

**MULTISENSOR IMAGE REGISTRATION
UTILIZING THE
LoG FILTER AND FWT**

Karl C. Walli

**Bachelors of Science, Michigan Technological University, 1991
Masters of Science, Joint Military Intelligence College, 1995**

**A thesis submitted in partial fulfillment of the
requirements for the degree of Masters of Science
in the Chester F. Carlson Center for Imaging Science
of the College of Science
Rochester Institute of Technology**

June, 2003

Signature of Author:_____

Accepted by:_____

Coordinator, M.S. Degree Program

Report Documentation Page

Form Approved
OMB No. 0704-0188

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

1. REPORT DATE 23 FEB 2004	2. REPORT TYPE N/A	3. DATES COVERED June 2003 - June 2003			
4. TITLE AND SUBTITLE Multisensor Image Registration Utilizing The LoG Filter And FWT		5a. CONTRACT NUMBER			
		5b. GRANT NUMBER			
		5c. PROGRAM ELEMENT NUMBER			
6. AUTHOR(S) Karl C. Walli		5d. PROJECT NUMBER			
		5e. TASK NUMBER			
		5f. WORK UNIT NUMBER			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Rochester Institute of Technology Rochester, New York		8. PERFORMING ORGANIZATION REPORT NUMBER CI04-124			
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) The Department of the Air Force AFIT/CCIA, BLDG 125 2950 P Street WPAFB OH 45433		10. SPONSOR/MONITOR'S ACRONYM(S)			
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)			
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited					
13. SUPPLEMENTARY NOTES The original document contains color images.					
14. ABSTRACT This thesis examines the utility of automated image registration techniques developed by the author. The major thrusts of this research include using the Laplacian of Gaussian (LoG) filter to automatically determine ground control points (GCPs) and wavelet theory for multiresolution analysis. Additionally, advances in both composite and predictive transformations will be covered. The defense will include an overview of the processes involved in general image registration and specifically how they pertain to automation with the techniques utilized in this thesis. Use of the LoG filter to extract semi-invariant GCPs, development of automated point matching schemas, and the use of matrix transformations for efficient management of affine image relationships will be explained in detail. Additionally, the ability to apply statistical analysis to both local and image wide sets of GCPs will be discussed. The student developed software application, LoG Wavelet Registration (LoGWaR). will demonstrate the utility of these techniques for processing large datasets such as LANDSAT and how integration of these features can provide both power and flexibility when registering multiresolution and/or multisensor images. Automation techniques will be highlighted, demonstrating the strengths and weaknesses when applied to images with high degrees of parallax, cloud-cover, and other types of temporal change. Specific applications, such as wavelet sharpening and spectral unmixing will be addressed as it pertains to current research.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 147	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

**CHESTER F. CARLSON
CENTER FOR IMAGING SCIENCE
COLLEGE OF SCIENCE
ROCHESTER INSTITUTE OF TECHNOLOGY
ROCHESTER, NEW YORK**

CERTIFICATE OF APPROVAL

M.S. DEGREE THESIS

The M.S. Degree Thesis of Karl C. Walli has been examined and approved by the thesis committee as satisfactory for the thesis requirement for the Master of Science Degree.

Dr. John Schott, Thesis Advisor

Dr. Harvey Rhody

Dr. Joel Kastner

Date

THESIS RELEASE PERMISSION
ROCHESTER INSTITUTE OF TECHNOLOGY
COLLEGE OF SCIENCE
CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE

Title of Thesis:

**MULTISENSOR IMAGE REGISTRATION
UTILIZING THE
LOG FILTER AND FWT**

I, Karl C. Walli, hereby grant permission to the Wallace Memorial Library of the Rochester Institute of Technology to reproduce my thesis in whole or part. Any reproduction will not be for commercial use or profit.

Signature: _____

Date: _____

TABLE OF CONTENTS

1. Abstract/Acknowledgements/Glossary.....	v-vii
2. Introduction.....	1
3. Background and Theory	
3.1 Image Registration/Warping.....	4
3.2 Resampling.....	7
3.3 Laplacian of Gaussian (LoG) Filter.....	10
3.4 Wavelet Analysis.....	13
4. Registration Process.....	20
4.1 Image Preparation – Finding “Common Ground”.....	21
4.2 LoG Thresholding.....	25
4.3 Point Matching.....	26
4.4 Transform Model.....	32
4.5 Image Warping/Sampling.....	34
4.6 Automated GCP Selection.....	35
4.7 Registration Accuracy.....	38
4.8 Predictive Transforms.....	42
5. Composite Transforms.....	44
6. Registration Testing.....	53
7. Additional Applications.....	77
8. Summary/Key Research Advancements.....	87
10. Bibliography.....	91
11. Wavelet Appendix A.....	92
11. Sharpening Appendix B.....	105
12. LoGWaR Manual Appendix C.....	108

Chapter 1

ABSTRACT

LOGWAR MULTISENSOR IMAGE
REGISTRATION

This thesis examines the utility of automated image registration techniques developed by the author. The major thrusts of this research include using the Laplacian of Gaussian (LoG) filter to automatically determine ground control points (GCPs) and wavelet theory for multiresolution analysis. Additionally, advances in both composite and predictive transformations will be covered.

The defense will include an overview of the processes involved in general image registration and specifically how they pertain to automation with the techniques utilized in this thesis. Use of the LoG filter to extract semi-invariant GCPs, development of automated point matching schemas, and the use of matrix transformations for efficient management of affine image relationships will be explained in detail. Additionally, the ability to apply statistical analysis to both local and image wide sets of GCPs will be discussed.

The student developed software application, LoG Wavelet Registration (LoGWaR). will demonstrate the utility of these techniques for processing large datasets such as LANDSAT and how integration of these features can provide both power and flexibility when registering multiresolution and/or multisensor images.

Automation techniques will be highlighted, demonstrating the strengths and weaknesses when applied to images with high degrees of parallax, cloud-cover, and other types of temporal change. Specific applications, such as “wavelet sharpening” and “spectral unmixing” will be addressed as it pertains to current research.

ACKNOWLEDGMENTS

The author wishes to thank his wife for her staunch support and unquestionable love; without which, this research and thesis would not be possible.

GLOSSARY

Affine Transform: A subset of polynomial transformations that include shift, rotation, scale, and skew.

Dyadic Power: A power of two; used in the Fast Wavelet Transform to maintain proper dimension constraints for multiresolution analysis.

Hyperspectral: Image datasets that contain tens, to hundreds, of spectral bands.

Hypertemporal: Video datasets that contain tens to hundreds of frames per second.

Multispectral: Image datasets with four to tens of spectral bands.

Multisensor: Images that contain similar spatial content taken from different sensors.

Multiresolution: The ability to decimate an image into several spatial frequency subbands for analysis lays the foundation for Wavelet theory.

Polynomial Transform: The generic form of a spatial transformation that can be utilized to relate two images via global equations.

Resampling: Changing the number of pixels in an image, normally done through pixel averaging, sampling or replication.

Wavelets: By iteratively stripping off the highest spatial frequency components from an image (decimating), it is possible to retain those frequencies in isolated subbands for analysis.

INTRODUCTION

With the rapid advancement of both hyperspectral and hypertemporal imaging capabilities, the need for automated registration of image bands and frames with each other and with an ever-growing database of related images is critical. Similarly, for low light conditions such as astronomy, analysts are often producing long dwell composite images (utilizing long integration times or by “stacking” several individual images). These techniques all require precise registration of images, whether it’s for change detection, spectral unmixing, or to maximize the S/N ratio of the output image.

This registration process can be very slow and tedious when done by supervised registration, when an analyst chooses similar reference locations within images as ground control points (GCPs) and generates the transformation operation necessary for registration. So, it is the attempt of this research to add automation to this registration process through the use of spatial frequency analysis, edge filtering, point matching, and statistical analysis. The proposed registration technique utilizes comparison of semi-invariant features (edge detail) within a scene to correlate images/spectral bands. With the increasing processing speeds of today’s computers and the continuing sophistication of edge detection/filtering techniques, point matching, and statistical analysis, it is possible to fully automate this task.

