ridge systems, but using ideas from classical Gabor theory, rather than coorbit theory.

As a starting point, consider the following *semi-discrete* representation formula for Gabor ridge systems [47]:

Theorem 3.1.8. *There exist $g, \psi \in \mathcal{S}(\mathbb{R})$ and $\alpha, \beta > 0$ such that for all $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, we have have:*

$$A\|f\|_2^2 \leq \int_{S^{d-1}} \sum_{m\in\mathbb{Z}} \sum_{t\in\mathbb{Z}} |\langle f, G_{\alpha m,\beta t,u}\rangle|^2 du \leq B\|f\|_2^2, \tag{3.1}$$

where the constants A, B depend only on g, α, β and du is the Lebesgue measure on S^{d-1}. Moreover, for this choice of g, ψ, we have:

$$f = \frac{1}{2} \int_{S^{d-1}} \sum_{m\in\mathbb{Z}} \sum_{t\in\mathbb{Z}} \langle f, G_{\alpha m,\beta t,u}\rangle \Psi_{\alpha m,\beta t,u} du.$$

3.2 Negative Results and a Toy Example

The representation (3.1) is called semi-discrete because there are two discrete sums, but also an integral over S^{d-1}. A first, somewhat naive, approach to developing a fully discrete directional Gabor system would be to start with the semi-discrete representation and choose certain points on the unit sphere, i.e. to discretize the integral in (3.1) as a finite sum. We show this approach never generates a frame for any reasonable class of functions.

We begin by noting that once discretized, the frequency weights used in the weighted Gabor ridge construction are not needed. So, we seek a discrete system of the form:

$$\{g_{m,t,u}\}_{(m,t,u)\in\Lambda},$$

along with a space of functions for which this set will be a discrete frame. The nomenclature used to describe such a system shall be *discrete directional Gabor frame*.

A first, somewhat naive, approach to developing a discrete directional Gabor frame could be to start with the semi-discrete representation and choose certain points on the unit circle at which to sample. This would essentially be replacing the integral in (3.1) with a finite sum. To investigate this approach, we consider systems of the form:

$$\{g_{\alpha m,\beta t,u}\}_{(m,t,u)\in\mathbb{Z}^2\times Q},$$

where α, β are fixed real numbers and $Q \subset S^{d-1}$ is a fixed set of finite cardinality. Note that since S^{d-1} is compact, Q is finite if and only if it is discrete. We show that such a system cannot be a frame for $L^2(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$.

Theorem 3.2.1. *Let $Q \subset S^{d-1}$ be finite. Then there exists $f \in \mathcal{S}(\mathbb{R}^d)$ such that $\|f\|_{L^2(\mathbb{R}^d)} = 1$ and*

$$\|R_u(f)\|^2_{L^2(\mathbb{R})} = 0, \ \forall u \in Q.$$

Proof. Note that

$$\|R_u(f)\|^2_{L^2(\mathbb{R})} = \int_{\mathbb{R}} \left| \widehat{R_u(f)}(\gamma) \right|^2 d\gamma$$

$$= \int_{\mathbb{R}} |\hat{f}(\gamma u)|^2 d\gamma.$$

These equalities follow from Parseval's identity and the Fourier slice theorem, respectively. We can choose $f \in \mathcal{S}(\mathbb{R}^d)$ of unit L^2 norm whose Fourier transform is supported off of the lines $\{\gamma u\}_{\gamma \in \mathbb{R}, u \in Q}$. Such an f has the desired properties.

□

Corollary 3.2.2. *Let $g \in W$ be an element of the classical Wiener class of functions, and let $\alpha, \beta > 0$, $Q \subset S^1$ be a fixed set of finite cardinality. Then the system*

$$\left\{g_{\alpha m, \beta t, u}\right\}_{(m,t,u) \in \mathbb{Z}^2 \times Q}$$

is not a frame for any subspace of $L^2(\mathbb{R}^2)$ which contains $S(\mathbb{R}^2)$. In fact, it cannot even be Bessel.

Proof. Note that $\langle f, g_{\alpha m, \beta t, u} \rangle = \langle R_u(f), g^{\alpha m, \beta t} \rangle$. Since $g \in W$, classical Gabor theory [49] guarantees that we may find α_1, β_1 such that $\{g^{\alpha_1 m, \beta_1 t}\}_{(m,t) \in \mathbb{Z}^2}$ is a Bessel system for $L^2(\mathbb{R})$. Choose α_1, β_1 sufficiently small so that α is an integer multiple of α_1 and β is an integer multiple of β_1. Then:

$$\begin{aligned}
&\sum_{u \in Q} \sum_{m \in \mathbb{Z}} \sum_{t \in \mathbb{Z}} |\langle f, g_{\alpha m, \beta t, u} \rangle|^2 \\
\leq &\sum_{u \in Q} \sum_{m \in \mathbb{Z}} \sum_{t \in \mathbb{Z}} |\langle f, g_{\alpha_1 m, \beta_1 t, u} \rangle|^2 \\
= &\sum_{u \in Q} \sum_{m \in \mathbb{Z}} \sum_{t \in \mathbb{Z}} |\langle R_u(f), g^{\alpha_1 m, \beta_1 t} \rangle|^2 \\
\leq &B \sum_{u \in Q} \|R_u(f)\|^2_{L^2(\mathbb{R})}.
\end{aligned}$$

By Theorem 3.2.1, we may find $\tilde{f} \in \mathcal{S}(\mathbb{R}^2)$, $\|\tilde{f}\|_{L^2(\mathbb{R}^2)} = 1$, such that

$$\sum_{u \in Q} \|R_u(\tilde{f})\|^2_{L^2(\mathbb{R})} = 0.$$

Noting that this forces

$$\sum_{u \in Q} \sum_{m \in \mathbb{Z}} \sum_{t \in \mathbb{Z}} |\langle \tilde{f}, g_{\alpha m, \beta t, u} \rangle|^2 = 0,$$

the result is shown. □

An even stronger negative result is possible. In essence, we cannot have a discrete directional Gabor system for any function space that includes $\mathcal{S}(\mathbb{R}^d)$.

Theorem 3.2.3. *Let $\{g_{m,t,u}\}_{(m,t,u) \in \Lambda}$ be any discrete directional Gabor system. Then there exists a sequence of Schwartz function $\{\phi_n\}_{n=1}^{\infty} \in \mathcal{S}(\mathbb{R}^d)$ such that*

1. $\lim_{n \to \infty} \|\phi_n\|_2 = 0.$
2. $\sum_{(m,t,u) \in \Lambda} |\langle \phi_n, g_{m,t,u} \rangle|^2 \geq 1, \forall n.$

Thus, $\{g_{m,t,u}\}_{(m,t,u) \in \Lambda}$ cannot be a frame for any function space containing $\mathcal{S}(\mathbb{R}^d)$.

Proof. We will show that ϕ_n may be chosen with the property that

$$|\langle \phi_n, g_{m,t,u} \rangle|^2 \geq 1,$$

for a fixed $(m, t, u) \in \Lambda$. By rotating ϕ_n as needed, we may assume WLOG that $u = (1, 0, ..., 0) \in S^{d-1}$. Then we compute:

$$|\langle \phi_n, g_{m,t,u} \rangle|^2 = |\langle R_u \phi_n, g^{m,t} \rangle|$$

$$= |\langle \hat{\phi}_n(u\gamma), \hat{g}^{m,t}(\gamma) \rangle|$$
$$= \left| \int_{\hat{\mathbb{R}}} \hat{\phi}_n(u\gamma) \overline{\hat{g}^{m,t}(\gamma)} d\gamma \right|.$$

Now, let us choose ϕ_n such that

$$\hat{\phi_n}(\gamma) = \hat{\phi_n}(\gamma_1, \gamma_2, ..., \gamma_d) = \hat{\psi_n}(\gamma_1) \cdot \hat{\eta_n}(\gamma_2, ..., \gamma_d),$$

where $\hat{\psi}_n \in \mathcal{C}^\infty(\mathbb{R})$ is chosen such that $\|\psi_n\|_2$ is uniformly bounded in n, and such that:

$$\left| \int_{\hat{\mathbb{R}}} \hat{\phi}_n(u\gamma) \overline{\hat{g}^{m,t}(\gamma)} | d\gamma \right| = \left| \int_{\hat{\mathbb{R}}} \hat{\psi}_n(\gamma) \overline{\hat{g}^{m,t}(\gamma)} d\gamma \right| = 1,$$

and choose $\hat{\eta}_n \in \mathcal{C}_c^\infty(\mathbb{R}^{d-1})$ such that $\|\hat{\eta}_n\|_2 \to 0$. This implies that $\|\phi_n\|_2 \to 0$, giving the desired result.

□

So, it is clear that in order to discretize a directional Gabor system to acquire a discrete frame, infinitely many directions on the sphere must be incorporated, and our space of functions must be chosen very specifically. In particular, the space of functions must not contain all of $\mathcal{S}(\mathbb{R}^d)$. Note also that no weights are used in this construction; these weights were introduced to account for the change of variables in the continuous integrals, and are not needed for discrete sums.

3.3 A Toy Example

We now give a *positive result*, showing that it is possible to produce a discrete directional Gabor system. For clarity, we take $d = 2$, i.e. we consider discrete directional Gabor systems for functions $\mathbb{R}^2$.

Definition 3.3.1. *Let $\mathcal{V}$ be the subset of $L^2(\mathbb{R}^2)$ of functions supported in $B_{\frac{1}{2}}(0)$.*

Theorem 3.3.2. *Let $g(x) = \chi_{[-\frac{1}{2},\frac{1}{2}]}(x)$. Let $\Gamma \subset S^1 \times \mathbb{R}$ be such that the mapping $\psi : \Gamma \to \mathbb{Z}^2$ given by $(u, m) \mapsto mu$ is a bijection. Set $\Lambda = \{(m, t, u)|(u, m) \in \Gamma, t \in \mathbb{Z}\}$. Then we have:*

$$\sum_{(m,t,u)\in\Lambda} |\langle f, g_{m,t,u}\rangle|^2 = \|f\|_2^2,$$

for all $f \in \mathcal{V}$.

Proof. We compute:

$$\begin{aligned}
\sum_{(m,t,u)\in\Lambda} |\langle f, g_{m,t,u}\rangle|^2 &= \sum_{(u,m)\in\Gamma} |\langle f, g_{m,0,u}\rangle|^2 \\
&= \sum_{(u,m)\in\Gamma} \left| \int_{\mathbb{R}^2} f(x)\chi_{[-\frac{1}{2},\frac{1}{2}]}(\langle u, x\rangle)e^{2\pi i m\langle u,x\rangle} dx \right|^2 \\
&= \sum_{(u,m)\in\Gamma} \left| \int_{B_{\frac{1}{2}}(0)} f(x)e^{2\pi i\langle mu,x\rangle} dx \right|^2 \\
&= \sum_{n\in\mathbb{Z}^2} |\hat{f}(n)|^2 \\
&= \|f\|_2^2.
\end{aligned}$$

□

Note that in this proof, we make no use of the translations in the directional Gabor system $g_{m,t,u}$, since we consider functions f supported only on the ball $B_{\frac{1}{2}}(0)$. It is of interest if this parameter set can be exploited non-trivially, to produce more interesting examples of discrete directional Gabor systems. Moreover, the support of g could be extended, and the support of functions in $\mathcal{V}$ increased, at the cost of having a frame which is not Parseval.

3.4 Sufficient Conditions for a Discrete System

Definition 3.4.1. *Let $\mathcal{U}$ be the subset of $\mathcal{S}(\mathbb{R}^2)$ of functions supported in $[-\frac{1}{2}, \frac{1}{2}]^2$.*

We emulate the approach in [61], in which the authors use the theory of absolutely convergent Fourier series to provide necessary and sufficient conditions for certain classes of functions to be frames. Among the classes of functions they characterize are classical discrete Gabor and discrete wavelet systems.

Definition 3.4.2. *For $g \in L^2(\mathbb{R})$, $(u, m) \in S^1 \times \mathbb{R}$, define a function on $\mathbb{R}^2$ by*

$$g_{(u,m)}(x) := e^{2\pi i m \langle u,x \rangle} g(\langle u, x \rangle).$$

For $h \in L^2(\mathbb{R}^2)$ and $n \in \mathbb{Z}$, define

$$T_n^{(u,m)} h(x) := h(x - \alpha_{u,m} n u),$$

where $\alpha_{u,m}$ is a constant assigned to the parameter (u, m).

In the arguments related to the existence of a discrete directional Gabor system, it is sufficient to take $\alpha_{u,m} = 1$, $\forall (u, m)$. However, we shall state certain lemmas as generally as possible.

Lemma 3.4.3. *Let $f \in \mathcal{S}(\mathbb{R}^2)$. Then for fixed $(u, m) \in S^1 \times \mathbb{R}$,*

$$\sum_{n \in \mathbb{Z}} |\langle f, T_n^{(u,m)} g_{(u,m)} \rangle|^2 = \frac{1}{|\alpha_{u,m}|^2} \int_{\frac{1}{\alpha_{u,m}}[0,1]} \left| \sum_{n \in \mathbb{Z}} \hat{f}\left(\left(\gamma + \frac{n}{\alpha_{u,m}} \right) u \right) \overline{\hat{g}\left(\gamma + \frac{n}{\alpha_{u,m}} - m \right)} \right|^2 d\gamma.$$

Proof. Notice that:

$$T_n^{(u,m)} g_{(u,m)}(x) = e^{2\pi i m \langle u,x \rangle - m i \alpha_{u,m} n} g(\langle u, x \rangle - \alpha_{u,m} n)$$

$$= g^{m,\alpha_{u,m}n}(\langle u, x\rangle)$$

$$= g_{m,\alpha_{u,m}n,u}(x).$$

So, by Lemma 1 in [47],

$$\begin{aligned}\langle f, T_n^{(u,m)} g_{(u,m)}\rangle &= \langle R_u f, g^{m,\alpha_{u,m}n}\rangle \\ &= \int_{\hat{\mathbb{R}}} \hat{f}(\gamma u)\overline{\widehat{g^{m,\alpha_{u,m}n}}(\gamma)}d\gamma \\ &= \int_{\hat{\mathbb{R}}} \hat{f}(\gamma u)\overline{\hat{g}(\gamma - m)}e^{-2\pi i\alpha_{u,m}n\gamma}d\gamma.\end{aligned}$$

Now, periodize this integral with respect to $\alpha_{u,m}$:

$$= \frac{1}{\alpha_{u,m}} \int_{\frac{1}{\alpha_{u,m}}[0,1]} \sum_{m\in\mathbb{Z}} \hat{f}\left(\left(\gamma + \frac{m}{\alpha_{u,m}}\right)u\right)\overline{\hat{g}\left(\gamma - m + \frac{m}{\alpha_{u,m}}\right)}e^{-2\pi i\alpha_{u,m}n\gamma}d\gamma.$$

Thus,

$$\begin{aligned}&\sum_{n\in\mathbb{Z}} |\langle f, T_n^{(u,m)} g_{(u,m)}\rangle|^2 \\ =&\sum_{n\in\mathbb{Z}} \left|\frac{1}{\alpha_{u,m}} \int_{\frac{1}{\alpha_{u,m}}[0,1]} \sum_{m\in\mathbb{Z}} \hat{f}\left(\left(\gamma + \frac{m}{\alpha_{u,m}}\right)u\right)\overline{\hat{g}\left(\gamma - m + \frac{m}{\alpha_{u,m}}\right)}e^{-2\pi i\alpha_{u,m}n\gamma}d\gamma\right|^2 \\ =&\frac{1}{|\alpha_{u,m}|^2} \int_{\frac{1}{\alpha_{u,m}}[0,1]} \left|\sum_{n\in\mathbb{Z}} \hat{f}\left(\left(\gamma + \frac{n}{\alpha_{u,m}}\right)u\right)\overline{\hat{g}\left(\gamma - m + \frac{n}{\alpha_{u,m}}\right)}\right|^2 d\gamma.\end{aligned}$$

To transition from the penultimate to the last line, notice we are summing the Fourier coefficients with respect to the orthonormal basis $\{e^{2\pi i\alpha_{u,m}n\gamma}\}_{n\in\mathbb{Z}}$ of the function

$$\gamma \mapsto \sum_{m\in\mathbb{Z}} \hat{f}\left(\left(\gamma + \frac{m}{\alpha_{u,m}}\right)u\right)\overline{\hat{g}\left(\gamma - m + \frac{m}{\alpha_{u,m}}\right)}.$$

This is equal to the square of the L^2 norm of the function by Parseval's formula, which appears on the final line.

□

Let $f \in \mathcal{U}$, $\Lambda \subset S^1 \times \mathbb{R}$ as before. We proceed by considering the function

$$w(x) := \sum_{(u,m)\in\Lambda} \sum_{n\in\mathbb{Z}} |\langle T_x f, T_n^{(u,m)} g_{(u,m)}\rangle|^2.$$

We shall show that w may be written as an absolutely convergent Fourier series, and exploit the properties of this special class of functions.

Definition 3.4.4. *For $(u,m) \in S^1 \times \mathbb{R}$ fixed, set*

$$H(x) := \sum_{n\in\mathbb{Z}} |\langle T_x f, T_n^{(u,m)} g_{(u,m)}\rangle|^2.$$

Lemma 3.4.5. *$H(x)$ is a trigonometric series of the form*

$$H(x) = \sum_{k\in\mathbb{Z}} \hat{H}(k) e^{2\pi i \frac{k}{\alpha_{u,m}}\langle u,x\rangle},$$

where the coefficients are given by

$$\hat{H}(k) = \frac{1}{\alpha_{u,m}} \int_{\hat{\mathbb{R}}} \hat{f}(\gamma u)\overline{\hat{g}(\gamma - m)}\,\overline{\hat{f}\left(\left(\gamma + \frac{k}{\alpha_{u,m}}\right)u\right)}\,\hat{g}\left(\gamma + \frac{k}{\alpha_{u,m}} - m\right) d\gamma.$$

Proof. Substituting $T_x f$ into Lemma 3.4.3, we compute:

$$\begin{aligned}
H(x) =& \frac{1}{|\alpha_{u,m}|^2} \int_{\frac{1}{\alpha_{u,m}}[0,1]} \left| \sum_{n\in\mathbb{Z}} \hat{f}\left(\left(\gamma + \frac{n}{\alpha_{u,m}}\right)u\right) \overline{\hat{g}(\gamma + \frac{n}{\alpha_{u,m}} - m)} e^{-2\pi i(\gamma + \frac{n}{\alpha_{u,m}})\langle u,x\rangle} \right|^2 d\gamma \\
=& \frac{1}{|\alpha_{u,m}|^2} \int_{\frac{1}{\alpha_{u,m}}[0,1]} \left| e^{-2\pi i\gamma\langle u,x\rangle} \sum_{n\in\mathbb{Z}} \hat{f}\left(\left(\gamma + \frac{n}{\alpha_{u,m}}\right)u\right) \overline{\hat{g}\left(\gamma + \frac{n}{\alpha_{u,m}} - m\right)} e^{-2\pi i \frac{n}{\alpha_{u,m}}\langle u,x\rangle} \right|^2 d\gamma \\
=& \frac{1}{|\alpha_{u,m}|^2} \int_{\frac{1}{\alpha_{u,m}}[0,1]} \left| \sum_{n\in\mathbb{Z}} \hat{f}\left(\left(\gamma + \frac{n}{\alpha_{u,m}}\right)u\right) \overline{\hat{g}\left(\gamma + \frac{n}{\alpha_{u,m}} - m\right)} e^{-2\pi i \frac{n}{\alpha_{u,m}}\langle u,x\rangle} \right|^2 d\gamma \\
=& \frac{1}{|\alpha_{u,m}|^2} \int_{\frac{1}{\alpha_{u,m}}[0,1]} \sum_{\ell,m\in\mathbb{Z}} e^{-2\pi i \frac{m}{\alpha_{u,m}}\langle u,x\rangle} \hat{f}\left(\left(\gamma + \frac{m}{\alpha_{u,m}}\right)u\right) \overline{\hat{g}\left(\gamma + \frac{m}{\alpha_{u,m}} - m\right)} \\
& \times e^{-2\pi i \frac{\ell}{\alpha_{u,m}}\langle u,x\rangle} \hat{f}\left(\left(\gamma + \frac{\ell}{\alpha_{u,m}}\right)u\right) \overline{\hat{g}\left(\gamma + \frac{\ell}{\alpha_{u,m}} - m\right)} d\gamma \\
=& \frac{1}{|\alpha_{u,m}|^2} \int_{\frac{1}{\alpha_{u,m}}[0,1]} \sum_{\ell,m\in\mathbb{Z}} e^{2\pi i \frac{\ell - m}{\alpha_{u,m}}\langle u,x\rangle} \hat{f}\left(\left(\gamma + \frac{m}{\alpha_{u,m}}\right)u\right) \overline{\hat{g}\left(\gamma + \frac{m}{\alpha_{u,m}} - m\right)}
\end{aligned}$$

$$
\begin{aligned}
&\times \hat{f}\left(\left(\gamma+\frac{\ell}{\alpha_{u,m}}\right)u\right)\overline{\hat{g}\left(\gamma+\frac{\ell}{\alpha_{u,m}}-m\right)}d\gamma \\
=&\frac{1}{|\alpha_{u,m}|^2}\int_{\frac{1}{\alpha_{u,m}}[0,1]}\sum_{k,m\in\mathbb{Z}} e^{2\pi i\frac{k}{\alpha_{u,m}}\langle u,x\rangle}\hat{f}\left(\left(\gamma+\frac{m}{\alpha_{u,m}}\right)u\right)\overline{\hat{g}\left(\gamma+\frac{m}{\alpha_{u,m}}-m\right)} \\
&\times \hat{f}\left(\left(\gamma+\frac{k+m}{\alpha_{u,m}}\right)u\right)\overline{\hat{g}\left(\gamma+\frac{k+m}{\alpha_{u,m}}-m\right)}d\gamma \\
=&\frac{1}{|\alpha_{u,m}|^2}\int_{\hat{\mathbb{R}}}\sum_{k\in\mathbb{Z}}\hat{f}(\gamma u)\overline{\hat{g}(\gamma-m)}\overline{\hat{f}\left(\left(\gamma+\frac{k}{\alpha_{u,m}}\right)u\right)}\hat{g}\left(\gamma+\frac{k}{\alpha_{u,m}}-m\right)e^{2\pi i\frac{k}{\alpha_{um}}\langle u,x\rangle}d\gamma.
\end{aligned}
$$

To go from the penultimate to ultimate line, we sum over m and exploit the periodization. Since f is Schwartz class, we may interchange the integral and sum. This gives the result.