As is often the case, it is desirable to register high-resolution (panchromatic) images with lower-resolution (multispectral) images. If this can be accomplished, it is possible to allow the strengths of each sensor to compensate for the inherent weaknesses of the other, so that analysts can efficiently exploit the spatial and spectral characteristics of the fused data simultaneously.

The goal of this thesis is to provide a robust technique for automated multisensor image registration and the development of software to implement these techniques (in the IDL programming environment). Correction for the basic geometric distortions such as shift, rotation, and scale between images will be covered in detail. Wavelet analysis (image resolution pyramids) will be utilized to decompose higher resolution images to the equivalent frequency content of a lower resolution image. This will allow automated registration of multi-sensor images utilizing the Laplacian of Gaussian (LoG) filter and automatic point matching techniques. The capabilities of this LoG Wavelet Registration (LoGWaR) technique will be demonstrated on both test data and real multisensor datasets.

Finally, useful applications for this technique such as “sharpening” and “spectral unmixing” will be developed and applied to datasets of interest so that analysts can efficiently exploit the spatial and spectral characteristics of the data simultaneously. This is necessary for applications such as “sharpening”, where the high frequency components of the higher resolution image are utilized to determine detail in the lower resolution image. Additional applications include pure “end member” selection for spectral “end-member” libraries (critical for “step-wise” unmixing), image stacking to increase the S/N, and change detection.

WORK STATEMENT

The Objectives of this research were to:

1. Develop a technique, using the Laplacian of Gaussian (LoG) filter, to identify regions with similar rates-of-variation within a scene as candidates points for relating the datasets (for automated Ground Control Point selection).
2. Develop/Implement point matching algorithms to automate the registration of images that vary with shift, rotation, and scale. Relate these images through the use of global affine and polynomial equations.
3. Develop/Implement criteria to determine “Goodness” of registration on both synthetic and real datasets.
4. Develop/Implement wavelet algorithms to deconstruct images by iteratively stripping off high frequency components.
5. Develop applications utilizing this registration technique that will be useful in “sharpening” and “spectral unmixing”.
6. Develop/Implement composite transforms to automatically manage and cascade numerous affine manipulations into a single mathematical expression to reduce transformation degradations.
7. Develop predictive transformation techniques, to relate multiresolution datasets, by registering the data at the lowest common resolution.
8. Implement registration code in a “user-friendly” Graphical User Interface (GUI); actualized in the LoGWaR software application.

BACKGROUND AND THEORY

Research for this thesis has centered around three critical areas necessary to accomplish the aforementioned objectives: these areas are image registration, wavelet analysis, and point matching theory. Pertinent areas within image registration and point matching include relating images using invariant characteristics within a scene, matching those characteristics and then using this information to develop a polynomial equation to transform one dataset into another with care given to the effects of resampling. Wavelet analysis is utilized here for the ability to relate multisensor/multiresolution images. This is due to its ability to gracefully degrade high-resolution imagery to a comparable frequency content of its lower-resolution counterpart. Wavelet analysis will also provide a useful mechanism for “sharpening” images, later in Chapter 7.

3.1 Image Registration/Warping

The goal of image registration is to provide spatial commonality between two datasets. Many aspects of remote sensing involve the comparison of similar datasets over time (temporal change) and/or how individual image planes vary spectrally (spectral change) within a spectral cube. With both of these examples, it is inherently necessary to register two or more images together so that their spatial values can be related directly, thus reducing the spatial variability as much as possible.

3.1.1 Polynomial Transformation

Schott maintains that, the formatting and processing of GIS data rely very heavily on our ability to transfer spatial data into a common coordinate system (registration) and to resample the data so that we can easily access and process information from the same spatial location simultaneous (Schott 1997). To do this, we can define a coordinate system such that, x_{ref} and y_{ref} designate points in the reference coordinate system. While, x_{warp} and y_{warp} can represent coordinates in the secondary image that we plan to warp (transform) into the coordinates of

the reference image. Registration of the two images requires us to relate both coordinate systems; which can often be accomplished with a least-squares polynomial fit. It has become common practice, in the user community to use a generic polynomial model for registering images to each other and to maps (Schowengerdt 1997). The general expression for this transform is:

$$(3.1) \quad x_{warp} = \sum_{i=0}^N \sum_{j=0}^{N-1} a_{ij} x_{ref}^j y_{ref}^i + \epsilon_x$$

$$(3.2) \quad y_{warp} = \sum_{i=0}^N \sum_{j=0}^{N-1} b_{ij} x_{ref}^j y_{ref}^i + \epsilon_y$$

Once expanded, these equations become:

$$(3.3) \quad x_{warp} = a_{00} + a_{01}x_{ref} + a_{10}y_{ref} + a_{11}x_{ref}y_{ref} + a_{02}x_{ref}^2 + a_{20}y_{ref}^2 \dots \epsilon_x$$

$$(3.4) \quad y_{warp} = b_{00} + b_{01}x_{ref} + b_{10}y_{ref} + b_{11}x_{ref}y_{ref} + b_{20}x_{ref}^2 + b_{02}y_{ref}^2 \dots \epsilon_x$$

Here, N represents the “order” of the polynomial that will be used for the transform. If the input imagery has been processed accurately for systematic distortions, a linear polynomial may suffice for further correction. At worst, a quadratic polynomial (N equals 2) is sufficient for most problems in satellite remote sensing where the terrain relief is small and the FOV is not large (Schowengerdt 1997).

Higher-order polynomial terms are required to correct for ever more complex relationships between the two images. If the transform from the warp coordinate system to the reference is represented by a fairly linear relationship, only the first three terms are necessary to relate them (plus any residual error represented by ϵ). The $N=0$ terms (a_{00} and b_{00}) would represent a shift of the origin, while the $N=1$ terms represents a scaling, rotation, shear, and perspective change from one coordinate system to the next. Some of these terms represent the “affine” (linear polynomial) transformation. Affines can simultaneously accommodate shift, scale and rotation and can be written in a compact vector-matrix form (Schowengerdt 1997). This general form will be utilized due to its ease of coding and general utility in registering remotely sensed images:

$$(3.5) \quad \begin{bmatrix} x_{warp} \\ y_{warp} \end{bmatrix} = \begin{bmatrix} a_{10} & a_{01} \\ b_{10} & b_{01} \end{bmatrix} \begin{bmatrix} x_{ref} \\ y_{ref} \end{bmatrix} + \begin{bmatrix} a_{00} \\ b_{00} \end{bmatrix}$$

3.1.2 Ground Control Point Matching

In order to determine the polynomial relationship between the images, it is necessary to select similar locations or “Ground Control Points” (GCPs) within the scenes that can be used to relate the two images (figure 3.1.1). Although this process can be done manually through “supervised” GCP selection, it is quite tedious and it is the intention of this thesis to add a robust automation to GCP selection and matching between reference and warp images. Techniques, most notably cross-correlation, have been successfully utilized for this purpose in the past, but require a sensor’s pointing information for accurate registration. Unfortunately, the cross-correlation technique is very sensitive to variations in rotation, scale and parallax. This paper develops a new approach, based on the LoG filter and wavelet techniques, to facilitate multisensor registration. The feature matching is shift, rotation, and scale invariant and has demonstrated a robust performance against images with localized parallax and terrain relief.

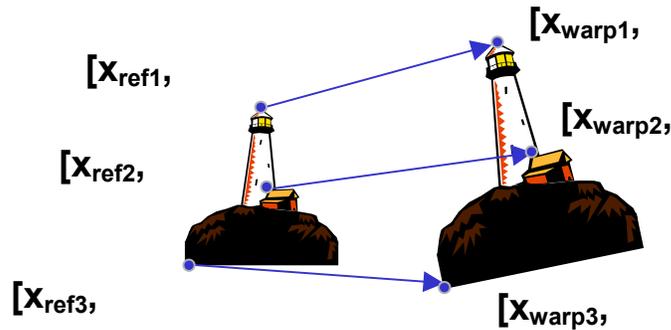


Figure 3.1.1: GCP selection for registration of two similar objects.

The x_{warp} , y_{warp} and x_{ref} , y_{ref} values of the GCPs from each image generate the input for a least-squares regression that is used to solve for the a_{ij} and b_{ij} coefficients of equations 3.1-3.5. As with any regression solution, the input data should cover the entire solution space, and in general, the solutions should not be extended beyond the sample space (Schott 1997). This becomes even more critical when higher order effects, like lens aberrations, may effect only a portion of one image.

Following this logic, the creation of a transform from matching point sets is relatively straightforward. This is especially true for supervised image registration, where an analyst selects similar feature pixels in both images and these point sets are utilized to feed into a regression solution. The difficulty is in developing robust algorithms that accurately and automatically relate the reference image to the warp image. Any automated registration scheme, will “live or die” on its ability to match like-features in the two images. This area is critical to the success of multisensor image registration and will be covered extensively in Chapter 4.

3.2 Resampling

Nearest neighbor resampling and bilinear interpolation are two well known techniques used for determining pixel intensity levels (grayscale values) when registering images. However, if a more robust resampling approach is required, cubic convolution will often produce the superior results. All have pluses and minuses associated with them that necessitate review.

Nearest neighbor transforms compute the location of the new pixel and associate that location’s grayscale value with the nearest pixel’s value from the original (pre-warped) image, (figure 3.2.1). This technique is easy to implement, computationally fast and radiometrically accurate, since it does not introduce any new grayscale values into the warped image. The negative aspect of this technique is primarily in its appearance, since “stair-stepping” artifacts can occur on edges (straight lines can have a blocky appearance) within an image.

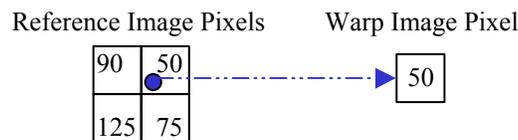


Figure 3.2.1: Nearest Neighbor grayscale resampling.

Bilinear interpolation computes the grayscale value at the transformed location by interpolating the value of the four nearest pixels based on their distance from the new location

(figure 3.2.1). This can be accomplished with the following formula, where $F(x,y)$ is the grayscale value of the warped pixel:

$$(3.6) \quad F(x,y) = \frac{\frac{1}{d_1} F(x_1, y_1) + \frac{1}{d_2} F(x_2, y_2) + \frac{1}{d_3} F(x_3, y_3) + \frac{1}{d_4} F(x_4, y_4)}{\frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3} + \frac{1}{d_4}}$$

$$F(x,y) = \frac{\frac{1}{\sqrt{.25^2 + .25^2}} 50 + \frac{1}{\sqrt{.75^2 + .25^2}} 90 + \frac{1}{\sqrt{.75^2 + .75^2}} 125 + \frac{1}{\sqrt{.25^2 + .75^2}} 75}{\frac{1}{\sqrt{.25^2 + .25^2}} + \frac{1}{\sqrt{.75^2 + .25^2}} + \frac{1}{\sqrt{.75^2 + .75^2}} + \frac{1}{\sqrt{.25^2 + .75^2}}} = 74.27$$

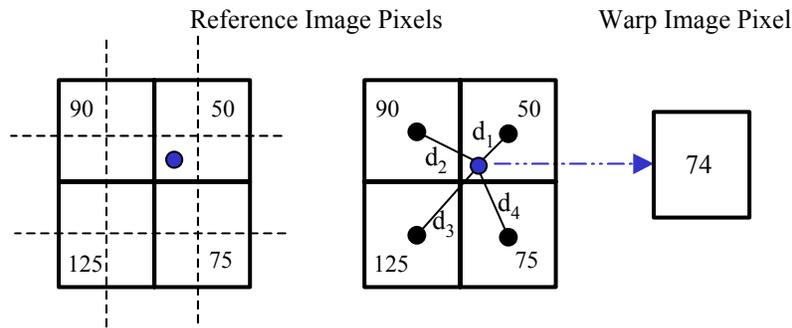


Figure 3.2.2: Example of grayscale resampling using Bilinear Interpolation.

The benefits of bilinear interpolation include moderate computational requirements and relatively good results, since the “stair-stepping” artifacts that occur in nearest neighbor resampling do not occur here. Unfortunately, radiometric integrity is not maintained since new grayscale values can be created through the interpolation (blurring) process, which can have very damaging results for spectral analysis. These new grayscale values could represent new object reflectances/radiances that might confuse classifying algorithms. Due to the pixel blurring results of any resampling besides nearest neighbor, Schott maintains that it is often desirable to perform radiometric calculations before any geometric transforms are applied to the data. (Schott 1997)

Smoother results can be obtained by using more sophisticated techniques, such as cubic convolution interpolation, which fits a surface of the $\sin(z)/z$ type through a much larger

number of neighbors (4, 8, 16) in order to obtain a smooth estimate of the gray level at any desired point (Gonzalez and Woods 2002). Cubic convolution resampling provides the closest approximation to the ideal, since it is designed to emulate the characteristics of the SYNC function, $\sin(z)/z$. Linear systems theory indicates that the ideal sampling kernel would leave all frequencies unaffected (i.e., a RECT function in frequency space) (Schott 1997), which results in a SYNC function in the space domain (figure 3.2.3). Although the SYNC function would require a convolution kernel over the whole image, a quantized version is normally utilized that affects only a 4x4 or 8x8 region around the target pixel for resampling. Even with this simplification computation time suffers compared to the previous methods since more neighboring pixels are sampled. Processing time aside, this technique will often provide the most “eye pleasing” results when resampling is required. The LoGWaR software utilizes the IDL implementation of the cubic convolution with a default value of -0.5 . Park and Schowengerdt (1983) suggest that this value can significantly improve the reconstruction properties of the algorithm that has been incorporated into IDL.

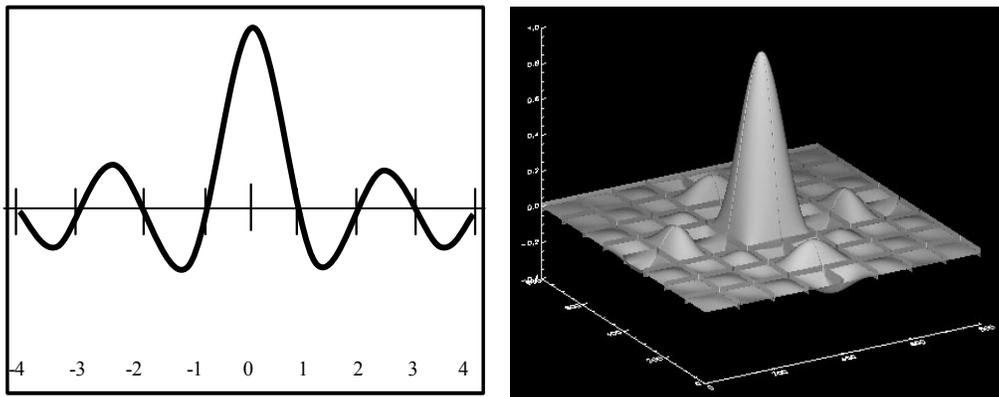


Figure 3.2.3: Cubic Convolution grayscale resampling over an 8x8 weighted pixel area.

The type of resampling we choose to implement will be heavily dependent on our application and desired outputs. If radiometric integrity is of primary importance, as it may be when registering and sharpening a multispectral image with a hyperspectral image, the nearest neighbor method can be utilized. However, if we wish to register a high-resolution panchromatic (or RGB) image with a lower-resolution spectral dataset, one of the alternate resampling techniques may be more attractive. Since radiometric integrity is less of an issue when transforming a panchromatic image, cubic convolution can be utilized unless processing time is excessive; in which case, bilinear interpolation would be the algorithm of choice.

3.3 Laplacian of Gaussian (LoG) Filter

The choice of filters to help identify and accentuate invariant features from within multisensor images is a critical design decision for this automated registration process. The idea for using the LoG filter for this task was sparked after noticing its effects on synthetic data during research on edge detection, for which this filter is traditionally utilized. During a Digital Image Processing exercise, it became apparent that the LoG filter could be utilized to consistently pinpoint features within an overhead image that might be utilized for image registration. By applying a threshold to the LoG filtered image, it is possible to isolate regions that have similar rates-of-variation within a scene and to do so in a repeatable fashion. This is due to the “second derivative” (∇^2) nature of the Laplacian filter which produces high output for well defined edges. Figure 3.3.1 demonstrates the effect of the LoG filter on a synthetic dataset that resembles the letter “X” but could represent a crossroads or building in an overhead image.