□

Lemma 3.4.6. *Let $g \in L^2(\mathbb{R})$ have supp$(\hat{g}) \subset K$ compact. Suppose $\Lambda = \{(m,u)\} \subset \mathbb{R}\times S^1$ is such that the map $\psi:(u,m)\mapsto mu$ is a bijection between Λ and $\mathbb{Z}^2$. Set*

$$w(x) := \sum_{(u,m)\in\Lambda}\sum_{n\in\mathbb{Z}}|\langle T_x, T_n^{(u,m)}g_{(u,m)}\rangle|^2.$$

Then $w(x)$ is an absolutely convergent Fourier series.

Proof. Set $w_{u,m}$ to be the sum over fixed (u,m):

$$
\begin{aligned}
w_{u,m}(x) =&\sum_{n\in\mathbb{Z}}|\langle T_x, T_n^{u,m}g_{(u,m)}\rangle|^2 \\
=&\sum_{k\in\mathbb{Z}}\hat{H}_{(u,m)}(k)e^{2\pi ik\langle u,x\rangle},
\end{aligned}
$$

where the coefficients of this trigonometric series are:

$$\hat{H}_{(u,m)}(k) = \int_{\hat{\mathbb{R}}}\hat{f}(\gamma u)\overline{\hat{g}(\gamma-m)}\overline{\hat{f}((\gamma+k)u)}\hat{g}(\gamma+k-m)d\gamma.$$

We want to show $\{\hat{H}_{(u,m)}(k)\}_{\Lambda\times\mathbb{Z}}\in\ell^1(\Lambda\times\mathbb{Z})$, that is,

$$\sum_{(u,m)\in\Lambda}\sum_{k\in\mathbb{Z}}\left|\int_{\hat{\mathbb{R}}}\hat{f}(\gamma u)\overline{\hat{g}(\gamma-m)}\overline{\hat{f}((\gamma+k)u)}\hat{g}(\gamma+k-m)d\gamma\right|<\infty.$$

Well,

$$\sum_{(u,m)\in\Lambda}\sum_{k\in\mathbb{Z}}\left|\int_{\hat{\mathbb{R}}}\hat{f}(\gamma u)\overline{\hat{g}(\gamma-m)}\overline{\hat{f}((\gamma+k)u)}\hat{g}(\gamma+k-m)d\gamma\right|$$

$$=\sum_{(u,m)\in\Lambda}\sum_{k\in\mathbb{Z}}\left|\int_{K}\hat{f}((\gamma+m)u)\overline{\hat{g}(\gamma)}\overline{\hat{f}((\gamma+k+m)u)}\hat{g}(\gamma+k)d\gamma\right|$$

$$\leq\left(\sum_{(u,m)\in\Lambda}\sum_{k\in\mathbb{Z}}\int_{K}\left|\hat{f}((\gamma+k+m)u)\hat{g}(\gamma)\chi_K(\gamma+k)\right|^2d\gamma\right)^{\frac{1}{2}}$$

$$\times\left(\sum_{(u,m)\in\Lambda}\sum_{k\in\mathbb{Z}}\int_{K}\left|\hat{f}((\gamma+m)u)\hat{g}(\gamma+k)\right|^2d\gamma\right)^{\frac{1}{2}}.$$

Let us consider the first factor. By positivity of the summands, we may interchange the order of summation integration. For a fixed $k\in\mathbb{Z}$, consider:

$$\sum_{(u,m)\in\Lambda}\int_{K}\left|\hat{f}((\gamma+k+m)u)\hat{g}(\gamma)\chi_K(\gamma+k)\right|^2d\gamma$$

$$=\int_{K}\sum_{(u,m)\in\Lambda}\left|\hat{f}((\gamma+k+m)u)\hat{g}(\gamma)\chi_K(\gamma+k)\right|^2d\gamma.$$

Let

$$F(k)=F_\gamma(k):=\sum_{(u,m)\in\Lambda}|\hat{f}((\gamma+k+m)u)|^2.$$

If $\forall k, F(k)<\infty$, then $\hat{g}\in L^2(\hat{\mathbb{R}})$ together with K compact gives the boundedness of the first factor. To see $F(k)<\infty,\ \forall k$, notice that since $f\in\mathcal{S}(\mathbb{R}^2)$, for all $\alpha>0$, there exists a constant $C>0$ such that:

$$|\hat{f}(\gamma u)|\leq\frac{C}{\|\gamma\|^\alpha}. \tag{3.2}$$

Then let $M>0$ be such that $K\subset B_M(0)$. Then:

$$\sum_{(u,m)\in\Lambda}|\hat{f}((\gamma+k+m)u)|^2$$

$$= \sum_{(mu)\in B_M(0)} |\hat{f}(\gamma+k+m)u|^2 + \sum_{(mu)\notin B_M(0)} |\hat{f}((\gamma+k+m)u)|^2.$$

The first sum is finite because there are only finitely many lattice points in $B_M(0)$. To see the second sum is finite, use (3.2). Indeed, for all $mu \notin B_M(0)$:

$$\begin{aligned} |\hat{f}(\gamma u + ku + mu)| \leq & \frac{C}{|\|mu\| - \|\gamma u + ku\||^\alpha} \\ \leq & \frac{C}{|\|mu\| - (M+k)|^\alpha} \\ \leq & \frac{C'}{\|mu\|^\alpha}, \end{aligned}$$

which is clearly ℓ^2 summable over the integer lattice for α sufficiently large.

Since $\chi_K(K+k) \equiv 0$ for k sufficiently large, the first factor is finite. The second factor is finite by the same argument. Thus, the coefficients of our series are summable and the series is absolutely convergent, i.e. w is an absolutely convergent Fourier series.

□

Note that with Λ as above, the sampling set $\Lambda \times \mathbb{Z}$ is indeed discrete. We are now prepared to give a sufficient condition for a discrete Gabor ridge system.

Theorem 3.4.7. *Let $g \in L^2(\mathbb{R})$ be such that $\operatorname{supp}(\hat{g}) \subset K \subset (-\frac{1}{4}, \frac{1}{4})$ compact. Then with Λ as above, $\{g_{k,m,u}\}_{k\in\mathbb{Z},(u,m)\in\Lambda}$ is a frame for $\mathcal{U}$.*

Proof. Let

$$\begin{aligned} w(x) &= \sum_{(u,m)} \sum_{n\in\mathbb{Z}} |\langle T_x f, T_n^{(u,m)} g_{(u,m)} \rangle|^2 \\ &= \sum_{(u,m)} \sum_{k\in\mathbb{Z}} \hat{H}_{(u,m)}(k) e^{2\pi i \langle ku, x\rangle}, \end{aligned}$$

where

$$\hat{H}_{(u,m)}(k) = \int_{\mathbb{R}} \hat{f}((\gamma+m)u)\overline{\hat{g}(\gamma)}\overline{\hat{f}((\gamma+k+m)u)}\hat{g}(\gamma+k)d\gamma.$$

Lemma 3.4.6 implies $w(x)$ is an absolutely convergent Fourier series, and thus continuous. It thus suffices to show that there exists $0 < A \leq B < \infty$ such that:

$$A\|f\|_2^2 \leq w(0) \leq B\|f\|_2^2.$$

Notice that if $k \neq 0$, then:

$$\hat{H}_{(u,m)}(k) = \int_{\mathbb{R}} \hat{f}((\gamma+m)u)\overline{\hat{g}(\gamma)}\overline{\hat{f}((\gamma+k+m)u)}\hat{g}(\gamma+k)d\gamma = 0,$$

due to the fact that $[-\frac{1}{4}+k, \frac{1}{4}+k] \cap [-\frac{1}{4}, \frac{1}{4}] = \emptyset,\ \forall k \neq 0$. Thus,

$$\begin{aligned}
w(0) &= \sum_{(u,m)\in\Lambda} \hat{H}_{(u,m)}(0) \\
&= \sum_{(u,m)\in\Lambda} \int_{\mathbb{R}} \hat{f}((\gamma+m)u)\overline{\hat{g}(\gamma)}\overline{\hat{f}((\gamma+m)u)}\hat{g}(\gamma)d\gamma \\
&= \sum_{(u,m)\in\Lambda} \int_{\hat{\mathbb{R}}} |\hat{f}((\gamma+m)u)|^2|\hat{g}(\gamma)|^2 d\gamma \\
&= \sum_{(u,m)\in\Lambda} \int_{[-\frac{1}{4},\frac{1}{4}]} |\hat{f}(\gamma+m)u)|^2|\hat{g}(\gamma)|^2 d\gamma \\
&= \int_{[-\frac{1}{4},\frac{1}{4}]} |\hat{g}(\gamma)|^2 \sum_{(u,m)\in\Lambda} |\hat{f}((\gamma+m)u)|^2 d\gamma.
\end{aligned}$$

Notice that

$$|\hat{f}((\gamma+m)u)|^2 = \int_{[-\frac{1}{2},\frac{1}{2}]^2} f(x)e^{-2\pi i\langle(m+\gamma)u,x\rangle}dx.$$

Now, because $\{mu\}_{(u,m)\in\Lambda} = \mathbb{Z}^2$ and $\|(m+\gamma)u - mu\|_2 = |\gamma| < \frac{1}{4}$, it follows from Kadec's theorem [95] that for each fixed $\gamma \in K$, $\{e^{-2\pi i\langle(m+\gamma)u,\cdot\rangle}\}_{(m,u)\in\Lambda}$ is a

Riesz basis for $L^2([-\frac{1}{2},\frac{1}{2}]^2)$, in particular for $\mathcal{U}$. Thus, for each $\gamma \in K$:

$$A_\gamma \|f\|_2^2 \leq \sum_{(u,m)\in\Lambda} |\hat{f}((\gamma+m)u)|^2 \leq B_\gamma \|f\|_2^2.$$

Note that because K is bounded uniformly away from $[-\frac{1}{4},\frac{1}{4}]$,

$$A := \inf_{\gamma\in K} A_\gamma > 0,$$

$$B := \sup_{\gamma\in K} B_\gamma < \infty.$$

Thus,

$$\int_{[-\frac{1}{4},\frac{1}{4}]} |\hat{g}(\gamma)|^2 \cdot B\|f\|_2^2 d\gamma \leq w(0) \leq \int_{[-\frac{1}{4},\frac{1}{4}]} |\hat{g}(\gamma)|^2 \cdot A\|f\|_2^2 d\gamma.$$

We conclude:

$$(A\|g\|_2^2)\|f\|_2^2 \leq w(0) \leq (B\|g\|_2^2)\|f\|_2^2,$$

as desired. □

This result relies on Kadec's theorem, which assumes nothing about the type of perturbation in the sampling set, only that the pairwise distances are sufficiently small. In this problem, we have rather precise knowledge concerning the perturbation, which may render Kadec's theorem too crude. This could be explored, to allow for $\hat{g}$ of larger support. Note that the frame bounds A, B can be precisely estimated, in accordance with Kadec's theorem.

Having shown the existence of discrete Gabor frames, along with precise estimates on the associated frame bounds, it is of interest to consider numerical implementations. Such implementations could be applied to many of the problems

in image processing that are studied in Chapters 4,5, and 6 of this thesis. It is also of interest to investigate what class of functions are *optimally represented by directectional Gabor systems.*

Chapter 4: Image Registration with Shearlets

4.1 Introduction

The process of image registration seeks to align two or more images of approximately the same scene, acquired at different times or with different sensors [11]. A variety of scientific fields make use of image registration, including biomedical imaging [80], microscopy [23], and remote sensing [72]. The purpose of studying image registration in all of these disciplines is to develop robust, accurate, and computationally efficient algorithms to align the relevant images. This can be for comparative purposes, or as a first step in a more general program. For example, many techniques for image fusion require registered images. In particular, fusion methods based on wavelets [74] and wavelet packets [25] require registered images. Chapter 6 shall give a more thorough treatment to image fusion.

Image registration is fraught with complications. Given the broad class of variations between types of images, degree of noise present in images, and initial knowledge of misregistration, an image registration technique could perform admirably in one set of circumstances and poorly in another. As such, the pursuit of a highly flexible, robust algorithm is valuable to the communities that rely on image registration.

This ambition has been pursued at the NASA Goddard Space Flight Center for some time. An automatic image registration technique was developed there [108], [73], based on wavelets and wavelet-like pyramids. This algorithm proved useful on a variety of remotely sensed image data, but sometimes failed to be robust to the initial registration guess. More precisely, if the initial guess for registration is very far from the truth registration, the algorithm could fail to converge to the correct registration transformation. This often happens if the images are severely misregistered to begin. Indeed, most registration techniques employ an optimization algorithm that requires an initial seed value; the optimization technique aligns the images by searching for a global minimizer to a non-linear least-squares problem. If the initial guess is too far from the global minimizer, the algorithm will converge to a local minimizer, rather than the global minimizer. This is a common issue with a large class of fitting algorithms to non-linear least squares problems [3]. It is important to have an image registration algorithm that is robust to initial guess, since many geophysical applications that require image registration have severely misregistered images. Providing distinct, sparse features for the optimization algorithm to use is a natural way to increase robustness of a registration algorithm.

In the years since this wavelet-based registration technique was developed, the mathematical discipline of harmonic analysis experienced a renaissance. The wavelet transform has been generalized to a growing family of transforms emphasizing different aspects of a signal. In particular, the shearlet transform generalizes the wavelet transform by providing increased directional sensitivity [39], [70]. Shearlet Theory was extensively reviewed in Chapter 1 of this thesis. Edge-like features such as

roads, rivers, mountain ridges, and land cover boundaries are very well-emphasized by the shearlet transform, both theoretically and in practice.

Our goal was to improve the wavelet automatic registration algorithm by registering images according to their shearlet features. Given the distinct features this mathematical technique produces, our expectation was the robustness of the algorithm would be improved. Theoretically, these sparse, well-defined features should allow a poorer initial guess, and still acquire accurate convergence, even in the case of severe misregistration. We justify this heuristic in Section 4.3. Our algorithm exploits this by first registering with shearlet features, then refining the registration by registering with wavelet features. This two-stage algorithm provides strong robustness, from the shearlet stage, and strong precision, from the wavelet stage.

After describing our algorithms in detail in Section 4.4, we tested them on a variety of remotely sensed data. In Section 4.5, we synthetically generated input data for which we had perfect knowledge of the transformation between the images. These experiments are useful, but are somewhat unrealistic, because of the limited variety in feature size, shape and contrast. Thus, in Section 4.6, we experiment on real, multimodal data. The first set of experiments involved registering a shaded-relief lidar image to an optical image of approximately the same scene. The second set of experiments involved registering one band of a multispectral urban scene to a panchromatic image of approximately the same scene. The lidar-to-optical registration is of particular interest, because the information content differences between elevation data and optical radiometric images render many pixel-based registration techniques ineffective. Registration of lidar-derived data, such as vegetation height

and radiometric images, such as solar reflectance, enable novel fusion studies of land cover properties and processes. Our automatic lidar-to-optical registration technique provides a solution for these disparate data sources. This makes the results of this chapter of significant interest to remote sensing scientists in the earth sciences community.

4.2 Background on Image Registration

The process of image registration aligns two images, called an *input image* and *reference image*. The reference is understood to be fixed, and the input image is transformed to match the reference image. Image registration may be viewed as the combination of four separate sub-processes [11]:

1. Selecting an appropriate *search space* of admissible transformations. This will depend on whether the images are at the same resolution, and what type of transformations will carry the input image to the reference image, i.e. rotation-translation (RT), or polynomial warping.

2. Extracting relevant *features* to be used for matching. These could be individual pixels or groups of pixels that are known to be in correspondence between the two images, or could be global structures in the images, such as roads, buildings, rivers, and textures.

3. Selecting a *similarity metric*, in order to decide if a transformed input image closely matches the reference image. This metric should make use of the

features which are extracted from the image, be they specific pixels or global structures.

4. Selecting a *search strategy*, which is used to match the images based on maximizing or minimizing the similarity metric.

The process of image registration aligns two images, called an *input image* and *reference image*. The reference image is understood to be fixed, and the input image is transformed to match the reference image. Image registration may be viewed as the combination of four separate sub-processes [11]:

1. Selecting an appropriate *search space* of admissible transformations. This will depend on whether the images are at the same resolution, and what type of transformations will carry the input image to the reference image, i.e. rotation-translation (RT), or polynomial warping.

2. Extracting relevant *features* to be used for matching. These could be individual pixels or groups of pixels that are known to be in correspondence between the two images, or could be global structures in the images, such as roads, buildings, rivers, and textural regions.

3. Selecting a *similarity metric*, in order to decide if a transformed input image closely matches the reference image. This metric should make use of the features which are extracted from the image, be they specific pixels or global structures.

4. Selecting a *search strategy*, which is used to match the images based on maximizing or minimizing the similarity metric.

Perhaps the most straightforward, yet inefficient, approach to image registration is manual registration. This involves a human examining the images to be registered and selecting pixel matches between the two images. That is, one selects a pixel or group of pixels in each image that are in correspondence, based on the features they represent and their location relative to other pixels. In remotely sensed images, these matching pixel pairs are called ground control points (GCPs). Once a suitable number of GCPs have been selected, a transformation can be found between the images by minimizing the least squares distance between GCPs. There are commercial software products, such as ENVI, that compute a registration based on user-selected GCPs. Unfortunately, manual registration has many drawbacks. It requires human resources and is time-consuming. Moreover, in the case of two images with vastly different information content, it may be very difficult to identify GCPs that correspond exactly to one another.

In distinction to manual registration is automatic registration: registration that requires no human selection of GCPs or features. Types of automatic registration can be broken into two classes: pixel-based and global feature-based. Pixel-based automatic registration techniques are very similar to manual registration, in that individual pixels or groups of pixels are selected to compute the image registration. In this case, however, the GCPs are selected based on an algorithm; the *scale invariant feature transform* (SIFT)-algorithm [78] is very popular in this re-

gard. While speed is no longer an issue with pixel-based automatic registration, the difficulties of registering multimodal images remains. In particular, if there are no obvious pixels in the input image to match with ones in the reference image, this scheme will suffer. We will examine a lidar-to-optical image experiment in Section 4.6, in which this problem is manifest.