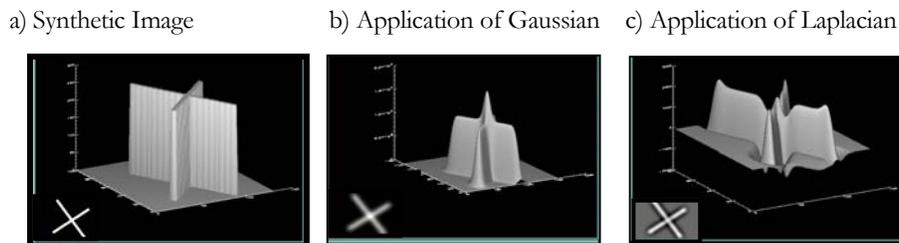


Figure 3.3.1: Demonstration of the effects of LoG filter to identify repeatable Maximums

The effect that the LoG filter has on an image is very similar to the lateral-brightness adaptation of the human eye (also known as lateral inhibition) that leads to the “Mach band effect”. Evidence of this is provided by Gonzalez and Woods, when they maintain that certain aspects of human vision can be modeled mathematically in the basic form of the LoG equation (Gonzalez & Woods 2002). This phenomenon is demonstrated in figure 3.3.2, with an exaggeration of grayscale step edges. In the human eye, this adaptation is due to cross-talk between different receptors [in the retina] and it makes the discrimination between objects more apparent by enhancing the detection of edges (Hailstone 2001). Schott maintains this is because, each cell reduces the sensitivity of adjacent cells when it is excited (Schott 1997).

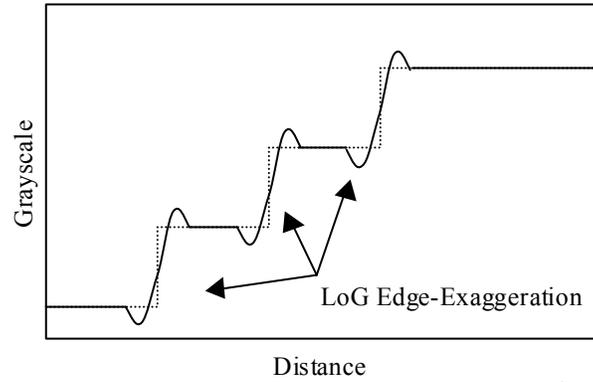


Figure 3.3.2: Edge-exaggeration resulting from convolution w/LoG Filter

The Laplacian is the second derivative of a function. This equation takes the following forms for both the 1-D (3.7) and 2-D (3.8) versions:

$$(3.7) \quad \nabla^2 f = \frac{\partial^2 f}{\partial x^2}$$

$$(3.8) \quad \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

This function can be approximated with the following 1-D & 2-D digital filters:

$$(3.9) \quad \nabla^2 = \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$$

$$(3.10) \quad \nabla^2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

A graphical representation of the effects of this filter when applied to a 1-D step function (3.4.a) that has been first convolved with a Gaussian low-pass filter (3.3.3.b) follows. It can be seen why the 2nd Derivative filters are also called “zero-crossing” edge detectors since the knife edge input (3.3.3.a) goes to unity precisely at the zero crossing between the positive and negative peaks of figure 3.3.3.d.

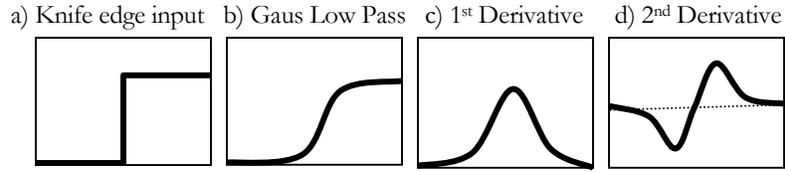


Figure 3.3.3: Visual effect of the Laplacian of Gaussian Filters in succession

Although the LoG filter can be easily deconstructed into its component parts as seen above, it is more commonly implemented in one convolution step with a kernel similar to figure 3.3.4. the 5x5 filter approximation and the “Mexican Hat” (LoG) function are shown below.

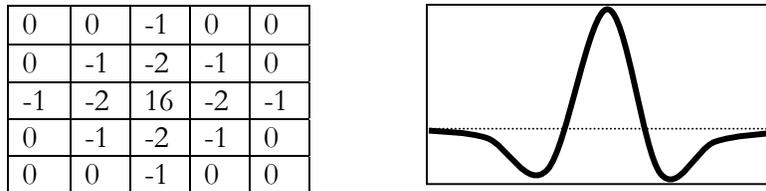


Figure 3.3.4: A composite 5x5 LoG Filter & 1-D representation of the function it approximates.

After successfully testing both techniques within the LoGWaR program, the incremental approach of applying first the Gaussian Filter and then the Laplacian has been primarily adopted due to the ease in which the width of the Gaussian smoothing kernel can be changed to mitigate the effects of noise within an image.

The Laplacian is very good at highlighting variation within an image. This result is useful if the variation is equivalent to information content or edges, but, detrimental if that variation is represented by noise. On its own, the Laplacian will accentuate all high frequency components, including noise along with the edges. For this reason the image is first convolved with a Gaussian filter, to diminish the effects of noise, before the Laplacian filter is applied. Gonzalez and Woods point out that the edges determined by the zero crossings form numerous closed loops within the filtered image (Gonzalez and Woods 2002). Although they suggest that this detracts from utilizing the LoG filter as an edge detector, it will actually work to our benefit in automated GCP selection since it will more easily allow for “connected components” analysis after thresholding of the resulting maxima and minima (figure 3.3.5).

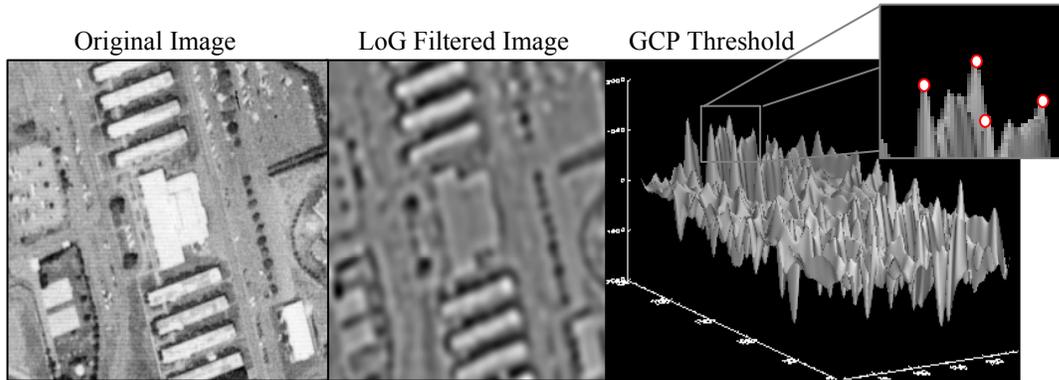


Figure 3.3.5: The results of the LoG filter and thresholding of maxima to create GCPs.

The results of this LoG thresholding process provide the automated GCP selection within each image. Once these GCPs have been identified, a point matching routine (section 4.4) will be utilized to relate the subset of similar points from each image. As mentioned in section 3.1.2, these related points can be used to develop a polynomial equation, for registration of the two images.

3.4 Wavelet Analysis

The ability of wavelet algorithms to gracefully decimate images by iteratively stripping off high frequency components, saving that information into orthogonal wavelet coefficients and a reduced “scale” image proves quite useful in multisensor image registration. The focus of this paper’s wavelet analysis will be on multiresolution or “resolution pyramid” schemes. Multiresolution theory gives a simple and fast method for decomposing a signal into its components at different scales. It is possible to progressively drain the signal of its information, beginning with small details and continuing on to larger and larger components. At each step the "details" are encoded as wavelet coefficients, while the next step analyzes the signal seen at half the previous resolution (Hubbard 1996).

Even as Fourier analysis deconstructs a periodic signal into a sum of orthogonal sines and cosines, so to wavelet analysis deconstructs a signal into its constituent wavelet coefficients. Wavelet coefficients are largest where they best match the signal being analyzed.