A second class of automatic image registration algorithms is those that take into account the entire image, not just prioritized GCPs. The basic idea behind these algorithms is to apply a feature-extraction algorithm to the input and reference image, that ideally isolates multi-pixel features common to both images. These features are then matched with an optimization scheme. Wavelet features and wavelet-like features, such as Simoncelli pyramids [94], have proven effective for this type of image registration [108], [96], [97], [22]. However, wavelets often fail to find the most robust features in an image. Roads, rivers, and other edge-like features are not well-captured by wavelet algorithms. This is because wavelets are essentially *isotropic*, meaning they are the same in all directions. Consequently, textural rather than directional features are emphasized by wavelets [30]. This lack of directional sensitivity leads to a lack of robustness in the corresponding registration algorithms: the feature space is too homogenous for the optimization scheme to avoid local minimizers near the global minimizer. To counter this, we have developed a feature extraction algorithm that has a strong directional emphasis, yielding a more robust registration solution.

We summarize our image registration algorithm in terms of the four components described in [11]:

1. *Search Space:* All of our examples feature images at the same scale, so effectively, we chose the search space to be restricted to the space of compositions of rotations and translations (RT). In future work, we shall examine incorporating polynomial transformations in our algorithm to address more complex, spatially varying distortions within images. Such images could include remotely sensed images which have not been ortho-rectified to remove distortions due to topography.

2. *Features*: Wavelet features in one case and shearlet features coupled with wavelet features in another.

3. *Similarity Metric*: Unconstrained least squares. That is, if F_R and F_I are the reference and input features, N the number of relevant pixels, (x_i, y_i) the integer coordinate of the ith pixel, and T_p the transformation associated to parameters p, we seek to minimize the similarity metric given by:

$$\chi^2(p) := \frac{1}{N} \sum_{i=1}^{N} \left(F_R(T_p(x_i, y_i)) - F_I(x_i, y_i) \right)^2 .$$

4. *Search Strategy:* Modified Marquadt-Levenberg method of solving non-linear least squares problems [84], [96].

4.3 Harmonic Analysis for Image Registration

Numerical wavelet algorithms decompose an image according to scale and translation. This has three principle benefits for image registration. First, it extracts features that are easier to match than the initial images, making search algorithms

more accurate and robust. Second, wavelet and wavelet-like algorithms often contain an iterative decimation step, which reduces the number of pixels in the images to be matched. This allows for faster computation time, when compared to methods without decimation [108]. Finally, wavelet algorithms represent textures very well, which can be very useful features for accurate image registration.

While wavelets and wavelet-like algorithms have had success in automatic image registration [108], they are lacking in certain regards. Indeed, the features they produce are often textural in nature [17]. Instead of producing sparse, distinct edges, wavelets produce soft, somewhat noisy looking features. As such, image registration techniques based on wavelets often suffer from a lack of robustness to initial transformation, since the lack of sharp features means there will be many local minimizers to the optimization algorithm. This is related to the isotropic nature of wavelets. These problems have been well-documented [104], and led to the burgeoning sub-discipline of harmonic analysis known as geometric multi-resolution analysis.

The shearlet construction adds a new parameter of decomposition to classical wavelet methods: direction. This means that shearlet coefficients $\{\langle f, \psi_{i,j,k}\rangle\}_{(i,j,k)}$ will contain information about a signal's behavior in different directions, which is not present in wavelet coefficients. This property makes shearlet coefficients optimally sparse for a broad class of signals, as has been discussed in Chapter 2. This anisotropy makes shearlets well-suited for problems in image processing, including image fusion [87], image denoising [38], and image in painting inpainting [66], [24].

4.3.1 Numerical Implementations of Wavelets and Shearlets

Wavelets have been numerically implemented in a variety of ways, and are pervasive in image processing. The wavelet features used in our algorithm are computed in C, but could just as easily have been computed in MATLAB or another high level programming language. The major difference between computing the features in C, as opposed to MATLAB, is speed; C is much faster. The shearlet features for our algorithm are computed in MATLAB. The algorithm that produces these features for a given image makes use of a recent MATLAB library [54], modified for computational purposes pertaining to optimization search strategies. For a given image, the toolbox in [54] is used to compute shearlet features in a variety of directions and scales. These are then thresholded and combined at each scale, to provide distinct features to be used in our optimization algorithm. Details of the precise construction are in Section 4.4.

As an demonstration of how wavelet and shearlet features differ, consider Figure 4.1, which features a 256×256 optical image of Washington state; the image contains many features which could be used for matching by an automatic image registration algorithm, such as textures from vegetation and edges from land cover boundaries.

To illustrate the directional character of discrete shearlet algorithms, and their utility for image registration, we show in Figure 4.2 the images produced by a MATLAB discrete wavelet algorithm using the 'db2' wavelet, and the shearlet feature algorithm we have developed.

Figure 4.1: *A grayscale optical image of Washington state containing both textural and edge-like features. The image is courtesy of Dr. David J. Harding of NASA Goddard Space Flight Center.*

The features produced by the isotropic wavelet transform are composed of diffuse speckle and the almost total absence of edge-like features. By contrast, the features produced by the shearlet transform highlight the distinct, linear features oriented in all directions. In the following sections, we demonstrate the improved registration performance of our shearlet approach using a variety of image sources.

4.4 Algorithms to be Tested

We conducted experiments with six algorithms. As a control, we consider the three algorithms used in [108], namely image registration based on feature-matching with spline wavelet pyramids, Simoncelli low-pass pyramids, and Simoncelli band-

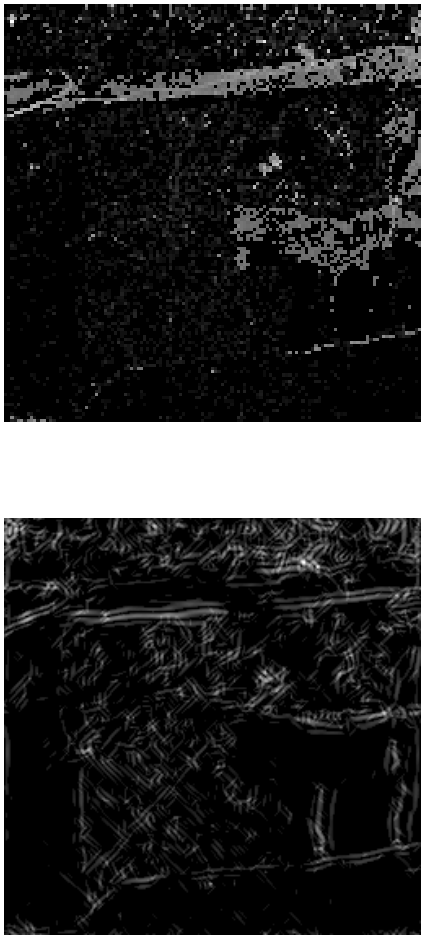


Figure 4.2: *Wavelet (top) and shearlet (bottom) features extracted from Figure 4.1.*

pass pyramids. We note that the Simoncelli features are derived from rotation-invariant and translation-invariant filters [94], and all three of these wavelet-like methods have the advantage of being translation-invariant. This means if an image I has wavelet features W_I, then the image $T_{x,y}(I)$ has wavelet features $T_{x,y}(W_I)$, where $T_{x,y}$ is a translation by the coordinates (x, y). Intuitively, if an image is shifted, then translation-invariant wavelet features will shift in exactly the same manner. Translation invariance is a very useful property for getting precise registration, but

does not contribute to robustness. The construction of these wavelet-like pyramids is beyond the scope of this article; we refer to [94], [100], and [108] for details. These three classical algorithms are collectively referred to as wavelet registration techniques. We compare these algorithms with a two-stage registration algorithm. First register with shearlet features to acquire a registration transformation. Then, set this registration as the initial guess and run the optimization algorithm again with each of the three wavelet techniques. These three two-stage techniques shall be collectively referred to as hybrid shearlet+wavelet registration techniques.

The motivation for this two-step registration algorithm is that using shearlet features for optimization should provide a highly robust, but less accurate registration. The shearlet algorithm is not translation invariant, resulting in small errors in registration computation. Thus, sub-pixel accuracy is not assured. Our two-step algorithm corrects for this by first providing a first stage registration, which is subsequently refined by the classical wavelet registration algorithms. This combines the strong robustness of shearlet features with the high precision provided by wavelet features. Note that both the wavelet and shearlet transforms are multi-scale, meaning each image is decomposed into sub-images with features of progressively finer scales. The key difference between the wavelet and shearlet algorithms is that a directional component is included in the latter but not the former. We summarize our hybrid algorithm below. θ denotes a counterclockwise rotation, T_x, T_y translations in the x and y directions respectively.

1. Input a reference image, I^r, and an input image I^i. These will be the images

to be registered.

2. Input an initial registration guess $(\theta_0, T_{x_0}, T_{y_0})$.

3. Apply shearlet feature algorithm and wavelet-like feature algorithm to I^r and I^i. This produces a set of shearlet features for both, denoted $S_1^r, ..., S_n^r$ and $S_1^i, ..., S_n^i$, respectively, as well as a set of wavelet features for both, denoted $W_1^r, ..., W_n^r$ and $W_1^i, ..., W_n^i$. Here n refers to the level of decomposition chosen. In general, n is bounded by the resolution of the images as

$$n \leq \left\lfloor \frac{1}{2} \log_2(\max\{M, N\}) \right\rfloor,$$

where I_r, I_i are $M \times N$ pixels. All our experiments are for 256×256 images, and we set $n = 4$.

4. Match S_1^r with S_1^i with a least-squares optimization algorithm and initial guess $(\theta_0, T_{x_0}, T_{y_0})$ to get a transformation T_1^S. Using T_1^S as an initial guess, match S_2^r with S_2^i, to acquire a transformation T_2^S. Iterate this process by matching S_j^r with S_j^i using T_{j-1}^S as an initial guess, for $j = 2, ..., n$. At the end of this iterative matching, we acquire our *final shearlet-based registration*, call it $T^S = (\theta^S, T_x^S, T_y^S)$.

5. Using T^S as our initial guess, match W_1^r with W_1^i with a least-squares optimization algorithm to acquire a transformation T_1^W. Using T_1^W as an initial guess, match W_2^r with W_2^i, to acquire a transformation T_2^W. Iterate this process by matching W_j^r with W_j^i using T_{j-1}^W as an initial guess, for $j = 2, ..., n$.

At the end of this iterative matching, we acquire our *final hybrid registration*, call it T^H.

6. Output $T^H = (\theta^H, T_x^H, T_y^H)$.

The wavelets-only algorithm is the same as above, without step 4.) and using $(\theta_0, T_{x_0}, T_{y_0})$ as the initial guess in step 5.).

As mentioned, the wavelet component of the algorithm is based on software coded in C, which is detailed in [108]. To produce the shearlet features for an image I, we proceed as follows:

1. Apply MATLAB script known as the fast finite shearlet transform (FFST) [54], which is part of the FFST library.

2. Perform hard thresholding on each shearlet coefficient to set the bottom 90% of coefficients in magnitude to 0. An example of these thresholded coefficients is in Figure 4.3.

3. Combine all the coefficients of a particular scale. Output the resulting n shearlet features, $S_1, ..., S_n$, where n is as above.

In addition to wavelet and shearlet features, we tested image registration with edge features produced by the classical edge detection algorithms of Canny [32] and Sobel [102]. These algorithms use discrete gradient methods to identify regions of sharp transition, which are then labeled as edges. Registration experiments with these algorithms were performed, with extremely negative results. Indeed, unless the truth registration is within a pixel or two of the initial guess for each parameter,

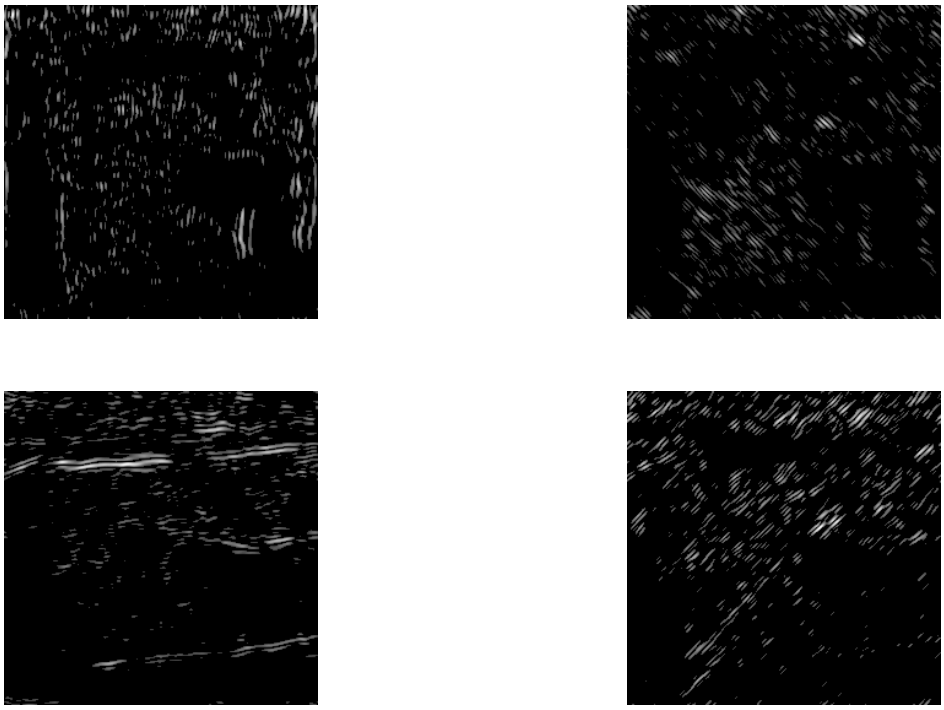


Figure 4.3: *Intermediate results of MATLAB shearlet features algorithm applied to Figure 4.1. Notice each image emphasizes a different direction. These can be combined to produce the shearlet image seen in Figure 4.2.*

these features could not be used to successfully register the images; the optimization algorithm immediately became trapped in a local minimum. Because of their very poor performance, we do not include the numerical results for these experiments.

The failure of classical edge detectors, which often produce many very thin features, in automatic image registration illustrates an interesting principle: that features ought to be substantial and relatively sparse in order to produce a robust registration algorithm. Indeed, although our shearlet algorithm also emphasizes edge-like features, only the most extreme global edges are produced. Our results corroborate this, as our experiments with shearlets indicate their effectiveness in pro-

ducing a highly robust registration algorithm, in stark contrast to the experiments with gradient-based edge detectors.

All of the registration algorithms tested were evaluated by computing the root-mean-square error (RMSE) between the truth registration parameters and those computed by the algorithms. In the case of our synthetic experiments, the truth registration parameters are the parameters of the misregistration we applied to the images. In the case of our lidar-to-optical experiments, the truth registration was computed using manual registration assisted by the software package ENVI. In the case of the multispectral-to-panchromatic experiments, the truth registration was a known, artificial misregistration. The details of the computation of RMSE are given in [108]. In each of our experiments, a different number of experiments are performed. The number of experiments was chosen to emphasize the robustness limits of the wavelets-only and hybrid shearlets+wavelets algorithms for each set of images.

We note that although we tested our algorithm only on images of the same resolution, the optimization scheme is capable of registering images of different resolutions. This is possible because our harmonic analysis algorithms perform resolution-decreasing decimations, which eventually bring the image of high resolution close to the resolution of the other image.

4.5 Experiments on Synthetically Generated Datasets

Our first set of experiments involved synthetic test images. We fixed a source image, then extracted a fixed reference image from this. We then rotated and translated this reference image within the larger source image, in order to acquire a collection of input images to register against the reference image. A diagram of this process is shown in Figure 4.4. The input images are these deliberately misregistered versions of the reference image.

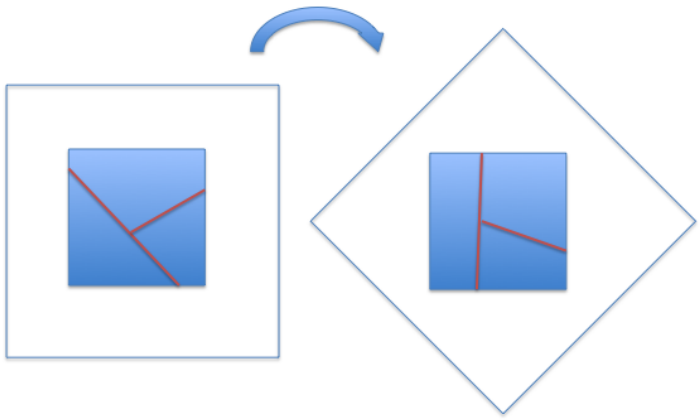


Figure 4.4: *In order to produce synthetic input images, we rotated and translated our reference image within the larger source image and extracted the resulting image. This extracted input image (right) is registered against the extracted reference image from the original image (left).*

This process has the benefit of simple error computation. Because we knew the

exact transformation that was applied to the reference image to acquire the input image, we could compare our registration result directly to these values. Moreover, since we are essentially comparing an image to a transformed version of itself, a convergent registration algorithm should be expected to produce sub-pixel accuracy. Thus, we expected to see very small RMSE values for the synthetic experiments. Sub-pixel accuracy should be attained to signify strong performance in this case, which is indicated by a RMSE value less than 1.

For ease of comprehension, we coupled the translation and rotation parameters together. We examined images that had been rotated and translated in the x and y direction by the same value. This joint parameter was denoted RT. For example, if RT= 5.5, then the image was rotated counter-clockwise by 5.5 degrees and translated in the x and y directions by 5.5 pixels. Nearest neighbor interpolation was used for rotations and non-integer translations. Letting RT increase from 0, we considered input images that were increasingly misregistered from the reference image. Consequently, the larger RT was, the more difficult it was for an optimization algorithm to derive the correct registration parameters. We were interested in allowing for larger values of RT, while maintaining good registration accuracy; this is the problem of robustness.

We performed synthetic experiments using as the source image a 1024×1024 image extracted from Band 4 of a Landsat Thematic Mapper scene of the Mount Hood National Forest; see Figure 4.5. We then extracted the center 256×256 sub-image to serve as the reference. To produce the input images, we rotated and translated the reference image within the source by a range of RT values.

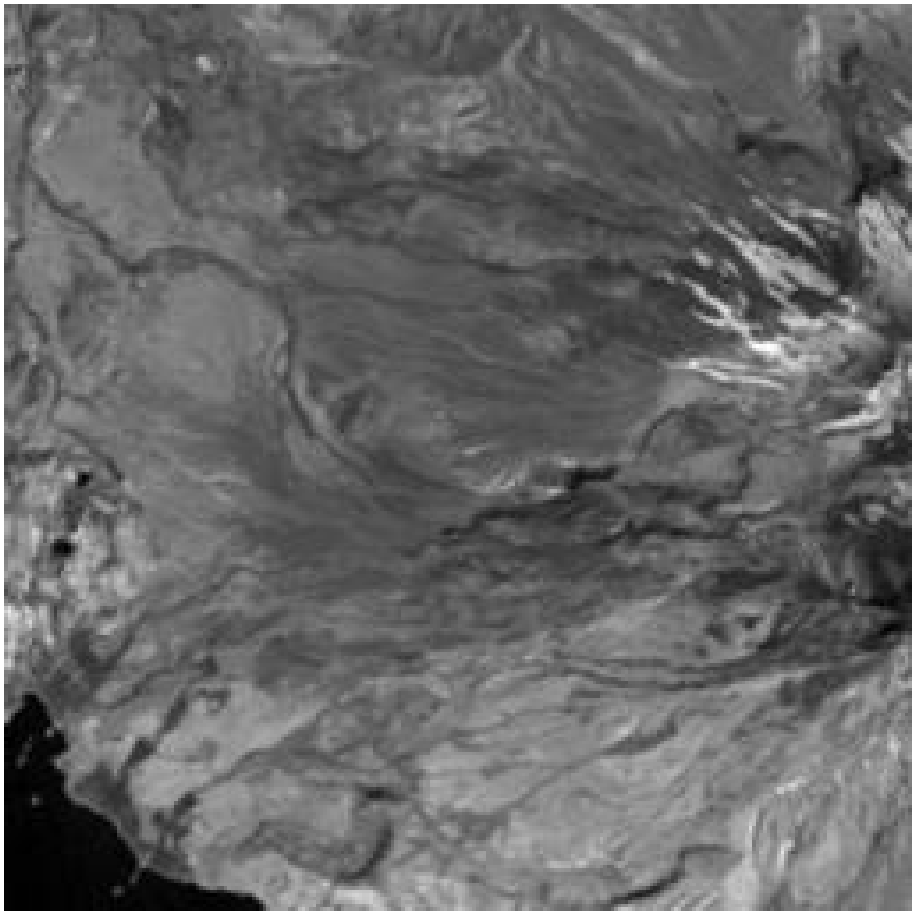


Figure 4.5: *Landsat-TM scene of Pacific Northwest used as source image for synthetic experiments. The image has been converted to grayscale. The image is courtesy of Dr. Jacqueline Le Moigne of NASA Goddard Space Flight Center.*

The RT parameter ranged from 0 to 40, with increments of 0.2. We performed 200 corresponding registration experiments, in which we used each of our six registration algorithms to find a registration transformation. Such a registration transformation is parametrized as a triple of rotation and translation values (θ, T_x, T_y). To analyze the quality of our algorithms, we computed the RMSE between the truth registration parameters and the computed registration parameters. Table 4.1 displays the number of converged experiments, the percentage of con-

Registration Technique	Number of Converged Experiments (out of 200)	Percentage of Converged Experiments	RMSE	Relative Improvement
Spline Wavelets	108	54.00%	.0019	-
Simoncelli Band-Pass	21	10.50%	.0045	-
Simoncelli Low-Pass	113	55.50%	.0040	-
Shearlet+ Spline Wavelets	154	77.00%	.0058	42.59%
Shearlet+ Simoncelli Band-Pass	154	77.00%	.0080	633.33%
Shearlet + Simoncelli Low-Pass	154	77.00%	.0081	36.28%

Table 4.1: *Comparison of registration algorithms for Landsat-TM synthetic experiment.*

verged experiments, and the average RMSE in converged cases. In the case of the hybrid shearlet+wavelet algorithms, the relative improvement in the number of converged experiments as compared to the corresponding wavelets-only technique was computed using the formula:

$$\text{Relative Improvement} := \frac{CV_{s+w} - CV_w}{CV_w},$$

where CV_{s+w} denotes the number of converged experiments with shearlet and wavelet features, and CV_w denotes the number of converged experiments with wavelets alone.

The RMSE between each wavelets-only technique and the corresponding hybrid technique for each RT iteration are displayed graphically in Figure 4.6. An experiment is considered convergent if its RMSE is less than 1, indicating sub-pixel registration was achieved. This is a very high degree of precision for synthetic experiments.

Our results indicate, based on the number of wavelets-only converged experiments, the Simoncelli low-pass features offer the most robust image registration of

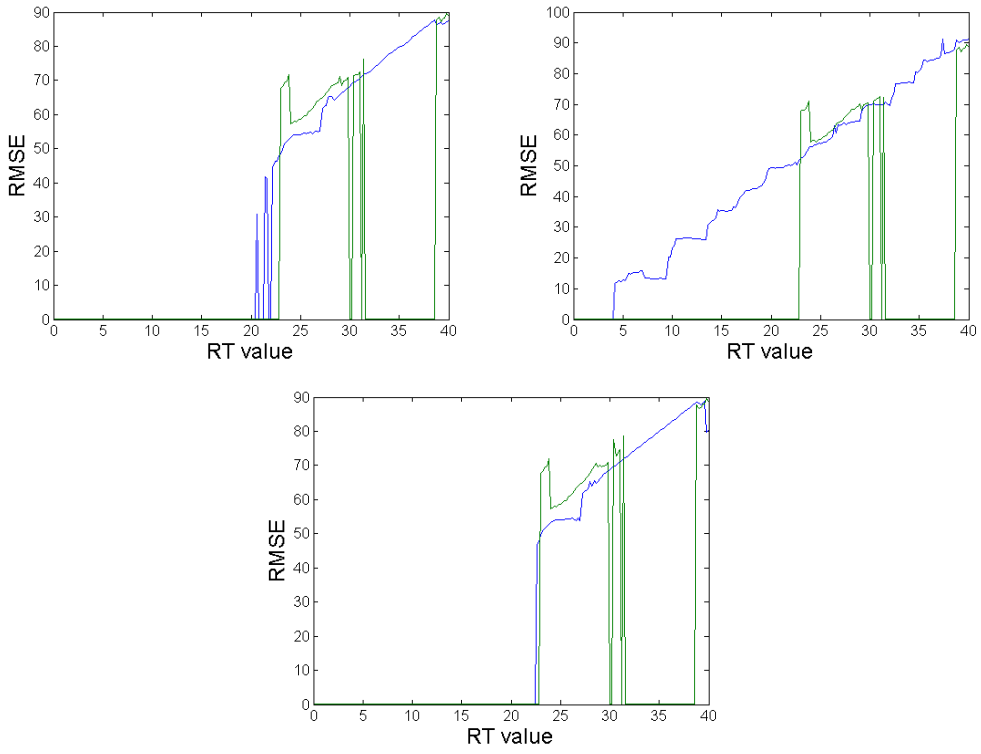


Figure 4.6: *Comparison of wavelet and shearlet+wavelet algorithms for Landsat-TM image synthetic experiments (top left: splines; top right: Simoncelli band-pass, bottom: Simoncelli low-pass); blue is wavelets, green is hybrid shearlets+wavelets.*

the wavelet methods, while Simoncelli band-pass features perform far worse than the others. All three offer strong accuracy when convergent. The hybrid shearlet+wavelet algorithms, however, offer substantially increased robustness at a very small error increase over the convergence set. Indeed, sub-pixel accuracy is achieved in all convergent cases. It is interesting to note that all three of the shearlet algorithms have the same number of converged experiments. This can be justified by recalling that the first stage of all three of the algorithms involves acquiring a first stage registration, based on shearlets. Thus, when one of the hybrid algorithms fails to converge, it is likely because the shearlet feature matching step fails to converge.

This failure would be the same for all algorithms which begin with a shearlet feature matching, since it is highly unlikely that the second-stage wavelet feature matching would compensate sufficiently. We conclude that for these synthetic experiments, the hybrid shearlet-wavelet registration algorithms provide increased robustness over the classical wavelets algorithms.

4.6 Experiments on Multimodal Images

We next considered experiments registering two real images with different modalities. This represented a more realistic test of the functionality of our algorithms, since in reality, image registration will be between two different images, not an image and a synthetic misregistration of itself. Moreover, many image registration problems in the geosciences involve registering images of different modalities, so our present experiments are relevant to an important class of registration problems. In particular, the modal differences can make finding GCPs exceedingly difficult. This renders pixel-based automatic registration algorithms, such as SIFT, sub-optimal.

As an illustration of this, consider two data sources for a WA state mixed land-cover scene: one lidar and the other optical. These images are shown in Figure 4.7. The lidar data was acquired in 2003 by Terrapoint, Inc., under contract to NASA, using a multi-return airborne laser swath mapping (ALSM) instrument. The optical data is a natural color aerial photograph, presented as a grey-scale image, collected in 2006 and obtained by Google from the United States Geological Survey. These

data sources have fundamentally different, but related, information content. A lidar image, commonly referred to as a digital elevation model (DEM), is a measure of the elevation of the components making up the surface. The data we used is a highest surface DEM. This represents vegetation canopy tops where vegetated, and ground, roads, and building tops where not vegetated. On the other hand, an optical image records solar radiance reflected from the surface. The latter is a function of the reflectance of the surface components and their three-dimensional organization. Together these define the patterns and brightness of illuminated and shadowed patches seen in optical images. The features in the two image types are markedly different and thus are not well suited for GCP identification.

We partially overcome this problem by generating a synthetic shaded-relief image by artificially illuminating the lidar elevation image with a light source directed in the same orientation as the solar illumination in the optical image as shown in Figure 4.7. Similar patterns of illuminated and shadowed patches are produced but the correspondence is not exact for the following reasons:

1. The lidar elevation image is not a perfect representation of the surface.

2. The shaded-relief modeling used is not a perfect representation of solar illumination.

3. There can be surface change between the times of lidar and optical image acquisition. The discrepancies between the images can cause erroneous selection of GCPs.

As an experiment, a MATLAB SIFT algorithm [101] was applied to these images,

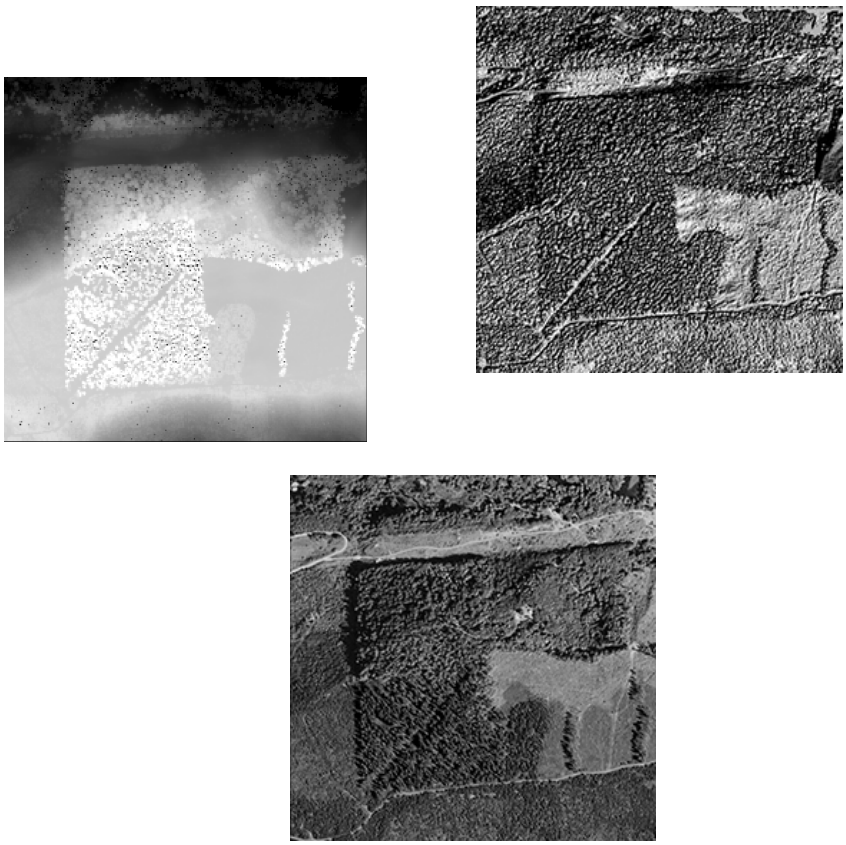


Figure 4.7: *Lidar ALSM elevation image (top left), the derived shaded relief image (top right) and aerial photograph for a scene in WA state (bottom). The shaded relief image, illuminated in the same direction as in the optical image, depicts some similar patterns of textures and edges. The images are courtesy of Dr. David J. Harding of NASA Goddard Space Flight Center.*

which computes pairs of points to use for pixel-based registration. This code is courtesy of Andrea Vedaldi, and is available at `http://www.robots.ox.ac.uk/~vedaldi/code/sift.html`. The results of this algorithm appear in Figure 4.8, where the corresponding pixels are linked with a green line. There are far more

incorrect pixel pairs than correct, making this method ineffective for image registration. This highlights a pitfall of automated SIFT for the registration of multimodal images: the visual similarity of features must be very high.

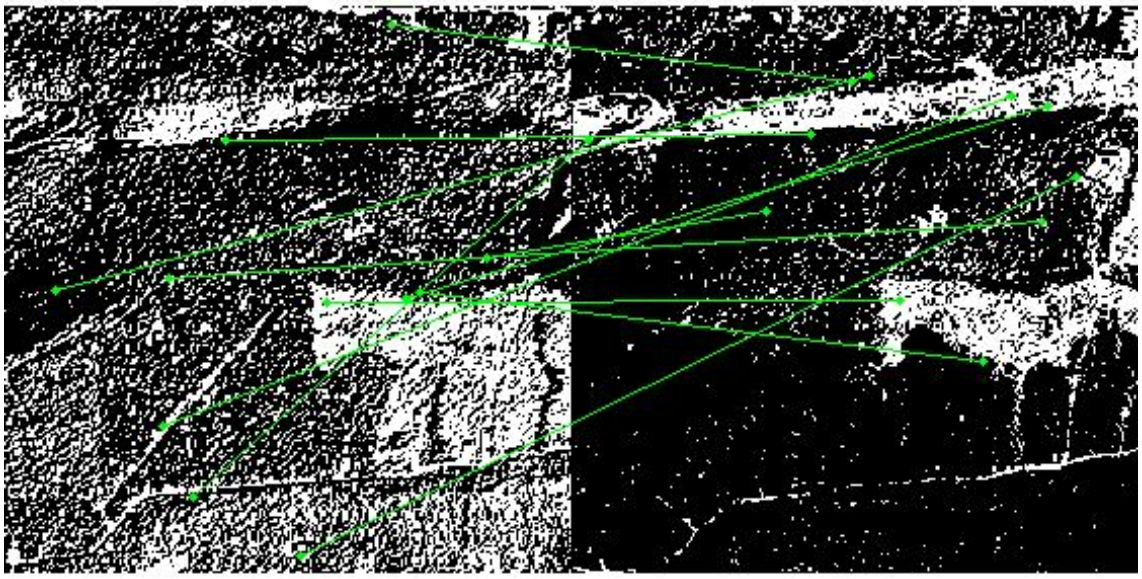


Figure 4.8: *The pixels computed by SIFT in the lidar shaded-relief (left) and optical (right) images of WA, connected by green lines. Note the lack of correspondence; such points are unusable for a registration algorithm.*

It is important to note that when conducting experiments to assess the accuracy of registration algorithms using multimodal images, knowing the truth registration between the images is not as straightforward as in synthetic image registration. In our synthetic experiments, we knew the truth registration of the input images

with respect to the reference image perfectly, since we designed the misregistration that produced the input images. Thus, we easily computed the RMSE between the truth registration and computed registration. In our multimodal image experiments, we did not know a priori the truth registration. We established this via manual selection of between 15 and 25 GCPs using the ENVI image processing software and applying its rotation-translation transformation solution.

However, as mentioned, this is very difficult for certain image pairs, as there is often little local pixel-to-pixel correspondence near key features. Indeed, consider the images in Figure 4.9. These images are subsets of the WA state lidar and optical images, depicting the same alignment of trees. However, there is little pixel-to-pixel correspondence. These images demonstrate that features in multimodal images can have global correspondence, but not pixel-to-pixel correspondence. Thus, establishing the truth registration using our manual GCP method was difficult in these circumstances. As such, the margin for the lidar-to-optical experiments should be increased from that used in the synthetic experiments, to account for approximations made in computing the truth registration. Near sub-pixel accuracy is the goal, corresponding to a RMSE of less than 2.

For these experiments, we tested for algorithm robustness by allowing the initial registration guess for the optimization algorithm to vary according to the RT parameter. The truth registration was modified by RT to produce the initial guess for the algorithm. So, if RT= 5.5, then the initial guess of the registration algorithm is a counterclockwise rotation of 5.5 degrees and a translation in both the x and y directions by 5.5 pixels, on top of the truth registration. This tests robustness

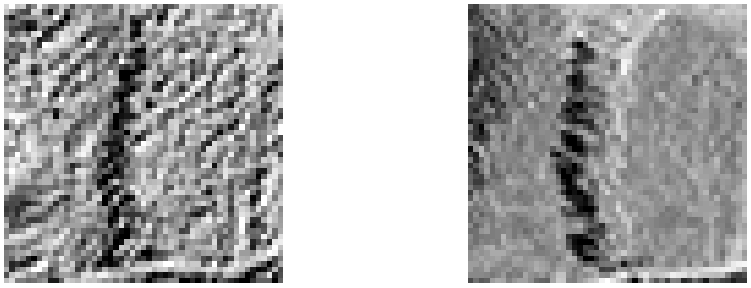


Figure 4.9: *The same alignment of trees in the lidar shaded-relief (left) and optical (right) images of WA. Although there is clear correspondence at the macroscopic level, it is difficult to find pixel-to-pixel correspondences.*

in a slightly different way than did the synthetic experiments. For the synthetic experiments, the initial guess was always set to 0, but the images themselves were synthetically misregistered to be increasingly far apart. Here, the images are at a fixed misregistration, and the initial guess is changed. We use this test for robustness because the images are not synthetic misregistrations of one another, but two distinct images that may or be not be co-registered. The RT applied to the truth registration shall be denoted RT_{IG} to indicate this RT parameter was not applied to create a synthetic input image, but was applied to the initial guess of the algorithm.

4.6.1 Lidar to Optical Registration Experiments

Our first set of multimodal experiments involved registering our WA state lidar shaded-relief image and optical image of approximately the same scene. The truth registration for these images was computed using our manual ENVI method to be $(\theta, T_x, T_y) = (2, 1, -3)$.

Registration Technique	Number of Converged Experiments (out of 101)	Percentage of Converged Experiments	RMSE	Relative Improvement
Spline Wavelets	55	54.46%	1.3439	-
Simoncelli Band-Pass	61	60.40%	1.5862	-
Simoncelli Low-Pass	86	85.15%	1.4868	-
Shearlet + Spline Wavelets	60	59.41%	1.3144	9.09%
Shearlet + Simoncelli Band-Pass	65	64.36%	1.5836	6.56%
Shearlet + Simoncelli Low-Pass	88	87.13%	1.4861	2.33%

Table 4.2: *Comparison of registration algorithms for lidar-to-optical experiment.*

In this experiment, we allowed the initial RT_{IG} parameter to vary from -25 to 25, with increments of .5. These RT_{IG} parameters are applied to the truth registration of $(2, 1, -3)$ to produce the initial guess. We then performed 101 image registration experiments with each of the six algorithms, then computed the RMSE between the algorithm solutions and the truth registration along with the number of converged experiments. The results appear in Table 4.2 and Figure 4.10.

Our results in this case show Simoncelli low-pass features to provide the best robustness out of the three wavelet algorithms, spline wavelets the least so. The hybrid shearlet+wavelet algorithms offer some improvement over the classical wavelets, but not to the same degree as with the synthetic image experiments. The increased RMSE in all experiments, when compared to the synthetic experiments, is attributed to both the visual differences between these images, and the fact that the truth registration was approximated using the imperfect method of manual registration via GCP selection.

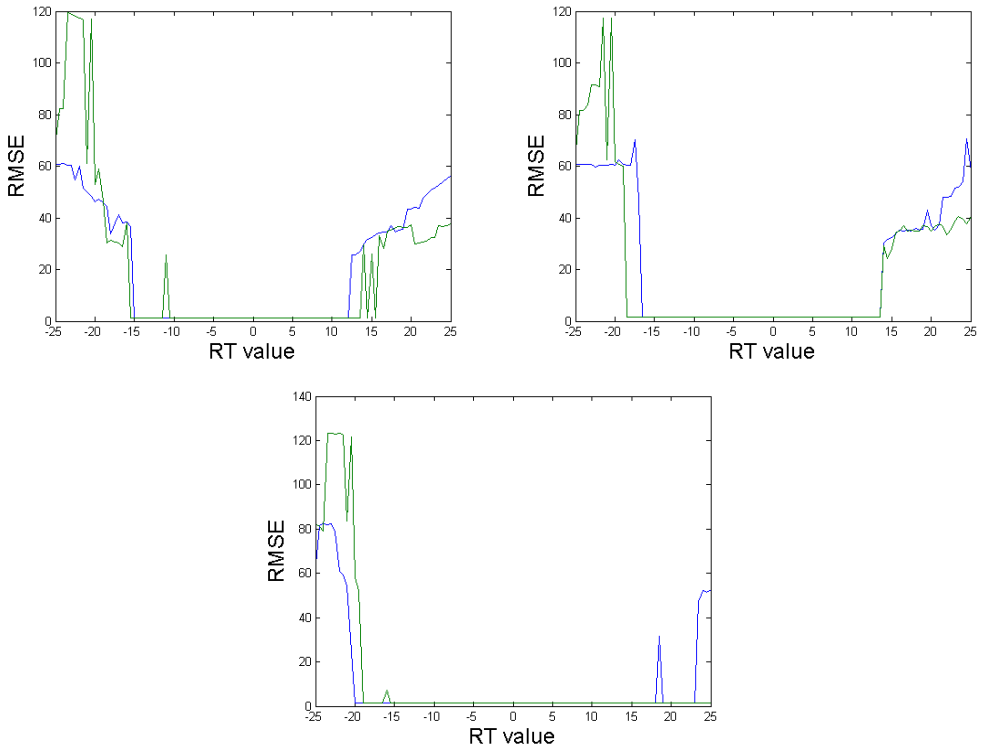


Figure 4.10: *Comparison of wavelet and shearlet+wavelet algorithms for WA lidar-to-optical experiment (top left: splines; top right: Simoncelli band-pass; bottom: Simoncelli low-pass); blue is wavelets, green is hybrid shearlets+wavelets.*

4.6.2 Multispectral to Panchromatic Registration Experiments

Our second set of multimodal experiments involved registering two bands of a multispectral image. These images of Hasselt, Belgium were acquired by the Landsat 7 Enhanced Thematic Mapping Plus (ETM+) sensor in 1999 and distributed as part of the IEEE Geoscience and Remote Sensing Society 2000 data fusion contest. The first seven bands of the sensor are multispectral, and produce images covering the visible and infrared spectra; the eighth band is panchromatic. We considered the registration of band 1 to band 8. Band 1 has a narrow spectral resolution of 450-515 nm, while band 8 has a broad spectral resolution of 520-900 nm. These images appear in Figure 4.11.