The main advantage to utilizing wavelet theory is that it allows concurrent analysis of scale and frequency whereas Fourier analysis is limited only to frequency.

$$(3.11) \quad f(x) = \int_{-\infty}^{\infty} F(\xi)e^{2\pi i\xi x} d\xi$$

Where $F(\xi)$ is the scaling function and $e^{2\pi i\xi x}$ is the basis function.

$$(3.12) \quad e^{i\theta} = \cos \theta + i \sin \theta$$

So, in the case of the discrete fourier series below, the a's and b's are the coefficients, which tell us how much of the integer multiples of cosines and sines are contained in the function $f(x)$.

$$(3.13) \quad f(x) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} (a_k \cos 2\pi kt + b_k \sin 2\pi kt)$$

For the continuous wavelet transform (CWT), we have:

$$(3.14) \quad W_x(s, \tau) \equiv \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} f(x)\psi\left(\frac{x+\tau}{s}\right)dx \quad (\text{Daubechies 1992})$$

Where s provides the scaling of the function W_x , τ shifts it and $\psi(x)$ is the “mother wavelet” that is utilized to deconstruct the function much as the sines and cosines of the Fourier series. The primary difference is that the wavelet must have relatively compact support, but like the sinusoidal Fourier series components, must integrate to zero. As is the case with the discrete Fourier series, the mother wavelet is translated and scaled only by integer values (s and τ are integers) for the discrete wavelet transform. Also, use of the Fast Wavelet Transform (FWT) requires scaling by a factor of two ($s=2^k$); this is sometimes referred to as dyadic constraint.

The use of wavelet analysis in this paper will be limited to the orthogonal Haar Wavelet (Daubechies order 1), which can be seen in figure 3.4.1. Both the ease with which it can be

applied in the form of the FWT and its application in multi-resolution pyramid analysis will aid greatly in its application and utility here.

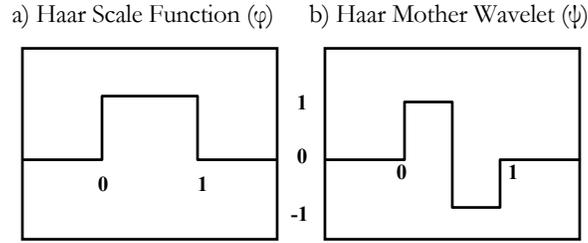


Figure 3.4.1: The Haar Wavelet used here for the Fast Wavelet Transform.

In addition, the FWT demonstrates both lossless decimation and maintains radiometric integrity (see Appendix A). Before addressing the FWT, it is useful to first investigate the mathematical formulation of the DWT. The following derivation for the Haar DWT was drawn from (Rao and Bopardikar 1998).

For the Haar Scale Function (φ), the “father wavelet”, let:

$$(3.15) \quad \varphi(t) \equiv \begin{cases} 1 & 0 \leq t < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$(3.16) \quad c(k, l) \equiv \frac{1}{2^k} \int_{2^k(l)}^{2^k(l+1)} f(t) dt$$

The equation for $c(k, l)$ simply represents the average value over the interval, so:

$$(3.17) \quad \varphi(2^{-k}t) \equiv \begin{cases} 1 & 0 \leq t < 2^k \\ 0 & \text{otherwise} \end{cases}$$

So, a function $f_k(t)$ can be approximated by:

$$(3.18) \quad f_k(t) = \sum_{l=-\infty}^{\infty} c(k, l) \varphi(2^{-k}t - l)$$

Now, if the Haar mother wavelet is defined as:

$$(3.20) \quad \Psi(t) = \begin{cases} 1 & 0 \leq t < \frac{1}{2} \\ -1 & \frac{1}{2} \leq t < 1 \\ 0 & \text{otherwise} \end{cases}$$

then, the detail and scaling wavelets can be expressed (in interval 0 to 1) as:

$$(3.21) \quad \Psi_{[0,1)}(t) = \varphi_{\left[0, \frac{1}{2}\right)}(t) - \varphi_{\left[\frac{1}{2}, 1\right)}(t)$$

$$(3.22) \quad \varphi_{[0,1)}(t) = \varphi_{\left[0, \frac{1}{2}\right)}(t) + \varphi_{\left[\frac{1}{2}, 1\right)}(t)$$

$$(3.23) \quad \varphi_{\left[0, \frac{1}{2}\right)}(t) = \frac{\left[\varphi_{[0,1)}(t) + \varphi_{[0,1)}(t) \right]}{2}$$

$$(3.24) \quad \varphi_{\left[\frac{1}{2}, 1\right)}(t) = \frac{\left[\varphi_{[0,1)}(t) - \varphi_{[0,1)}(t) \right]}{2}$$

and if detail in the wavelet subband is expressed as:

$$(3.25) \quad d(k, l) \equiv \frac{1}{2} [c(k-1, 2l) - c(k-1, 2l+1)]$$

The detail function is then given by:

$$(3.26) \quad g_k(t) = \sum_{l=-\infty}^{\infty} d(k, l) \Psi(2^{-k}t - l)$$

Appendix 1 contains a simple 1-D example of the Haar FWT and how it can be utilized to deconstruct a signal into its frequency and space components for multiresolution analysis.

The fast wavelet transform works from fine resolution to coarse. At each resolution, the signal is analyzed with both wavelets and scaling function. The wavelets encode the details, while the scaling function produces an image of the signal at half resolution, taking one sample out of

two. The process is repeated until nothing, or virtually nothing, is left (Hubbard 1996). In this way, an image is deconstructed iteratively into half resolution components.

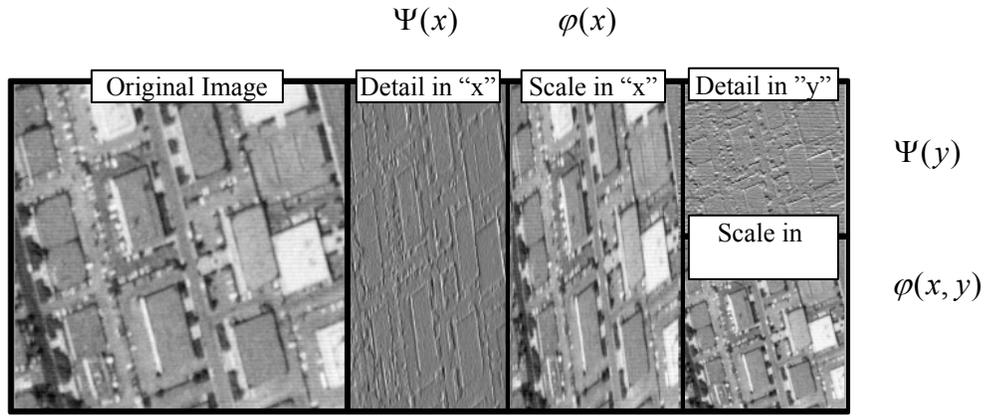


Figure 3.4.2: A minimal processing approach to the FWT (one less step than traditional).

Figure 3.4.2 demonstrates the “separability” of the FWT algorithm in both the “x” and “y” axis when utilizing an approach that minimizes the number of operations needed to maintain an orthogonal decimation. This figure shows how the FWT first strips off the highest frequency data ($\Psi(x)$) from the scale image ($\phi(x)$) in the “x” and then strips off the remaining highest frequency data ($\Psi(y)$) from the scale image ($\phi(y)$) in the “y”. The resulting scale plane ($\phi(x, y)$) and detail planes ($\Psi(x)$ & $\Psi(y)$) contain all the information of the original image. This process is continued by utilizing the scale plane ($\phi(x, y)$) of this process as the input image to the next FWT reduction. The reverse application of the process yields the inverse Fast Wavelet Transform (FWT^{-1}) and orthogonality is proven, by generating the original image. A detailed analysis of how the FWT algorithm accomplishes this feat is incorporated in Appendix A.

The ability to decimate a high resolution panchromatic image into increasingly lower resolution “scale images” allows for direct comparison to lower resolution spectral images, thus increasing the potential to automatically register these images. An added benefit is that once registered in this manner, it is possible to transfer the high frequency wavelet coefficients from the panchromatic dataset to the related spectral image planes for the purpose of spatial “sharpening”. The same applies to sharpening of hyperspectral bands with multispectral, but

with the added potential for spectral interpolation/extrapolation (commonly referred to as “crossband correlation”). The more traditional Mallat Representation, displays the subbands for individual analysis of horizontal, vertical & diagonal detail as well as the reduced resolution scale subband (figure 3.4.3).