These images were artificially misregistered to have truth registration of $(\theta, T_x, T_y) = (5, 10, 10)$. We let the initial RT_{IG} parameter vary from -50 to 50, incrementing by 1. We then performed 101 corresponding image registration experiments with each of the six algorithms, then computed the RMSE and number of converged experiments. Results from this experiment appear in Table 4.3 and Figure 4.12.

Our results indicate that among the three wavelet algorithms, Simoncelli band-pass features are the most robust, spline wavelets the least so. In all three cases, the shearlet registration algorithm outperforms the standard wavelet registration algorithm. The average RMSE is consistent among all the algorithms, and is higher than in the synthetic experiments of Section 4.5, but lower than for the lidar-to-optical experiments. This is attributed to the fact that features in the multispectral and panchromatic images are more similar than the features in the lidar-to-optical ex-

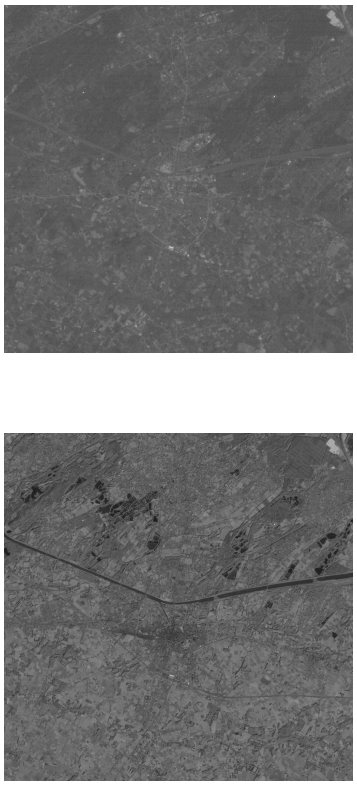


Figure 4.11: *Multispectral band 1 (top) and panchromatic band 8 (bottom) images of Hasselt, Belguim acquired by Landsat ETM+. The images have been converted to grayscale. A subset is extracted from these images to ease computation. The images are courtesy of the IEEE 2000 GRSS Data Fusion Contest.*

periments, but are less similar than in the synthetic experiments. Indeed, the RMSE for the multispectral-to-panchromatic experiments falls in between the RMSE for the synthetic experiments and the RMSE for the lidar-to-optical experiments.

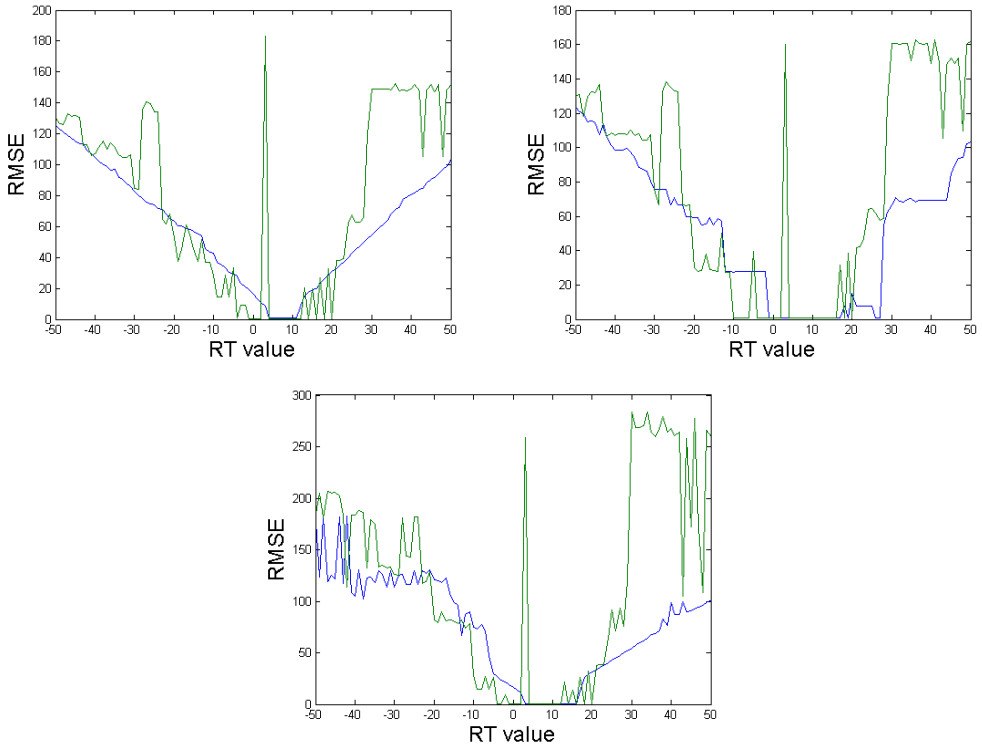


Figure 4.12: *Comparison of wavelet and shearlet+wavelet algorithms for panchromatic to multispectral experiment (top left: splines; top right: Simoncelli band-pass; bottom: Simoncelli low-pass); blue is wavelets, green is hybrid shearlets+wavelets.*

Registration Technique	Number of Converged Experiments (out of 101)	Percentage of Converged Experiments	RMSE	Relative Improvement
Spline Wavelets	8	7.92%	.6376	-
Simoncelli Band-Pass	19	18.81%	.7534	-
Simoncelli Low-Pass	14	13.86%	.6034	-
Shearlet + Spline Wavelets	20	19.80%	.5185	150.00%
Shearlet + Simoncelli Band-Pass	27	26.73%	.6494	42.11%
Shearlet + Simoncelli Low-Pass	20	19.80%	.5513	42.86%

Table 4.3: *Comparison of registration algorithms for panchromatic to multispectral experiment.*

4.7 Summary and Conclusions

We have demonstrated in synthetic and real experiments, with both unimodal and multimodal images, that shearlet features can be used to increase robustness of image registration algorithms. In our synthetic experiments, the hybrid algorithms produced an average increase of 237.40% in number of converged experiments. When comparing the best wavelet algorithm to its corresponding hybrid algorithm, the increase was 36.28%. In our lidar-to-optical registration experiment, the hybrid algorithms produced an average increase of 5.99% in number of converged experiments. When comparing the best wavelet algorithm to its corresponding hybrid algorithm, the increase was 2.33%. In the multispectral-to-panchromatic registration experiments, the hybrid algorithms produced an average increase of 78.32%. When comparing the best wavelet algorithm to its corresponding hybrid algorithm, the increase was 42.11%.

Although registration robustness was improved in all three image experiment sets by the use of the hybrid algorithm, the extent of improvement varied. In particular, the improvement was noticeably lower in the lidar-to-optical experiments than in the synthetic experiments or the multispectral-to-panchromatic experiments. In both the Landsat image used for the synthetic experiments and the multispectral bands, edge features were dominant. These edges are optimally represented theoretically by shearlets, and the improvement in robustness for these experiments manifests this. Conversely, the lidar shaded-relief and optical images have fewer common edges and many more isotropic textures. These textures are not strongly directional, and are not theoretically optimized by shearlets. That the lidar and optical images display fewer shared edge features is clear from examining the shearlet features produced by our algorithm; see Figure 4.13.

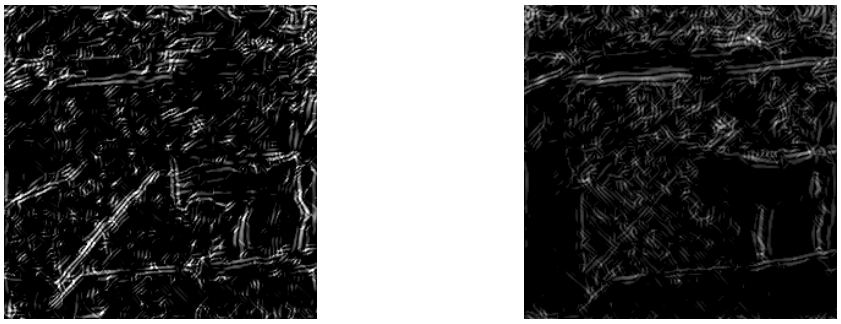


Figure 4.13: *Shearlet features produced for the lidar (left) and optical (right) images of WA, respectively. Notice that the strongest edge-like features in each image are emphasized: a diagonal of trees for lidar and and land-cover change near the top for optical. Due to the information content differences between lidar and optical data, these features are not represented in both images. Some edges are shared by both, but not all.*

Consequently, it is reasonable that shearlets would offer comparatively little improvement over wavelets in terms of registration robustness for these type of images, as they lack the strong, shared directional features shearlets are known to optimize.

We conclude that the experiments performed are practical confirmation of the theoretical properties of shearlets. While they are an effective tool for registering images with strong edge features, they are somewhat less effective for registering images that are texturally dominant. Our hybrid algorithms produce a solution to this difference, by incorporating shearlets to capture edge features, and also wavelets, to capture textural features.

In light of the success of the current experiments, it would be of interest to test other directionally-sensitive representation systems in place of shearlets, such as curvelets or contourlets. These systems produce sparse features that represent edge optimally, in a manner theoretically similar to shearlets. How they would perform for image registration, compared to shearlets, is not clear. They are numerically implemented by not one basis function ψ, but by a finite family of functions. This variety of generating functions is often considered a disadvantage, but could offer flexibility in registering a wide variety of images, with basis functions adapted for certain non-linear edge features found in images.

It is also of interest to apply these harmonic analytic techniques to the registration of more complicated data types, such a three-dimensional representations of data, extending beyond the two-dimensional images considered in this article. For example, lidar measures the three-dimensional distribution of vegetation com-

ponents from which the two-dimensional highest surface DEM used in this article was derived; see Figure 4.7. Registration in the z direction, as well as x, y, of multi-temporal $3D$ lidar cubes could aid in the identification of areas of vertical vegetation change due to processes such as growth, tree mortality, fire and human land use activities. This approach could also be applied to hyperspectral image cubes in which the z dimension is a record of spectral absorption features. $3D$ registration of multi-temporal data could aid in the identification of locations that have undergone spectral change. Three dimensional shearlet implementations exist [71], and could be applied to this problem.

In addition, developing more efficient techniques for registering extremely large images is of interest, since the methods described in this article become computationally burdensome for large images. An approach of partitioning the image into a few sub-images, each containing many features, then running the algorithm described in this article on these sub-images, shall be investigated in future work.

Chapter 5: Superresolution with Shearlets

This chapter develops a novel algorithm for the superresolution of images. A superresolution algorithm increases the resolution of an image, while attempting to preserve smoothness and important information in the image, and without introducing artifacts. There may be additional information incorporated into the new superresolved image, such as lower resolution sub-pixel shifts of the same scene, or images of different modalities. Many classical methods for superresolution employ an interpolation scheme, based on some form of weighted local averaging [64]. More sophisticated methods exploit the geometry inherent in the image to augment these interpolation schemes by improving smoothness [8].

In order to analyze the geometry of an image, anisotropic harmonic analysis techniques are useful. These methods provide directionally sensitive computational tools for decomposing an image, and efficiently and accurately encoding its most salient features. In particular, the construction of *shearlets* offers a computationally efficient method for analyzing the directional content of an image. This information can be used to provide smoother superresolved images, as our algorithm's results demonstrate; see Section 5.3.

The structure of this chapter is as follows. We begin with relevant background

on the problem of image superresolution in Section 5.1. Our shearlet-based superresolution algorithm is detailed in Section 5.2, and is tested on a remotely sensed image in Section 5.3 [27], [9]. We conclude and explore future work related to this algorithm in Section 5.4.

5.1 Background on Superresolution

The problem of superesolution is significant in image processing. The goal of superresolution is to increase the resolution of an image I, while preserving detail and without producing artifacts. The outcome of a superresolution algorithm is an image $\tilde{I}$, which is of the same scene as I, but at a higher resolution. We restrict ourselves here to greyscale images, hence we can consider our images as real-valued matrices. Let I be an $M \times N$ matrix and $\tilde{I}$ an $\tilde{M} \times \tilde{N}$ matrix, with $M < \tilde{M},\ N < \tilde{N}$. We consider the common case where $\tilde{M} = 2M$ and $\tilde{N} = 2N$, which corresponds to doubling the resolution of the original image in both the horizontal and vertical directions.

Superesolution can be done by using information in addition to I, such as low resolution images at sub-pixel shifts of the scene [90], or images of the scene with different modalities. The latter method is related to the specific problem of pan-sharpening [25]. Alternatively, superresolution can be performed using only I. The first type of superresolution requires additional data, and is thus less desirable than the second type. In this chapter, we shall develop a superresolution method of the second type, which requires as input only a single image.

There are several standard approaches to superresolving I without using additional information. Among the most common are *nearest neighbor interpolation* and *bicubic interpolation*. Let us consider the superresolved version of $I = \{a_{m,n}\}$, denoted $\tilde{I} = \{\tilde{a}_{i,j}\}$. Here, the values $\tilde{a}_{i,j}$ and $a_{m,n}$ may be understood as entries of a real matrix representing the images. We must compute each pixel value in the new image, namely $\tilde{a}_{i,j}$, from the pixel values of the original image, $a_{m,n}$.

In the case of nearest neighbor interpolation, new pixel values are computed simply by replicating current pixel values. This method is simple and computationally efficient, but leads to extremely jagged superresolved images. It is unsuitable when a high-quality, smooth $\tilde{I}$ is required. Other methods involve convolving the image with an interpolation kernel, which amounts to taking a weighted average of pixel values within some neighborhood. For example, bicubic interpolation determines $\tilde{I}$ by computing each $\tilde{a}_{i,j}$ as a weighted average of the 16 nearest neighbors in I; the weights are chosen to approximate the derivative values at the pixels being analyzed. A precise description of the algorithm may be found in [64].

A novel method for improving the smoothness of images superresolved using these methods was demonstrated in [8]. In this algorithm, local dominant directions are computed using tight frames derived from circulant matrices. After using nearest neighbors or bicubic upsampling, a motion blur filter is applied in the dominant direction. Areas with low variance are assumed to have no dominant direction. This method resulted in superresolved images with much smoother edge features when compared to the interpolation techniques alone. This method proved effective, but required new tight frames to be computed for each application of the algorithm,

depending on the image size and the structure of the features present, thus limiting its efficacy. Essentially, this method uses frame theory to compute locally dominant directions; we shall compute locally dominant directions in another, more efficient manner.

The aim of this chapter is to develop a superresolution algorithm that computes dominant directions efficiently and accurately, using the harmonic analytic construction of shearlets; see Chapter 2. This method is quite general, and can be applied to images of any size, and has few tunable parameters.

5.2 Description of Shearlet Superresolution Algorithm

The goal of our algorithm is to exploit the directional sensitivity of shearlets to efficiently find the location and orientation of edges in images, which are then superresolved smoothly. Our algorithm for shearlet-based superresolution is coded in MATLAB. It is described below. Consider an $M \times N$ image matrix I; the superresolved image shall be denoted $\tilde{I}$ as above.

1. Apply the Fast Finite Shearlet Transform [56], [55] to I. This produces shearlet coefficients up to $\left\lfloor \frac{1}{2}\log_2(\max\{M,N\}) \right\rfloor$ scales. If we label the scales from coarsest to finest scale starting at $j = 1$, we have 2^{j+1} matrices of size $M \times N$ at the jth scale, each corresponding to a different direction from 90° to $(90 + 180(1 - 1/2^{j+1}))^\circ$, equally spaced. Denote these directional matrices $D_1, ..., D_{2^{j+1}}$. For the experimental images, we used the $j = 3$ scale since it best captured the edges, giving us 16 directions. This parameter may be set

differently depending on the size of the image under analysis.

2. Upsample by a factor of 2 each of $D_1, ..., D_{2^{j+1}}$, using the upsampling method of bicubic interpolation to acquire $\tilde{D}_1, ..., \tilde{D}_{2^{j+1}}$. These contain the directional information that will be used later.

3. Upsample by a factor of 2 the original image I, using the upsampling method of bicubic interpolation to acquire $\tilde{I}$. This upsampled $\tilde{I}$ will be modified using the directional information present in $\tilde{D}_1, ..., \tilde{D}_{2^{j+1}}$.

4. Assign each pixel in $\tilde{I}$ a local direction based on which matrix contained the shearlet coefficient of largest magnitude, i.e. which entry in that location is maximal among $\tilde{D}_1, ..., \tilde{D}_{2^{j+1}}$. Pixels which have no dominant direction are determined by one of the following three methods:

   (a) Pixels whose maximum coefficients were in the bottom 10% of all max coefficients were assigned no direction.

   (b) Apply a standard deviation filter of size 5 to $\tilde{I}$ to acquire $\tilde{I}_\sigma$. If a pixel in $\tilde{I}_\sigma$ has value less than .05, this pixel is assigned no dominant direction. Intuitively, pixels with low $\tilde{I}_\sigma$ value are locally constant, and should not be assigned a dominant direction. The parameter .05 can be tuned as needed.

   (c) Apply the Canny edge detector with default parameters to $\tilde{I}$, then thicken the edges using the MATLAB function 'imdilate'. Apply this mask to each of $\tilde{D}_1, ..., \tilde{D}_{2^{j+1}}$ and proceed as in a.) This has the effect of pick-

ing no dominant direction if a pixel is far from an edge-like feature, as determined by the Canny edge detector.

5. Apply a motion blur filter of length ℓ in each of the 2^{j+1} directions $\tilde{I}$, to produce $\tilde{I}_1, ..., \tilde{I}_{2^{j+1}}$. The parameter ℓ is tunable, and was set to $\ell = 5$ for our experiments.

6. Replace the pixel values of $\tilde{I}$ by their corresponding blurred version based on the previously assigned local direction, i.e. with the pixel value in $\tilde{I}_m$ where the pixel has dominant direction corresponding to m.

7. Output the superresolved image $\tilde{I}$.

5.3 Experiments and Results

We test our algorithm against standard bicubic interpolation by running both superesolution algorithms on a remotely sensed image. Our test image is from an aerial view of the University of Mississippi-Gulf Park near Gulfport, acquired with a CASI-1500 sensor with spectral range $375 - 1050$ nm in 72 bands. The image is courtesy of Paul Gader, Alina Zare, Ryan Close, J. Aitken, and G. Tuell [46]. We choose to analyze the 30^{th} band, due to its relatively high contrast. The image has a spatial resolution of 1 m and consists of 325×337 pixels. We perform our algorithm on the full spacial image, though we extract a 125×125 subset for visualization purposes; this subset is shown in Figure 5.1.

The 250×250 image produced from a simple bicubic interpolation is shown

Figure 5.1: *Original* 125×125 *image of Gulfport, MS. This is* I *in the algorithm as described in algorithm description.*

in Figure 5.2.

Figure 5.2: *Superresolved* 250×250 *image with bicubic interpolation. Notice that the edge-like features are jagged.*

We consider three shearlet-based superresolved images, based on the three methods for determining the pixels with no dominant direction (see Step 4 in algorithm description). In the interest of space, we show only the results for methods a.) and c). The superresolved images using our algorithm are shown in Figures 5.3

and 5.4.

Figure 5.3: *Superresolved* 250×250 *image using our algorithm with a.) in Step 4.*

Figure 5.4: *Superresolved* 250×250 *image using our algorithm with c.) in Step 4.*

Figures 5.5 and 5.6 illustrate the local directions chosen by our algorithm with methods a.) and c.) for determining pixels with no dominant direction (see Step 4 in description of our algorithm). In these images, the colors vary from dark blue (corresponding to a dominant direction of 90°) to dark red (corresponding to a dominant direction of 258.75°). In the case of method a.), only a few pixels, indicated by the darkest blue, were not assigned a direction. Some of these pixels

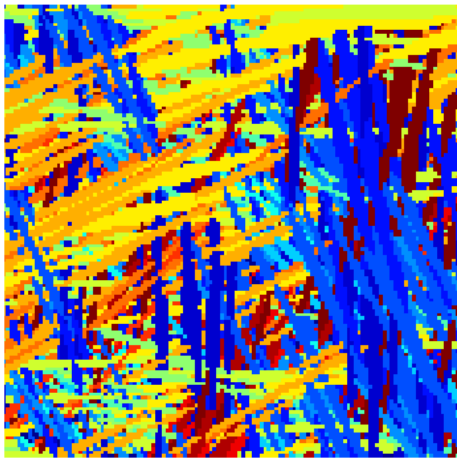


Figure 5.5: *Assignment of local directions, based on maximal shearlet coefficients and a.) in Step 4. Direction varies from dark blue* (90°) *to dark red* (258.75°). *The darkest blue corresponds to no assigned direction.*

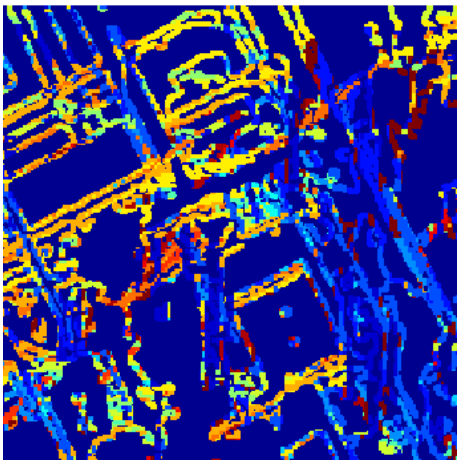


Figure 5.6: *Assignment of local directions, based on maximal shearlet coefficients and c.) in Step 4. In this case, a Canny edge detector is applied to determine which pixels have no dominant direction. Direction varies from dark blue* (90°) *to dark red* (258.75°). *The darkest blue corresponds to no assigned direction.*

can be seen in the lower left corner of the figure. In the case of method c.), far more pixels were not assigned a direction. Method c.) seems to be more accurate

in finding all edges, when compared to method b.), which is not pictured in this chapter.

5.4 Conclusions and Future Work

Our shearlet algorithm produces smoother superresolved images with fewer artifacts, when compared to bicubic interpolation. Notably, the method using c.) in Step 4 in the description of our algorithm, namely using a Canny edge detector to find areas with no dominant direction, provided very good results. In particular, the dominant direction map seen in Figure 5.6 is quite convincing in this case. We would like to find a quantitative measure that demonstrates the superiority of our algorithm over bicubic interpolation, which is clear visually. Towards this end, we cut out a 125×100 subset out of the top left of each superresolved image, an area consisting mainly of edges and flat regions. See Figure 5.7 for this region in the superresolved image using bicubic interpolation.

We then averaged the length of the gradient vector over all pixels. The idea is that jagged edges lead to longer edges and hence larger gradients. Smaller gradient vectors are then associated with smoother, more accurate edges. Under this metric, the bicubic upsampling did the worst with an average gradient length of 0.0244. Methods b.) and c.) performed better, averaging 0.0218 and 0.0213, respectively. Method a.) performed the best with an an average gradient of 0.199. Recall, however, that we restricted ourselves to an area with primarily edges and flat regions. For more complicated areas, method a.) tends to blur excessively, so we conclude

that method c.) is the best overall. The images used for this quantitative analysis appear in Figures 5.7, 5.8, and 5.9.

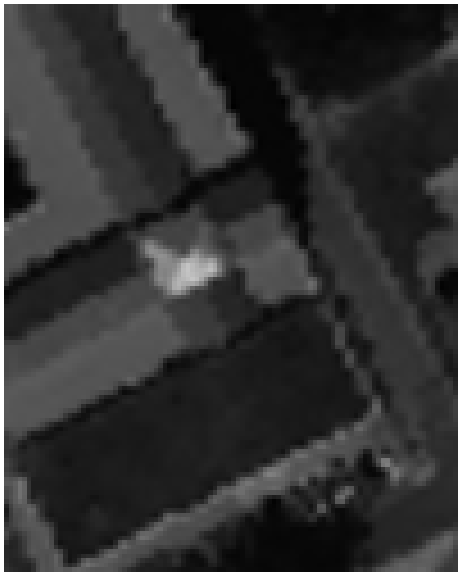


Figure 5.7: *The upper left* 125×100 *pixels of the superresolved image using bicubic interpolation. Average gradient is* 0.0244.

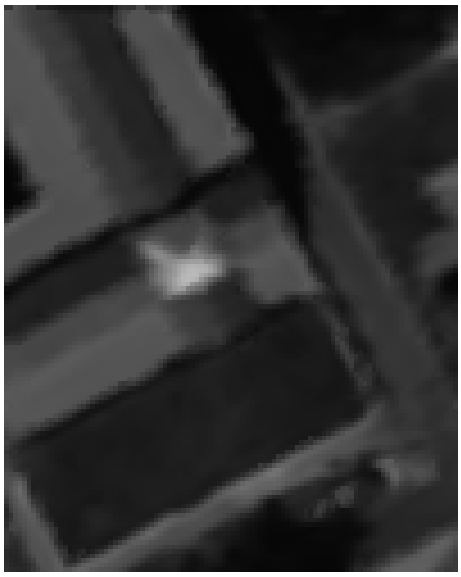


Figure 5.8: *The upper left* 125×100 *pixels of the superresolved image using our algorithm with a.) in Step 4. Average gradient is* 0.199.

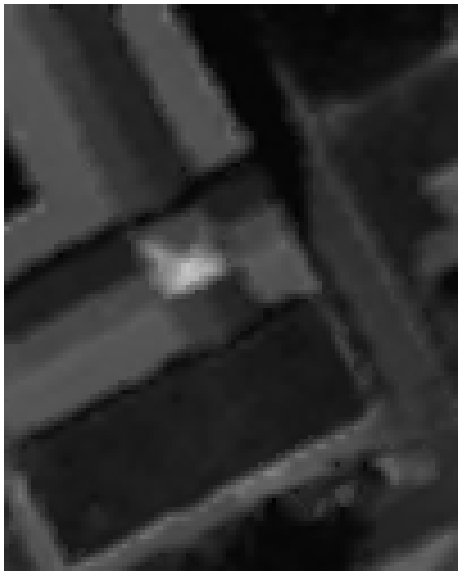


Figure 5.9: *The upper left* 125×100 *pixels of the superresolved image using our algorithm with c.) in Step 4. Average gradient is* 0.0213.

Generalizing this approach by using other anisotropic representation systems beyond shearlets, such as curvelets [12] and composite wavelets [52], is of interest. In addition, finding more sophisticated ways to implement the local directional information, beyond motion blurring, has the potential to improve superresolution results. This could be done through a variety of cutting-edge interpolation techniques [81], [1].

Chapter 6: Image Fusion with Wavelet Packets

Image fusion is, broadly construed, the merger of two or more images with different features in order to obtain a single image with the most desired features of each. A particular paradigm in the broad field of image fusion is that of pan-sharpening. The ambition in this case is to imbue the the low spatial resolution image with details from the high spatial resolution image, while retaining the spectral qualities of the low spatial resolution image. These images could be registered using the techniques developed Chapter 4, or could be a priori co-registered because both sensors are part of the same device, so that the images are captured almost simultaneously. Often, one considers co-registered multispectral and pan-chromatic images, and the ambition is to sharpen the multispectral images with details from the pan-chromatic band. Many contemporary satellite imaging systems produce co-registered multispectral and pan-chromatic images, including IKONOS, Quickbird, GeoEye and Landsat, making pan-sharpening a significant field of research. To create pan-sharpened multispectral images there are several well known algorithms, including IHS [98], [14], [57], PCA [93], [92], Brovey [110], Compressive Sensing [112], [76], and Framelet [40] based. For this study we concentrate on wavelet based methods and see what improvements can be realized using the wavelet packet trans-

form.

It has been known for some time that wavelet-based methods offer a powerful technique for image fusion [67], [111]. In wavelet-based image fusion, the images to be fused are decomposed according to a wavelet algorithm, for example, the Meyer-Mallat algorithm [105] which was discussed at a theoretical level in Chapter 2. Then, certain wavelet coefficients from one image decomposition are combined with those from the other image. The inverse wavelet algorithm is then applied, yielding a new image. If the coefficients to be mixed are chosen wisely, relevant features from one image can be integrated into the other image, giving a fused image that has features from both.

One situation that lends itself particularly well to wavelet-based approaches is the problem of pan-sharpening by fusing one image of high spatial resolution and another of lower resolution spatial resolution. After performing the wavelet decomposition, the high-pass coefficients from the high spatial resolution image are put in place of the high pass coefficients for the low spatial resolution image. This allows us to retain the textural features of the lower spatial resolution image, while adding high-pass features, thus sharpening the image. Many approaches to general image fusion have some degree of success in the specific problem of pan-sharpening, but suffer from color-distortion issues. Wavelet-based methods have the benefit of avoiding these distortions, and can be flexibility integrated with other fusion methods [2]. Wavelet-based pan-sharpening approaches have also given rise to techniques that counteract the electromagnetic spectral responses of sensors during the fusion process, thus circumventing a major technological stumbling point of

image fusion [89].

6.1 Wavelets for Image Fusion

Wavelets generated from the multiresolution analysis method have proved particularly significant in image processing applications. In practice, the discrete wavelet transform decomposes an image into high-pass and low-pass coefficients, then iteratively decomposes the high-pass coefficients. One thus acquires a coefficient tree, in which high and low pass features are represented, as seen in Figure 6.4. The node labels (a, b) correspond to the ath level and the bth node within the level. We note here that we are considering two dimensional signals, so at each level, our tree bifurcates into $2^2 = 4$ branches, as opposed to the $2^1 = 2$ branches in the examples in Chapter 2.

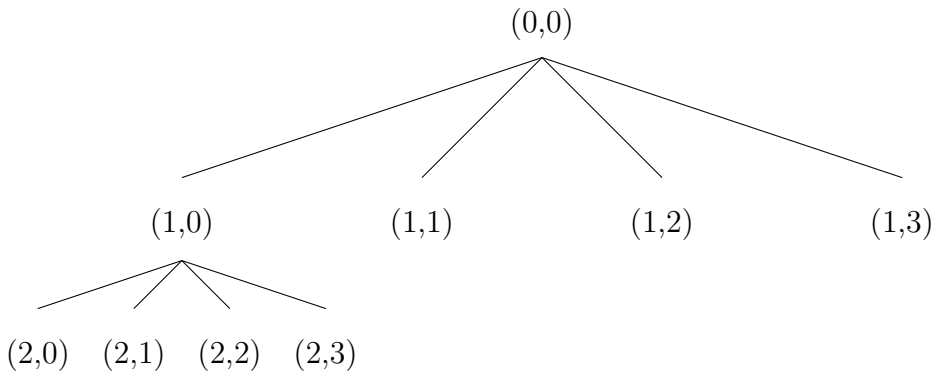


Figure 6.1: *An example of a 2D level 2 wavelet tree.*

Coifman, Meyer, and Wickerhauser addressed this issue in extending the discrete wavelet transform to decompose both the high and low pass coefficients at

each step, leading to the concept of the *discrete wavelet packet transform*. This is in distinction to the standard discrete wavelet transform, in which only the high pass coefficients are decomposed. One thus acquires a coefficient tree that is flat, in contradiction to the wavelet coefficient tree, as seen in Figure 6.2.

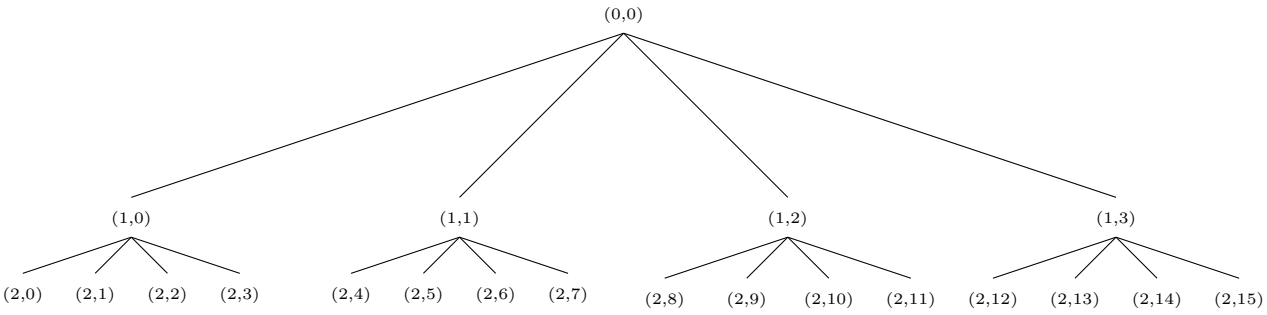


Figure 6.2: *An example of a 2D level 2 wavelet packet tree.*

Now any subset of the coefficients adhering to a disjoint dyadic decomposition will produce a wavelet basis. The best basis algorithm [21] developed by Wickerhauser gives us the ability to choose an optimal ONB, according to some entropy, to choose for the appropriate wavelet packet family. The best basis algorithm essentially works by recursively ascending the wavelet packet from the lowest level and furthest node to the first level and checking whether the parent or the collection of children nodes have a lower information cost. This yields a subset of the wavelet packet coefficient tree, as seen in Figure 6.3.

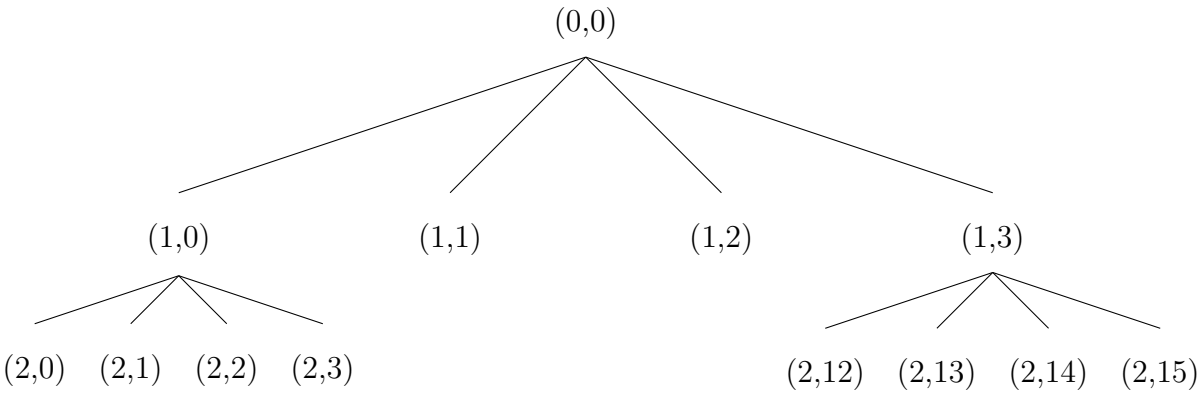


Figure 6.3: *An example of a 2D level 2 best basis wavelet packet tree.*

Building on the idea of Coifman and Wickerhauser, in [6], Benedetto, Czaja, and Ehler introduce the joint best basis algorithm. This algorithm sought a way to find the best basis representation among a collection of wavelet packet trees. To do this, first a joint entropy must be defined. The joint entropy $\mathcal{E}$ for a collection of wavelet packet trees $\mathcal{T} = \{T_1, \ldots, T_D\}$ is defined through a weighted ℓ_p norm:

$$\mathcal{E} := \sum_{i=1}^{D} w_i |E_i|^p,$$

where $0 < p \leq 2$ and $\{w_i\}_{i=1}^{D}$ is a collection of non-negative weights. Now with a joint entropy defined, the best basis algorithm can find an optimal basis for the collection of wavelet packet trees.

6.2 Wavelet Packets for Pan-Sharpening

Our approach [36] is to allow for significantly greater flexibility in wavelet-based fusion, by using wavelet packets instead of standard wavelets. The wavelet

packet tree has many coefficients that represent high pass features, so we can seek a more optimal way to infuse this information into the low spatial resolution image. While image fusion using wavelet packets has been studied [109], it is not at all clear how to optimally choose which coefficients to use in the fusion. We consider two variations of how to fuse using wavelet packets, depending on which coefficients to manipulate. The differing algorithms produce different pan-sharpened images, depending on which coefficients we choose to mix.

There is a natural question of how to choose which coefficients to change in the process of pan-sharpening. The wavelet decomposition iteratively decomposes the high-pass coefficients, so there is a natural coefficient to fuse with, as seen in Figure 6.4.

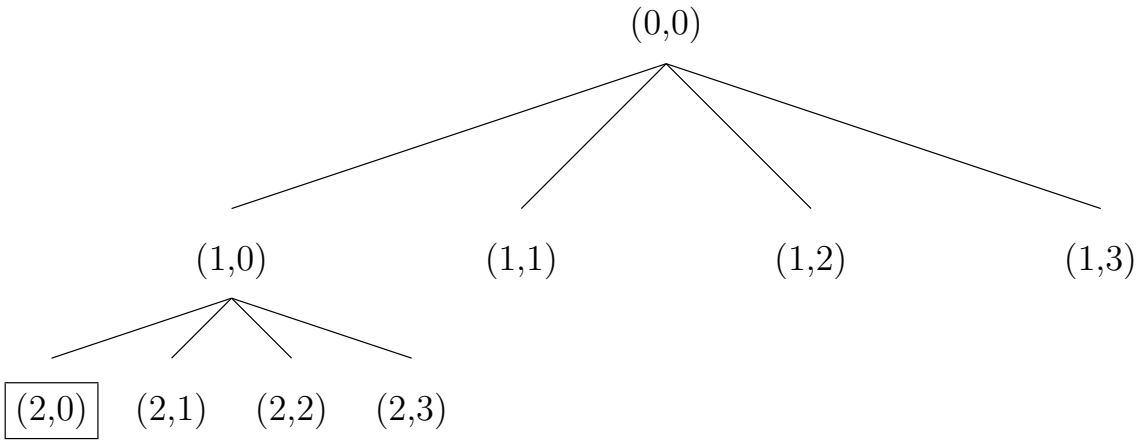


Figure 6.4: *An example of a 2D level 2 wavelet fusion. The boxed node is taken from the MS data, and unboxed nodes are taken from the panchromatic data*

In the case of wavelet packets, there are many coefficients that contain some degree of high pass information: those that represent high pass components of de-

composed low pass coefficients, and the decompositions of high pass components. The case in which we swap all coefficients that involve high frequencies shall be called WP2 fusion which is depicted in Figure 6.5; the case in which we swap only the high pass coefficients at the bottom of our tree shall be called WP1 fusion and is depicted in Figure 6.6.

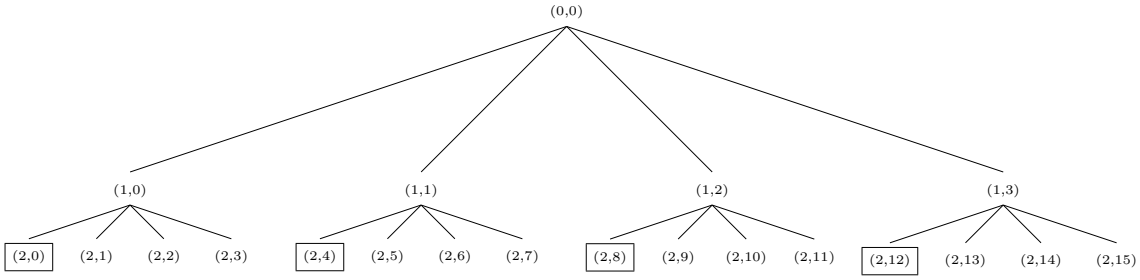


Figure 6.5: *An example of a 2D level WP1 fusion. Boxed nodes are taken from the MS data and unboxed nodes are taken from the panchromatic data.*

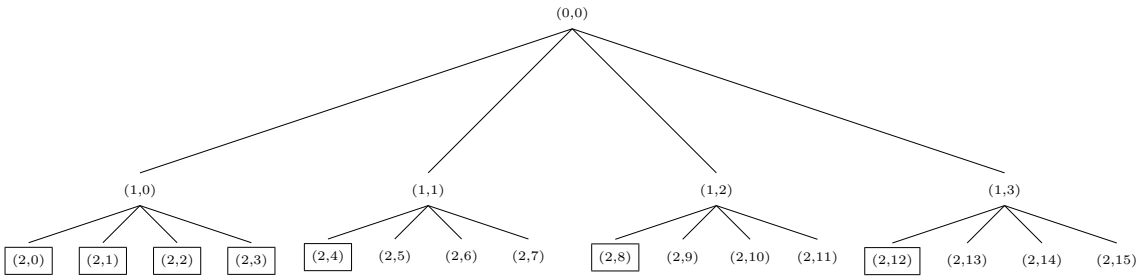


Figure 6.6: *An example of a 2D level WP2 fusion. Boxed nodes are taken from the MS data and unboxed nodes are taken from the panchromatic data.*

It is not clear a priori that one of these fusion flavors should be optimal, we will look at the analysis of a remote sensing data set to illustrate the benefits inherent in both WP1 and WP2.