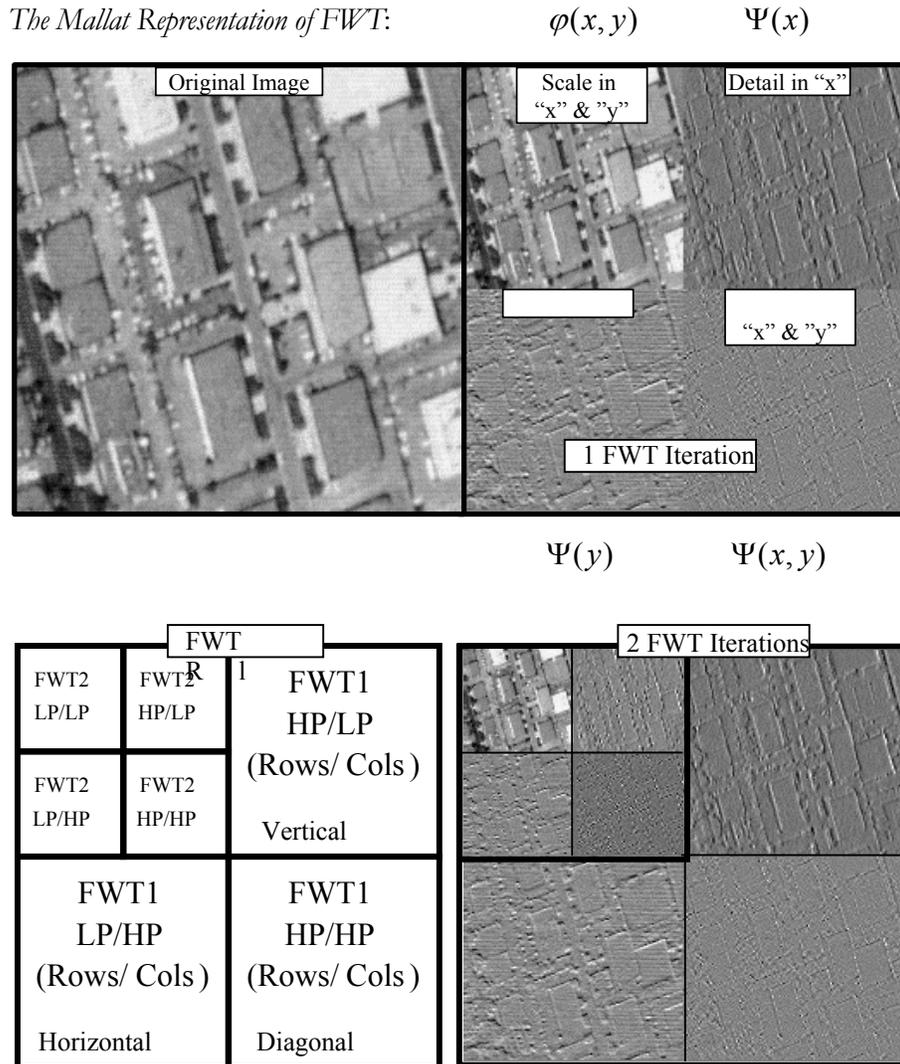


Figure 3.4.3: The Haar Wavelet orthogonally extracts high frequency “details” from x & y .

The biggest drawback in utilizing the FWT is that image registration is constrained to dyadic dimensions in both the x and y . Hubbard maintains that with an orthogonal wavelet transform, a signal is analyzed at scales varying always by a factor of two (obeying, without excess, the Shannon sampling theorem, since each time the frequency doubles, one doubles the number of wavelets used to sample the signal) (Hubbard 1996). Maintaining orthogonality

is important here, since it is imperative that all high frequency information can be completely recovered in order to maintain spectral integrity. Although utilizing the FWT necessitates a dyadic constraint on the dimensions of the common image chips that can be registered with each other, savings in processing time and complexity greatly outweigh the limitations incurred.

THE REGISTRATION PROCESS

The proposed process for registering multisensor images is most clearly explained through the use of a flowchart. Figure 4.1.1 gives an overview of the essential steps and the process flow that is required for multisensor image registration, as developed in this research. This process relies on the theory developed in Chapter 3 and represents the critical elements used to register images with the software developed as part of this thesis. The Laplacian of Gaussian Wavelet Registration (LoGWaR) software was crafted as a graphical user interface, in IDL, to help automate the process of image registration. Although the steps in the registration process have been laid out in a linear fashion, many of the initial steps (in 4.1 & 4.2), could occur in an order other than what is presented.

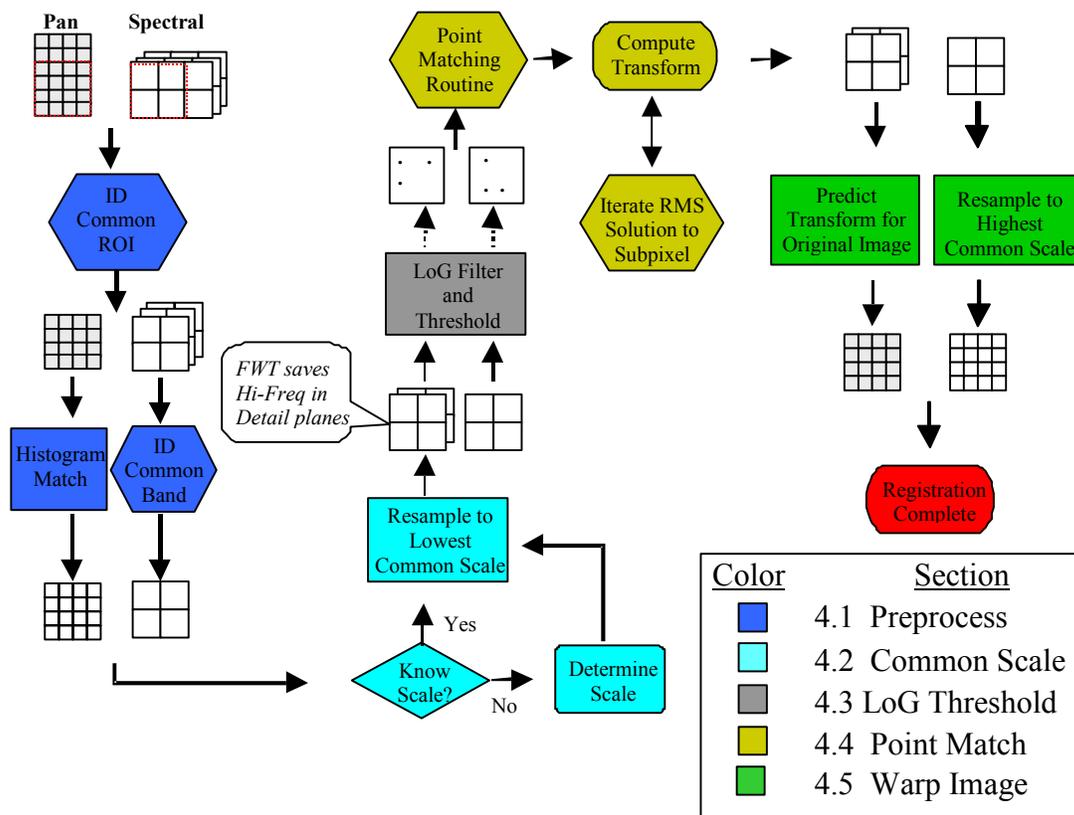


Figure 4.1.1: The Multisensor Image Registration Process.

Whenever possible throughout this process, changes to grayscale values (digital count) will be constrained to the grayscale images (Panchromatic) in order to maintain the integrity of our spectral image radiometry. When this cannot be done, special precautions may be instituted to reduce contamination of the radiometric information such as utilizing nearest neighbor resampling (section 4.5). In the case of using multispectral data for the sharpening of hyperspectral cubes, resampling will be kept to a minimum and if possible, accomplished in one composite step as opposed to multiple single steps.

4.1 Image Preparation – Finding “Common Ground”

Although one of the primary thrusts of this research is to automate the registration process as much as possible, one undeniable fact remains...the higher the degree of correlation between the two datasets (in dimension, scale, orientation, and grayscale range, etc.), the easier it is to get a good automated registration. The commonality, of the datasets, directly impacts the ability to automatically identify and solve for the degrees-of-freedom that exist between the data to be registered. So, any preparation that can easily be accomplished to relate the dataset in advance of automated techniques, is often worthwhile. This step is often referred to as image pre-processing or preparation.

Some steps, such as histogram matching and LoG filtering, are utilized only to help relate the two datasets. The common theme here is that it is totally acceptable to change the digital count values to identify common features in the images, but, the spatial coordinate relationship must be carefully maintained/managed to ensure that the final warped output is spatially registered. Once the images have been registered and the transform has been determined, it will be applied to the original image to minimize unnecessary changes to the data.

Define Common Regions of Interest (ROI): One of the first steps, identified in the registration process (figure 4.1.1), is to identify regions within each image that are common. However, this step may not be necessary if the amount of overlap between the reference and warp image is large. Image subregions can be utilized to compare related regions of interest using supervised registration (user defined ROIs). But, for automated registration, the two images must have a general orientation relationship, before subregion analysis can commence.

If they do not, it is necessary to analyze the entire image at its native resolution or at a reduced scale, using wavelet decimation or resampling. This concept is explored in more detail in section 5.1.

User defined ROIs often make it easier for the automated algorithms to arrive at an acceptable registration since the degrees-of-freedom have been reduced by identifying the areas of commonality visually. This is especially necessary when mosaicing (stitching) two datasets, when the area of overlap is small (figure 4.1.2). The reason for this is that the point matching algorithms work best when the point sets are most similar.

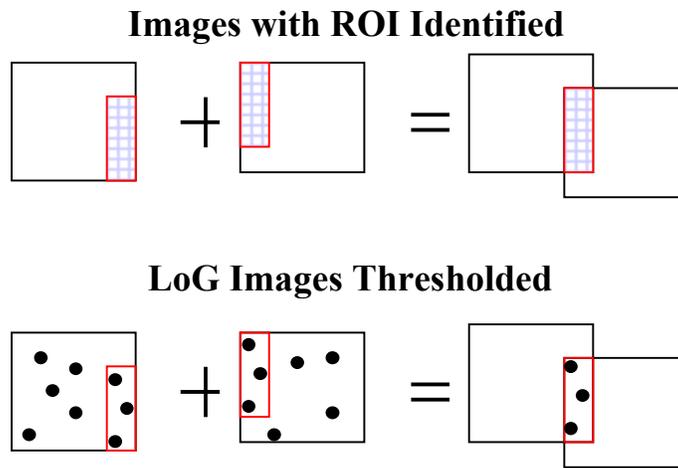


Figure 4.1.2: Image Mosaics may require user identification of commonality due to small overlap.

Similar Spectral Band/Region Selection: Another essential preparatory step is to extract the spectral bands with the highest degree of correlation. This may require some inspection since spectral bands can vary greatly from sensor to sensor. A good example is in the registration of a PAN image, which has a bandpass containing VIS and NIR spectra, with an MS image that normally has this spectra isolated into individual spectral bands. In this case it is often beneficial to derive a synthetic Pan image from the individual RGB and NIR bands.

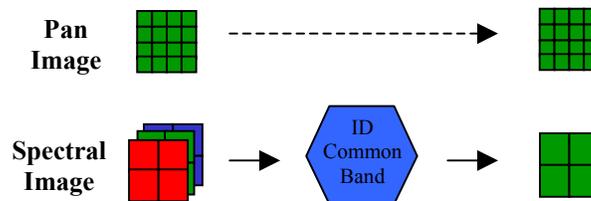


Figure 4.1.3: Select Spectral Regions that match most closely for Registration.

Histogram Matching the Images: Once this is accomplished, it is often useful to perform a histogram matching of the ‘warp’ image to the ‘reference’ image using a direct band-to-band comparison. This is done to ensure that the grayscale distributions are as similar as possible to promote automatic registration (figure 4.1.4). It is important to remember that the LoG filter is, in essence, an edge detector. The quest for commonality in grayscale and spectra is just a precursor to obtaining similar edge information in the two datasets! Once the images have been registered, this interim processing can be disregarded, since the registration transform that has been developed can be applied to the original image. It is always important to try and maintain radiometric integrity whenever possible!

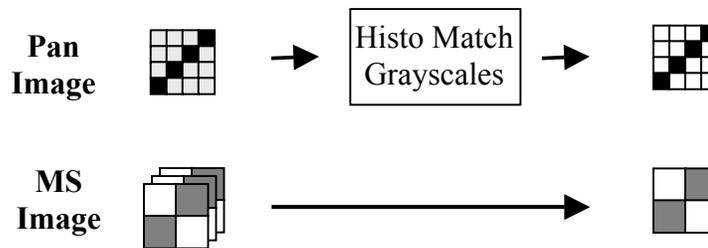


Figure 4.1.4: Histogram match images to enhance grayscale similarity.

Scale Similarity with Wavelet Decimation or Resampling: A final and critical requirement in the pre-processing stage is to obtain the same resolution for both datasets. This is important for two reasons. First, the point-matching algorithms exploit distance between like features to relate the datasets. Secondly, the results of the LoG filtering operation change depending on the edge/scale relationship and provided the initial incentive to explore multiresolution pyramids and wavelet domain analysis. Since, wavelet analysis allows for multi-scale analysis, it seemed a good choice for addressing this problem. In fact, "a multiresolution decomposition enables us to have a scale-invariant interpretation of the image" (Mallat 1989) and can be very useful in image registration.

Although the point matching algorithms, discussed in section 4.3, have been built to automate the determination of scale differences between the reference and warp images, we can often assume that this knowledge is available in the header information of the remotely sensed data or is already known. This assumption is often acceptable since the sensor imaging resolution is often well known and is one of the primary characteristics of most remote sensing

systems. The simplified results of the FWT as utilized in our registration process are shown in figure 4.1.5 to attain images at similar scale.

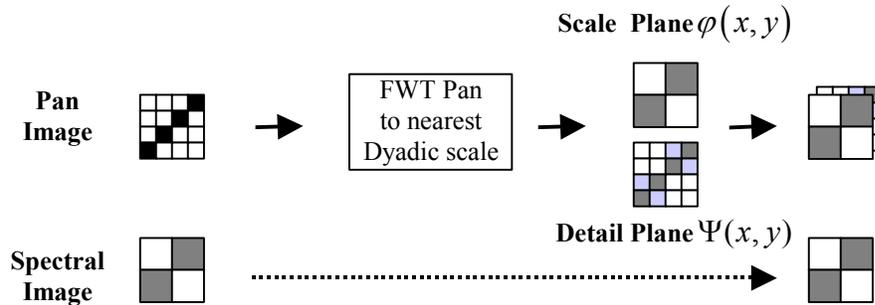


Figure 4.1.5: FWT image decomposition into orthogonal scale and detail planes.

When it is not necessary to retain the high-frequency information contained in the detail planes of the wavelet decimated image, it is then possible to perform a simple resampling of the image to the lowest common resolution. Once this is accomplished, the images are ready for registration utilizing the automated techniques encompassed in LoGWaR. Although the images will be registered at the lower resolution, it is possible to retain the original image and scale relationship for warping at a later time. This is possible through a predictive transformation process outlined in section 4.8 and Chapter 5.

It should be noted that degrading the high-resolution image to the same dimensions as the lower resolution data decreases the precision of the registration. However subpixel accuracy registration is still possible if enough GCPs are generated to over-define the least-squares solution provided by the psuedo-inverse process. So, it is possible to register at the lowest common resolution (common frequency information) and then predict the transform needed for the original high-resolution image.

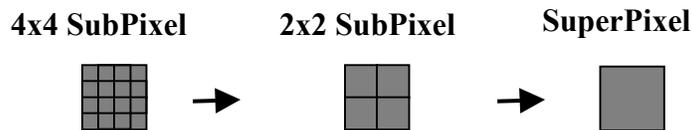


Figure 4.1.6: Resampling a pixel array into a superpixel with the same overall intensity.