6.2.1 Algorithms to be Tested

Given a data set containing a set of D multispectral bands, $MS = \{MS_i\}_{i=1}^{D}$ with corresponding panchromatic band, P, we seek a pan-sharpened version $MS^+ = \{MS_i^+\}_{i=1}^{D}$. Our general algorithm is summarized below for a chosen wavelet packet level ℓ and wavelet and scaling function pair ψ and ϕ.

1. Upsample MS to be the same spatial resolution of P. Denote this new collection of bands as $\widetilde{MS} = \{\widetilde{MS}_i\}_{i=1}^{D}$. The upsampling method can be chosen by the user; in our experiments, bicubic interpolation is used because it is simple, effective, and implemented on MATLAB.

2. For $i = 1, \ldots, D$, create a histogram matched version P_i of P in relation to $\widetilde{MS}_i$.

3. For $i = 1, \ldots, D$, find the level ℓ $2D$ wavelet packet decomposition of $\widetilde{MS}_i$ using ψ and ϕ. Denote the collection of wavelet packet trees as $T_{MS} = \{T_{MS}^i\}_{i=1}^{D}$.

4. For $i = 1, \ldots, D$, find the level ℓ $2D$ wavelet packet decomposition of P_i using ψ and ϕ. Denote the collection of wavelet packet trees as $T_P = \{T_P^i\}_{i=1}^{D}$.

5. For $i = 1, \ldots, D$, mix the wavelet packet coefficients from T_{MS}^i with T_P^i according to either WP1 or WP2. Denote the fused tree as T_i.

6. For $i = 1, \ldots, D$, find the inverse $2D$ wavelet packet transform for each T_i. Denote the pan-sharpened image as MS_i^+.

Optionally, we could also concentrate the spectral information by using the best basis algorithm and PCA to find a pan-sharpened spectrally concentrated representation which we denote as $MS^{+*} = \{MS_i^{+*}\}_{i=1}^{d}$. The authors in [6] explore how spatial/spectral fusion may be accomplished using wavelet packets and the joint best basis algorithm. Given a collection of wavelet packet trees, $\mathcal{T}$, and a joint best basis representation for the trees, $\bar{\mathcal{T}}$, we can collect the wavelet packet coefficients from $\bar{\mathcal{T}}$ into a $D \times N_T$ matrix C. D is the number of the wavelet packet trees which corresponds to the number of dimensions in the original data set. N_T is the number of wavelet packet coefficients, which is dependent on the size of the image decomposed by the wavelet packet algorithm and the type of wavelet used. Now a dimension reduction algorithm, such as PCA, can be used to reduce the dimensionality from D to d. The reduced dimension coefficients can now be reassembled into a tree format and then the inverse wavelet packet transform can be used to return the data to its original space. The algorithm for this method is summarized below

1-5. Same as steps 1-5 above.

6. Find the joint best basis representation of the collection of trees $\{T_i\}_{i=1}^{D}$ and denote as $\{T_i^B\}_{i=1}^{D}$.

7. Collect the coefficients from the terminal nodes of $\{T_i^B\}_{i=1}^{D}$ into a $N_T \times D$ matrix C where N_T is the number of coefficients in terminal nodes for each tree.

8. Using PCA followed by a projection down to $d \leq D$ dimensions, find the lower dimensional representation of C, denoted C'.

9. Create new wavelet packet trees $\{T_i^*\}_{i=1}^d$ with the joint best basis structure and coefficients corresponding C'.

10. For $i = 1, \ldots, d$, find the inverse $2D$ wavelet packet transform for each T_i^* and denote the pan-sharpened spectrally concentrated image as MS_i^{+*}.

6.3 Experiments and Results

6.3.1 Data

In conducting our experiments, we used the 2000 Data Fusion Contest data, which consisted of images taken over Hasselt, Belgium in 1999. The image was acquired from the Landsat 7 Enhanced Thematic Mapped Plus (ETM+). The data set contains seven multispectral bands covering the visual and infrared spectra, denoted bands 1-7, and one panchromatic band denoted band 8. The spectral windows and spatial resolution for the bands can be found in Table 6.1 and gray scale images for each band can be found in Figure 6.7.

Band Number	Spectral Window (nm)	Spatial Resolution (m)	Entropy
1	450-515	30	3.9904
2	525-605	30	4.3416
3	630-690	30	4.8394
4	750-900	30	6.0074
5	1550-1750	30	5.8962
6	1040-1250	60	3.5980
7	2090-2350	30	5.5004
8	520-900	15	4.8442

Table 6.1: *Spectral window, spatial resolution, and band entropy for Hasselt data set.*

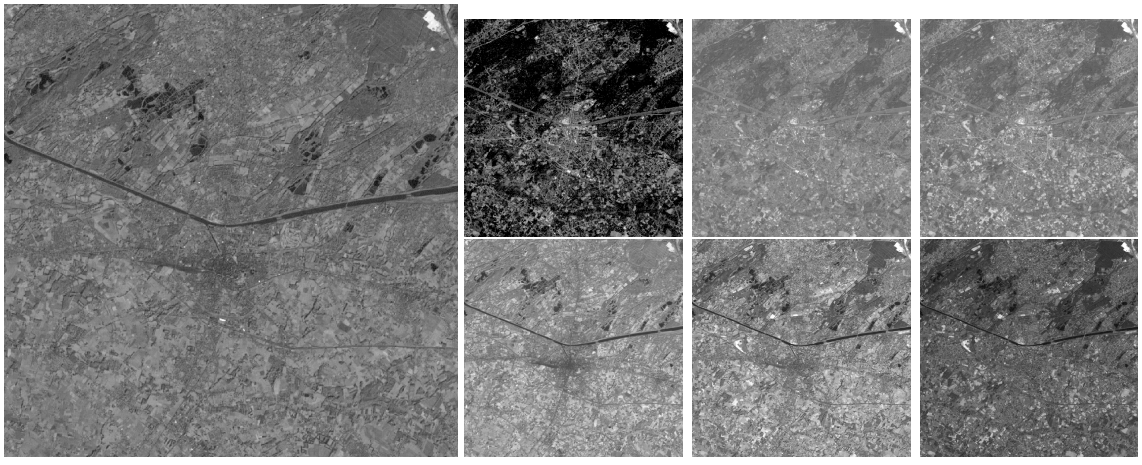


Figure 6.7: *The collection of bands from the Hasselt data set. On the left is the panchromatic band and then left to right top to bottom are multispectral bands 1,2,3,4,5,and 7. The data is courtesy of the 2000 IEEE GRSS Data Fusion Contest.*

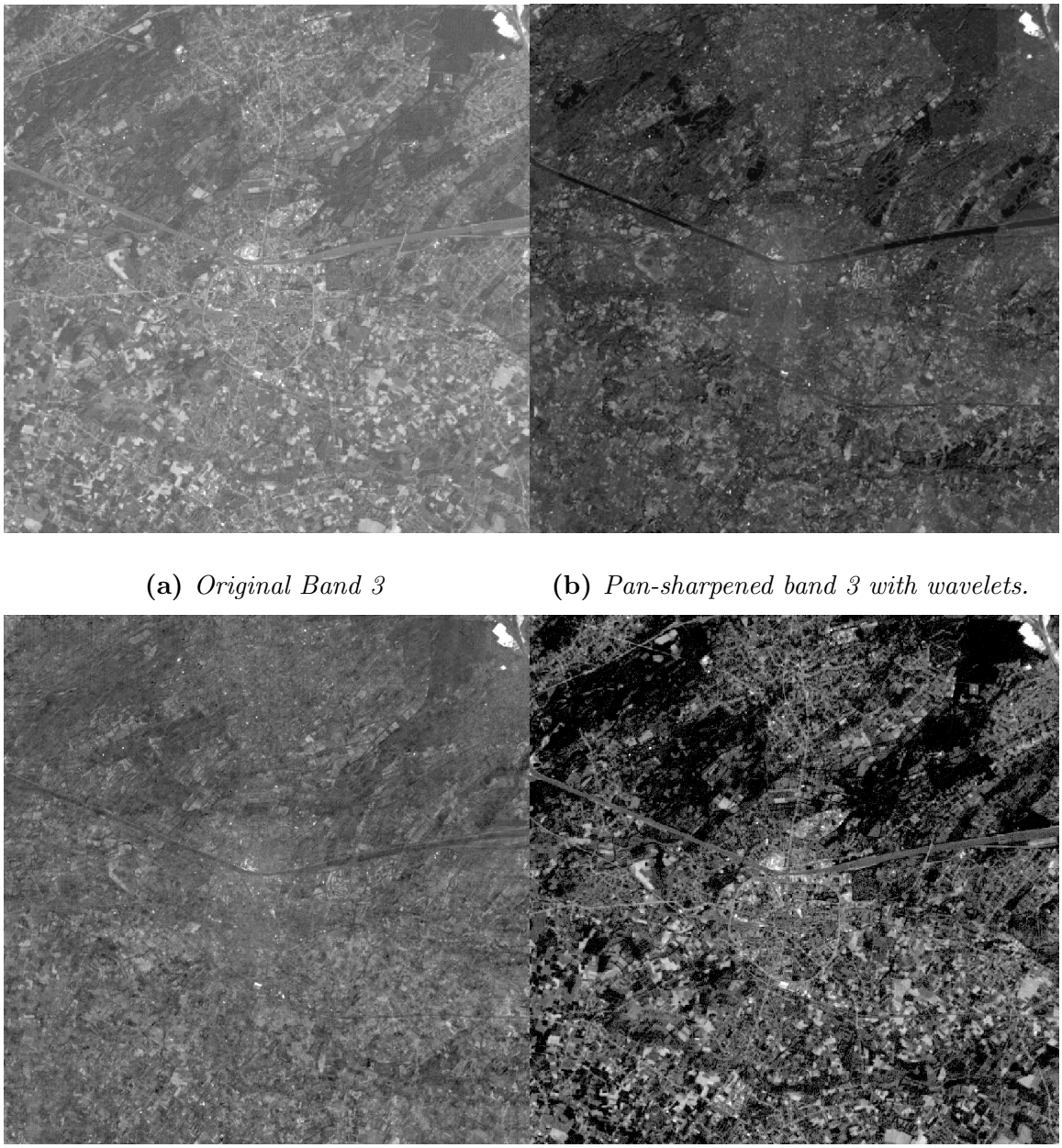
6.3.2 Pan Sharpening Experiment

We shall perform our experiments by fusing multispectral bands 1, 2, 3 with the panchromatic band 8. Bands 1, 2 and 3 have a spatial resolution of 30 meters and band 8 has a spatial resolution of 15 meters, so that our fusion can be understood as pan-sharpening.

Ideally, the quality of the fused results would be evaluated by classification against ground truth. No ground truth is available for this data, so we must use numerical results to judge the fusion algorithms. One approach to evaluating the quality of a fused image in the absence of ground truth is the use of correlation with the original image and entropy of the image [85], in order to indicate the extent to which the spectral quality of the bands was preserved by the fusion, and the extent to which additional spatial information was added. The spectral preservation quality was computed by measuring the correlation of a fused band with the original band; the higher this correlation, the more faithfully the spectral properties are preserved. To measure spatial information, entropy of the original image can be compared with the entropy of the fused images. We computed this entropy by first converting our images to grayscale, then computing the Shannon entropy $\epsilon := \sum_i p_i \log_2(pi)$, where p_i denotes the probability of a pixel being in the i^{th} bin.

We performed our decompositions by computing 6 and 7 levels of wavelet packet coefficients using the Daubechies wavelet of length 4. We then performed the coefficient mixing according to either the WP1 or WP2 algorithms, followed by an inverse wavelet packet transform. For reference, we also computed the entropy

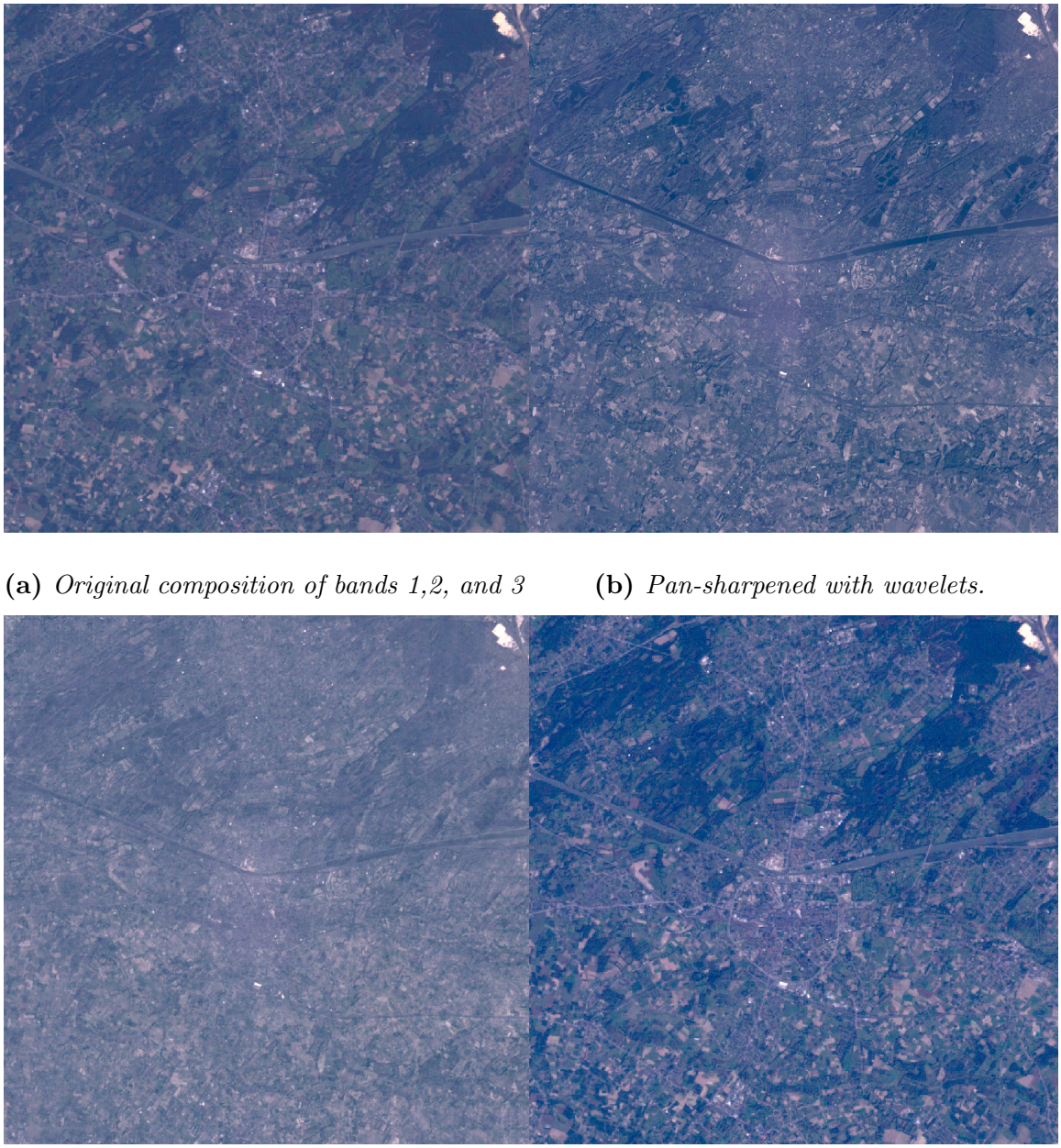
of the original bands, and the correlation and entropy values for standard wavelet fusion at levels 6 and 7.



(a) *Original Band 3* **(b)** *Pan-sharpened band 3 with wavelets.*

(c) *Pan-sharpened band 3 with WP1.* **(d)** *Pan-sharpened band 3 with WP2.*

Figure 6.8: *A comparison of the wavelet, WP1, and WP2 fusion methods for band 3.*



(a) *Original composition of bands 1,2, and 3*

(b) *Pan-sharpened with wavelets.*

(c) *Pan-sharpened with WP1.*

(d) *Pan-sharpened with WP2.*

Figure 6.9: *A comparison of the wavelet, WP1, and WP2 fusion methods for composition of bands 1, 2, and 3.*

Our results indicate that WP1 fusion achieves superior correlation and entropy values in all cases when compared to wavelet-based fusion. The WP2 fusion improves

upon the entropy of band 1 at level 7, but fails to match the increase given by wavelet fusion and WP1 fusion. However, the WP2 fusion gives very high correlation with the original bands, thus indicating excellent spectral preservation quality.

Band #	Correlation w/ Wavelet	Correlation w/ WP1	Correlation w/ WP2
1	.4114	.5364	.9739
2	.5749	.6610	.9746
3	.4414	.5531	.9791

Table 6.2: *Correlation values for fusion at level 6*

Band #	Original Entropy	Wavelet Entropy	WP1 Entropy	WP2
1	3.9904	4.2385	4.3777	4.0338
2	4.3416	4.4898	4.5749	4.2926
3	4.8394	4.9821	5.0896	4.7680

Table 6.3: *Entropy values for fusion at level 6*

Band #	Correlation w/ Wavelet	Correlation w/ WP1	Correlation w/ WP2
1	.3760	.5074	.9794
2	.5524	.6400	.9807
3	.4112	.5296	.9834

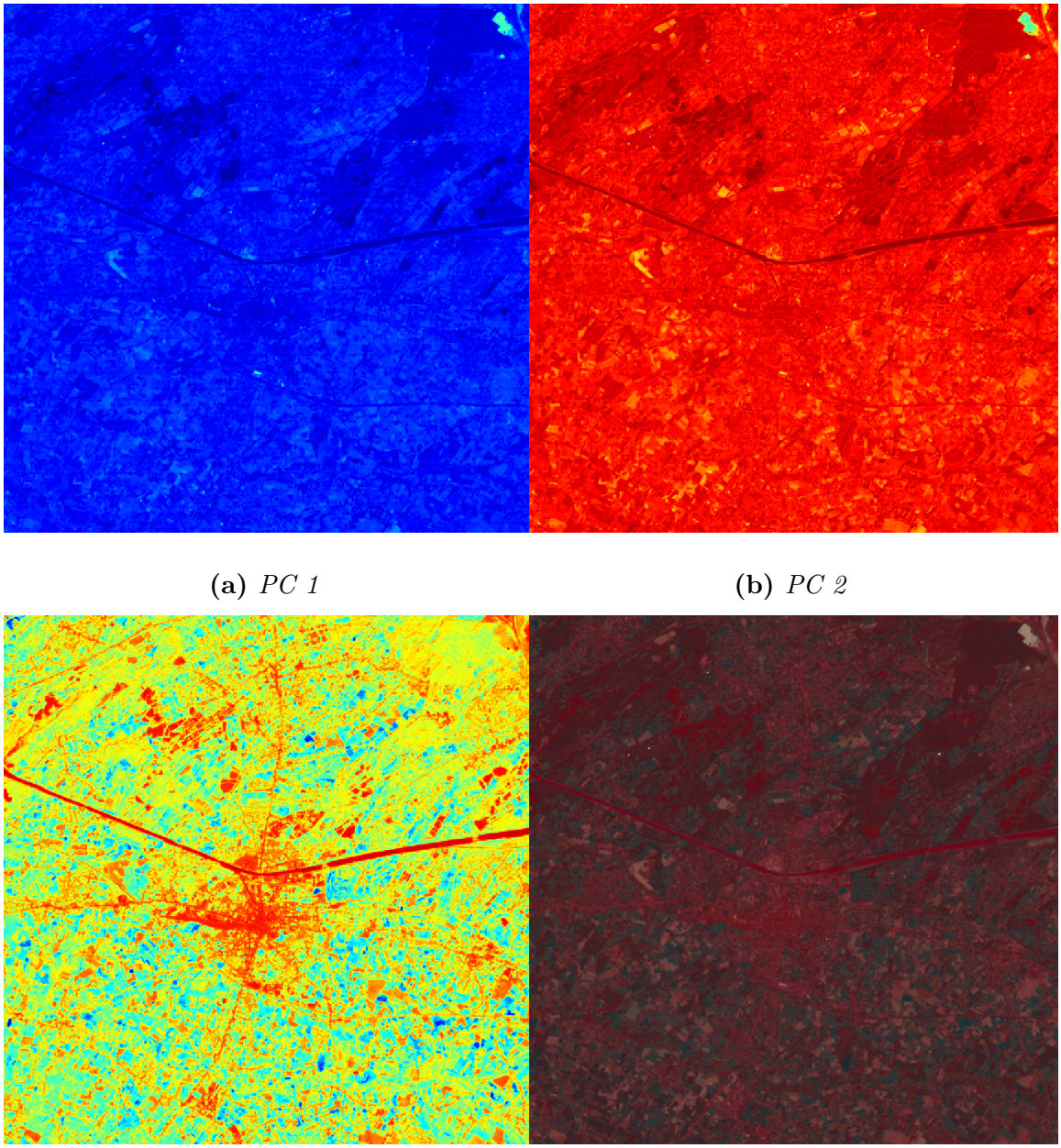
Table 6.4: *Correlation values for fusion at level 7*

Band #	Original Entropy	Wavelet Entropy	WP1 Entropy	WP2
1	3.9904	4.1855	4.3727	4.0680
2	4.3416	4.4523	4.5848	4.3204
3	4.8394	4.9744	5.0668	4.8128

Table 6.5: *Entropy values for fusion at level 7*

6.3.3 Spectrally Concentrated Pan-sharpening

For this experiment we looked at producing pan-sharpened versions of the seven multispectral bands then spectrally concentrating the information using PCA, as detailed above. In Figure 6.10 we see the first 3 principle components, together with a false-color image produced by combining these three principle components.



(a) *PC 1* **(b)** *PC 2*

(c) *PC 3* **(d)** *False Color RGB*

Figure 6.10: *Principle components 1,2, and 3, and a false color RGB rendering using the principle components.*

6.4 Conclusions and Future Directions

Wavelet packets offer a flexible approach to image fusion, particularly pan-sharpening. Our experiments have shown that wavelet packet WP1 fusion outperforms wavelet-based fusion by achieving superior spectral preservation quality and increased spatial information. While WP2 fusion was unable to achieve the same increased spatial information, it had excellent preservation of spectral preservation quality. The mixed results for WP2 suggest that the optimal wavelet packet fusion approach for pan-sharpening applications might be a mix between the WP1 algorithm and the WP1 algorithm, where some of the high pass coefficients are chosen but others are not. An implementation of an optimal wavelet packet fusion algorithm would offer a powerful new tool to the study of pan-sharpening.

Bibliography

[1] V. Algazi, G. E. Ford, and R. Potharlanka. Directional interpolation of images based on visual properties and rank order filtering. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1991.

[2] K. Amolins, Y. Zhang, and P. Dare. Wavelet based image fusion techniques-an introduction, review and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62:249–263, 2007.

[3] M. Avriel. *Nonlinear programming: analysis and methods*. Courier Dover Publications, 2003.

[4] J. J. Benedetto. *Harmonic Analysis and Applications*. CRC Press, 1996.

[5] J. J. Benedetto and R. L. Benedetto. A wavelet theory for local fields and related groups. *The Journal of Geometric Analysis*, 14(3):423–456, 2004.

[6] J. J. Benedetto, W. Czaja, and M. Ehler. Wavelet packets for time frequency analysis of multispectral data. *International Journal on Geomathematics*, 4(2):137–154, 2013.

[7] M. Berra. Gabor frame decomposition of evolution operators and applications. *Journal of Pseudo-Differential Operators and Applications*, 5(2), 2014.

[8] E. H. Bosch, A. Castrodad, J. S. Cooper, W. Czaja, and J. Dobrosotskaya. Multiscale and multidirectional tight frames for image analysis. In *Proceedings of SPIE, vol. 8750*, 2013.

[9] E. H. Bosch, W. Czaja, J. Murphy, and D. Weinberg. Anisotropic representations for superresolution of hyperspectral data. In *Proceedings of SPIE Defense+ Security*, 2015.

[10] W. L. Briggs. The DFT: An owners' manual for the discrete Fourier transform. *SIAM*, 1995.

[11] L. G. Brown. A survey of image registration techniques. *ACM computing surveys*, 24(4):325–376, 1992.

[12] E. J. Candès, L. Demanet, D. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Modeling & Simulation*, 5(3):861–899, 2006.

[13] E. J. Candès and D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise $\mathcal{C}^2$ singularities. *Communications on pure and applied mathematics*, 57(2):219–266, 2004.

[14] J. W. Carper, T. M. Lillesand, and R. W. Kiefer. The use of intensity-hue-saturation transformations for merging spot panchromatic and multispectral image data. *Photogrammetric Engineering and Remote Sensing*, 56(4):459–467, 1990.

[15] P. G. Casazza and G. Kutyniok. *Finite frames: Theory and applications*. Springer Science & Business Media, 2012.

[16] S. G. Chang, B. Yu, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing*, 9(9):1532–1546, 2000.

[17] T. Chang and C. Kuo. Texture analysis and classification with tree-structured wavelet transform. *IEEE Transactions on Image Processing*, 2(4):429–441, 1993.

[18] O. Christensen. *Frames and Bases: An Introductory Course*. Springer, 2008.

[19] O. Christensen, H. G. Feichtinger, and S. Paukner. *Handbook of Mathematical Methods in Imaging*, chapter Gabor analysis for imaging, pages 1271–1307. Springer New York, 2011.

[20] O. Christensen, B. Forster, and P. Massopust. Directional time frequency analysis via continuous frame. *arXiv preprint*, 2014.

[21] R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2):713–718, 1992.

[22] A. A. Cole-Rhodes, K. L. Johnson, J. LeMoigne, and I. Zavorin. Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *IEEE Transactions on Image Processing*, 12(12):1495–1511, 2003.

[23] N. Cornille, D. Garcia, M. A. Sutton, S. McNeill, and J.-J. Orteul. Automated 3-D reconstruction using a scanning electron microscope. In *SEM conference on experimental and applied mechanics*, 2003.

[24] W. Czaja, J. Dobrosotskaya, and B. Manning. Composite wavelet representations for reconstruction of missing data. In *International Society for Optics and Photonics*, 2013.

[25] W. Czaja, T. Doster, and J. M. Murphy. Wavelet packet mixing for image fusion and pan-sharpening. In *SPIE Defense+ Security*, 2014.

[26] W. Czaja and E. J. King. Isotropic shearlet analogs for $L^2(\mathbb{R}^k)$ and localization operators. *Numerical functional analysis and optimization*, 33(7-9):872–905, 2012.

[27] W. Czaja, J. M. Murphy, and D. Weinberg. Superresolution of remotely sensed images with anisotropic features. In *Proceedings of SAMPTA*, 2015.

[28] S. Dahlke, G. Kutyniok, P. Maass, C. Sagiv, H.-G. Stark, and G. Teschke. The uncertainty principle associated with the continuous shearlet transform. *International Journal of Wavelets, Multiresolution and Information Processing*, 6(02):157–181, 2008.

[29] S. Dahlke, G. Kutyniok, G. Steidl, and G. Teschke. Shearlet coorbit spaces and associated Banach frames. *Applied and Computational Harmonic Analysis*, 27(2):195–214, 2009.

[30] I. Daubechies. *Ten lectures on wavelets*. Society for industrial and applied mathematics, 1992.

[31] I. Daubechies, A. Grossman, and Y. Meyer. Painless nonorthogonal expansions. *Journal of Mathematical Physics*, 27(5):1271–1283, 1986.

[32] L. Ding and A. Goshtasby. On the Canny edge detector. *Pattern Recognition*, 34(3):721–725, 2001.

[33] M. N. Do and M. Vetterli. Contourlets: a directional multiresolution image representation. In *Proceedings of 2002 IEEE International Conference on Image Processing*, 2002.

[34] M. N. Do and M. Vetterli. The finite ridgelet transform for image representation. *IEEE Transactions on Image Processing*, 12(1):16–28, 2003.

[35] D. L. Donoho. Sparse components of images and optimal atomic decompositions. *Constructive Approximation*, 17(3):353–382, 2001.

[36] T. Doster. *Harmonic analysis inspired data fusion for applications in remote sensing*. PhD thesis, University of Maryland, College Park, 2014.

[37] R. J. Duffin and A. C. Schaeffer. A class of nonharmonic Fourier series. *Transactions of the American Mathematical Society*, pages 341–366, 1952.

[38] G. R. Easley, D. Labate, and F. Colonna. Shearlet-based total variation diffusion for denoising. *IEEE Transactions on Image Processing*, 18(2):260–268, 2009.

[39] G. R. Easley, D. Labate, and W.-Q. Lim. Sparse directional image representations using the discrete shearlet transform. *Applied and Computational Harmonic Analysis*, 25(1):25–46, 2008.

[40] F. Fang, G. Zhang, F. Li, and C. Shen. Framelet based pan-sharpening via a variational method. *Neurocomputing*, 2013.

[41] H. G. Feichtinger and K. H. Gröchenig. Banach spaces related to integrable group representations and their atomic decompositions, I. *Journal of Functional Analysis*, 86(2):307–340, 1989.

[42] M. Fornasier and H. Rauhut. Continuous frames, function spaces, and the discretization problem. *Journal of Fourier Analysis and Applications*, 11(3):245–287, 2005.

[43] J. Frikel. *Reconstructions in limited angle x-ray tomography: Characterization of classical reconstructions and adapted curvelet sparse regularization.* PhD thesis, Technische Universität München, 2013.

[44] H. Führ, L. Demanet, and F. Friedrich. Document and image compression. In *Beyond wavelets: New image representation paradigms*, pages 179–206. 2006.

[45] D. Gabor. Theory of communication. Part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26):429–441, 1946.

[46] P. Gader, A. Zare, R. Close, J. Aitken, and G. Tuell. Muufl gulfport hyperspectral and lidar airborne data set. REP 2013-570, University of Florida, Gainesville, FL,, 2013.

[47] L. Grafakos and C. Sansing. Gabor frames and directional time–frequency analysis. *Applied and Computational Harmonic Analysis*, 25(1):47–67, 2008.

[48] D. Griffin and J. S. Lim. Signal estimation from modified short-time Fourier transform. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 32(2):236–243, 1984.

[49] K. Gröchening. *Foundations of Time-Frequency Analysis*. Birkhäuser, 2001.

[50] P. Grohs, S. Keiper, G. Kutyniok, and M. Schäfer. α-molecules. *arXiv preprint arXiv:1407.4424*, 2014.

[51] K. Guo, G. Kutinyok, and D. Labate. Sparse multidimensional representations using anisotropic dilation and shear operators. In G. Chen and M. Lai, editors, *Wavelets and Splines (Athens, GA, 2005)*, pages 189–201. Nashboro Press, 2006.

[52] K. Guo, D. Labate, W.-Q. Lim, G. Weiss, and E. Wilson. Wavelets with composite dilations. *Electronic research announcements of the American Mathematical Society*, 10(9):78–87, 2004.

[53] A. Haar. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 69(3):331–371, 1910.

[54] S. Häuser. Fast finite shearlet transform. *arXiv preprint*, arXiv:1202.1773, 2012.

[55] S. Häuser and G. Steidl. Convex multiclass segmentation with shearlet regularization. *International Journal of Computer Mathematics*, 90(1):62–81, 2013.

[56] S. Häuser and G. Steidl. Fast finite shearlet transform: a tutorial. *Arxiv*, 1202.1773, 2014.

[57] R. Haydn, G. Dalke, J. Henkel, and J. Bare. Application of the IHS color transform to the processing of multisensor data and image enhancement. In *Proceedings of the International Symposium on Remote Sensing of Environment, First Thematic Conference: 19-25 January, 1982, Cairo, Egypt*, pages 599–616, 1982.

[58] C. Heil. History and evolution of the density theorem for Gabor frames. *Journal of Fourier Analysis and Applications*, 13(2):113–166, 2007.

[59] C. Heil. *A Basis Theory Primer: Expanded Edition*. Springer, 2010.

[60] S. Helgason. *The Radon Transform*, volume 2. Boston: Birkhäuser, 1999.

[61] E. Hernández, D. Labate, and G. Weiss. A unified characterization of reproducing systems generated by a finite family, II. *The Journal of Geometric Analysis*, 12(4):615–662, 2002.

[62] K. H. Hofmann and S. A. Morris. *The Structure of Compact Groups: A Primer for Students-a Handbook for the Expert*, volume 25. Walter de Gruyter, 2006.

[63] A. H. Karp. Bit reversal on uniprocessors. *SIAM review*, 38(1):1–26, 1996.

[64] R. Keys. Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 29(6):1153–1160, 1981.

[65] E. J. King. *Wavelet and frame theory: frame bound gaps, generalized shearlets, Grassmannian fusion frames, and p-adic wavelets*. PhD thesis, University of Maryland, 2009.

[66] E. J. King, G. Kutyniok, and X. Zhuang. Analysis of inpainting via clustered sparsity and microlocal analysis. *Journal of mathematical imaging and vision*, 48(2):205–234, 2014.

[67] R. King and J. Wang. A wavelet based algorithm for pan sharpening Landsat 7 imagery. In *Geoscience and Remote Sensing Symposium, 2001. IGARSS '01. IEEE 2001 International (Volume:2)*, pages 849–851, 2001.

[68] G. Kutyniok and D. Labate. *Shearlets: Multiscale analysis for multivariate data*. Springer Birkäuser, 2012.

[69] G. Kutyniok, J. Lemvig, and W.-Q. Lim. Compactly supported shearlets. In *Approximation Theory XIII: San Antonio 2010*, pages 163–186. Springer New York, 2012.

[70] D. Labate, W. Q. Lim, G. Kutinyok, and G. Weiss. Sparse multidimensional representation using shearlets. In *Proceedings of International Society for Optics and Phototronics: Optics and Phototronics*, 2005.

[71] D. Labate and P. Negi. 3D discrete shearlet transform and video denoising. In *SPIE Optical Engineering + Applications*, 2011.

[72] J. Le Moigne. Parallel registration of multisensor remotely sensed imagery using wavelet coefficients. In *SPIE's International Symposium on Optical Engineering and Photonics in Aerospace Sensing*, 1994.

[73] J. Le Moigne, W. J. Campbell, and R. P. Cromp. An automated parallel image registration technique based on the correlation of wavelet features. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8):1849–1864, 2002.

[74] J. Le Moigne and R. F. Cromp. Wavelets for remote sensing image registration and fusion. In *International Society for Optics and Photonics: Aerospace/Defense Sensing and Controls*, 1996.

[75] T. S. Lee. Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):959–971, 1996.

[76] S. Li and B. Yang. A new pan-sharpening method using a compressed sensing technique. *Geoscience and Remote Sensing, IEEE Transactions on*, 49(2):738–746, 2011.

[77] A. P. Lobo and P. C. Loizou. Voiced/unvoiced speech discrimination in noise using Gabor atomic decomposition. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*, volume 1, 2003.

[78] D. G. Lowe. Object recognition from local scale-invariant features. In *The proceedings of the seventh IEEE international conference on computer vision*, volume 2, 1999.

[79] Y. Lu, S. Joshi, and J. M. Morris. Noise reduction for NMR FID signals via Gabor expansion. *IEEE Transactions on Biomedical Engineering*, 44(6):512–528, 1997.

[80] J. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical image analysis*, 2(1):1–36, 1998.

[81] S. Mallat and G. Yu. Super-resolution with sparse mixing estimators. *IEEE Transactions on Image Processing*, 19(11):2889–2900, 2010.

[82] S. G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[83] B. Manning. *Composite multi resolution analysis wavelets*. PhD thesis, Washington University in Saint Louis, 2012.

[84] D. Marquadt. An algorithm for least-squares estimation of non-linear parameters. *Journal of SIAM*, 11:431–441, 1963.

[85] N. Memarsadeghi, J. Le Moigne, and D. Mount. Image fusion using cokriging. pages 2518–2521, 2006.

[86] Y. Meyer and R. Coifman. *Wavelets: Calderón-Zygmund and multilinear operators*, volume 48. Cambridge University Press, 1997.

[87] Q.-G. Miao, C. Shi, P.-F. Xu, M. Yang, and Y.-B. Shi. A novel algorithm of image fusion using shearlets. *Optics Communications*, 284(6):1540–1547, 2011.

[88] J. Nunez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol. Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Transactions on Geoscience and Remote Sensensing*, 37(3):1204–1211, 1999.

[89] X. Otazu, M. González-Audícana, O. Fors, and J. Núñez. Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods. *IEEE Transactions on Geoscience and Remote Sensing*, 43(10):2376–2385, 2005.

[90] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: a technical overview. *IEE Signal Processing Magazine*, 20(3):21–36, 2003.

[91] W. Rudin. *Fourier analysis on groups*. Number 12 in Tracts in Mathematics. John Wiley & Sons, 1990.

[92] V. P. Shah, N. H. Younan, and R. L. King. An efficient pan-sharpening method via a combined adaptive PCA approach and contourlets. *Geoscience and Remote Sensing, IEEE Transactions on*, 46(5):1323–1335, 2008.

[93] V. K. Shettigara. A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set. *Photogrammetric Engineering and Remote Sensing*, 58(5):561–567, 1992.

[94] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger. Shiftable multiscale transforms. *IEEE Transactions on Information Theory*, 38(3):587–607, 1992.

[95] W. Sun and X. Zhou. On Kadec's 1/4-theorem and the stability of Gabor frames. *7*, 7(2):239–242, 1999.

[96] P. Thévenaz, U. E. Ruttiman, and M. Unser. A pyramid approach to subpixel registration based on intensity. *IEEE Transactions on Image Processing*, 7(1):27–41, 1998.

[97] P. Thévenaz and M. Unser. Optimization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing*, 9(12):2083–2099, 2000.

[98] T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang. A new look at IHS-like image fusion methods. *Information Fusion*, 2(3):177–186, 2001.

[99] M. Unser. Texture classification and segmentation using wavelet frames. *IEEE Transactions on Image Processing*, 4(11):1549–1560, 1995.

[100] M. Unser, A. Aldroubi, and M. Eden. The L^2-polynomial spline pyramid. *IEEE Transactions in Pattern Analysis and Machine Intelligence*, 15(4):364–379, 1993.

[101] A. Vedaldi. An open implementation of the SIFT detector and descriptor. *UCLA CSD Tech. Report 070012*, 2006.

[102] O. R. Vincent and O. Folorunso. A descriptive algorithm for sobel image edge detection. In *Proceedings of Informing Science & IT Education Conference (InSITE)*, 2009.

[103] D. Walnut. Continuity properties of the Gabor frame operator. *Journal of Mathematical Analysis and Applications*, 165(2):479–504, 1992.

[104] M. V. Wickerhauser. *Adapted Wavelet Analysis from Theory to Software*. AK Peters Ltd., 1994.

[105] M. V. Wickerhauser. *Mathematics for Multimedia*. Academic Press, 2004.

[106] N. Wiener. Tauberian theorems. *Annals of mathematics*, pages 1–100, 1932.

[107] P. Wojtaszczyk. *A Mathematical Introduction to Wavelets*. Cambridge University Press, 1997.

[108] I. Zavorin and J. Le Moigne. Use of multiresolution wavelet feature pyramids for automatic registration of multisensor imagery. *IEEE Transactions on Image Processing*, 14(6):770–782, 2005.

[109] W. Zhang and J. Kang. Quickbird panchromatic and multi-spectral image fusion using wavelet packet transform. In *Intelligent Control and Automation*, pages 976–981, 2006.

[110] Y. Zhang. Problems in the fusion of commercial high-resolution satellite as well as Landsat 7 images and initial solutions. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(4):587–592, 2002.

[111] J. Zhou, D. Civco, and J. Silander. A wavelet transform method to merge Landsat TM and SPOT panchromatic data. *International Journal of Remote Sensing*, 19(4):743–757, 1998.

[112] X. X. Zhu and R. Bamler. A sparse image fusion algorithm with application to pan-sharpening. *Geoscience and Remote Sensing, IEEE Transactions on*, 51(5):2827–2836, 2013.

[113] M. Zibulski and Y. Y. Zeevi. Analysis of multiwindow Gabor-type schemes by frame methods. *Applied and Computational Harmonic Analysis*, 4(2):188–221, 1997.