“Resampling” is a process that replicates an existing pixel into several smaller subpixels that reside in the same space and have the same intensity as the original pixel or that averages

several smaller subpixels into a superpixel. So, as seen in figure 4.1.6, one pixel could be “subsamped” into matrices of smaller subpixels such as: 2x2, 3x3, 4x4, ...10x10, as long as all pixels in the matrix retain the same intensity (grayscale) level as the original pixel or vice versa. When this is applied to an entire image plane, overall radiometry can be preserved if each subpixel matrix retains the same intensity value as its parent super-pixel.

So the motto for registration is this: “Change whatever you want on the interim images to attain as much commonality as possible between the datasets to be registered. This may including adjusting scale, histogram matching grayscale values, dynamic range adjustment, and even band integration (it is often useful to create synthetic panchromatic image from RGB when registered RGB to Pan). But, keep tract of all spatial changes so that they can be incorporated into the composite model and perform the final transform on the original image or dataset”.

4.2 LoG Thresholding

The multiscale wavelet decimation/resampling, of section 4.2, allows the LoG filter to process images with similar spatial frequency content for the registration process. Since the LoG filter is essentially a robust edge detection tool, the detail information is much different in a low-resolution spectral cube in comparison to a high-resolution panchromatic image. In fact, we will be disregarding all of the highest frequency information contained in the high-resolution image during the registration process with the lower resolution image (figure 4.2.1).

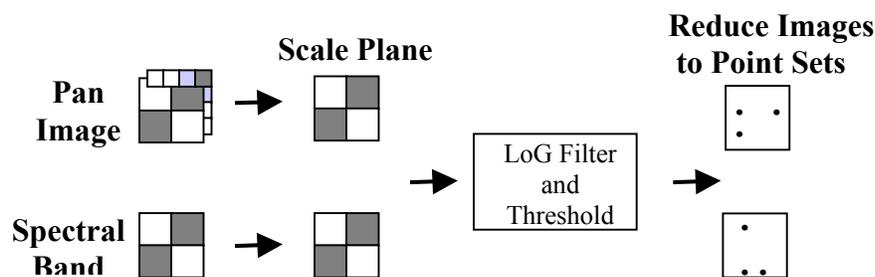


Figure 4.2.1: Reduction of the images to point sources (using the LoG filter) for matching.

The objective of this step is to reduce both images into sets of points. These points are the thresholded maxima and minima of the LoG filter output. They represent areas within the

image that portray extreme rates-of-variation and should be relatively similar for like images that have been taken under similar conditions. Once the LoG output has been thresholded, the resulting regions are aggregated using ‘connected components’ analysis and a distinct pixel is isolated as the extrema for each region. These point sets can then be related using point matching schemes and utilized, as ground control points (GCPs), to register the two images.

4.3 Point Matching Schemes

The accuracy of registering images utilizing the LoG technique boils down to how well related areas of both images can be identified, isolated, and matched. Even though the LoG thresholding procedure simplifies the registration process by reducing the images to point sets. It is the accurate matching of points, from dissimilar point sets, that will determine the utility and ultimate success of this registration process!

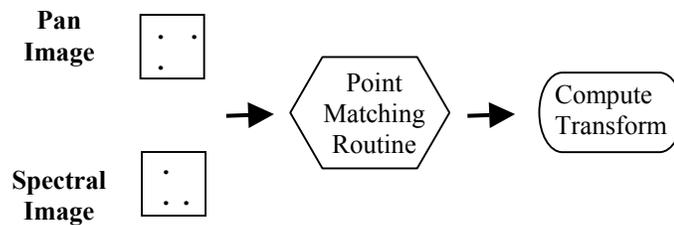


Figure 4.3.1: Matching points to determine the Polynomial Transform.

Throughout this section, a robust point matching technique will be introduced and applied to the task of image registration. An important concept to keep in mind is that the matched points will provide the matrix equation inputs to solve for the geometric distortions. So, if our panchromatic image is shifted (both horizontally and vertically), rotated, and scaled then we require three sets of matched points to solve for the geometric transform. If we have more matched points than required, the solution is over-determined and it is possible to either select a subset of the “best” point matches that uniquely determine the solution or utilize a linear regression model to estimate the best fit to the data and obtain a subpixel registration.

Point Distance Comparison: The method utilized here to match shifted and rotated points is adapted from the realm of astronomy, where it was used to mosaic “star fields” (Chandrasekhar 1999). The “star field” datasets are not dissimilar to the outputs of the

thresholded, LoG filtered images. This led the author to believe the technique could be cross-pollinated into the area of remote sensing image registration. This process utilizes a point's distance from every other point in a scene and creates an array of distances with this data. This is done with each point in the image, from which a matrix of distances is created. The point distance matrices, from each image, are then compared row-to-row for total number of matching distances. The two rows that have the greatest number of distance matches (within some predestinated error) are considered matched points (figure 4.3.2). In real image registration cases, it is quite possible to have a one-pixel variability across the focal plane due to quantization variability. So the default error for matching has been set to 2 pixels, to allow for the diagonal case of $\sqrt{2}$ pixel distance. This value can be changed by a slider bar in the lower left corner of the LoGWaR GUI, reference Appendix C.

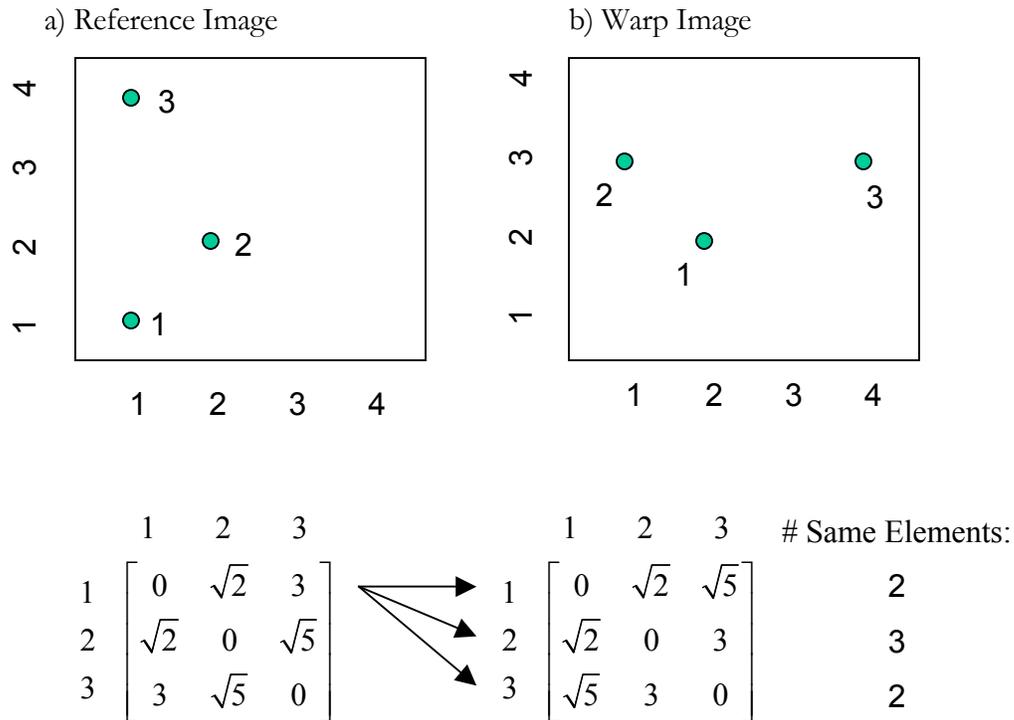


Figure 4.3.2: Determining matching points through equivalent distances to other points.

The distance between any two points is equal to the square root of the sum of the squares: Distance, $d = \sqrt{[(x_1 - x_2)^2 + (y_1 - y_2)^2]}$. For our reference image points 2 and

3, this becomes: $d = \sqrt{[(2-1)^2 + (2-4)^2]} = \sqrt{5}$. In the matrix each row and column (i.e. column & row 1) represents that point's (point 1) distance from the other points, which are also related to their equivalent row and column. In our example above, point 1 from the reference image would match point 2 from the warp image since they have the greatest number of matching distances in their equivalent rows.

Point Scale Comparison: This point distance matrix technique works well for images that can be related via shift and rotation transforms. However, the technique that will be developed here, to address scale difference, depends on the ratio of distances between sets of points, regardless of relative size. The implementation of this concept is similar to the technique above and is illustrated below (figure 4.3.3).

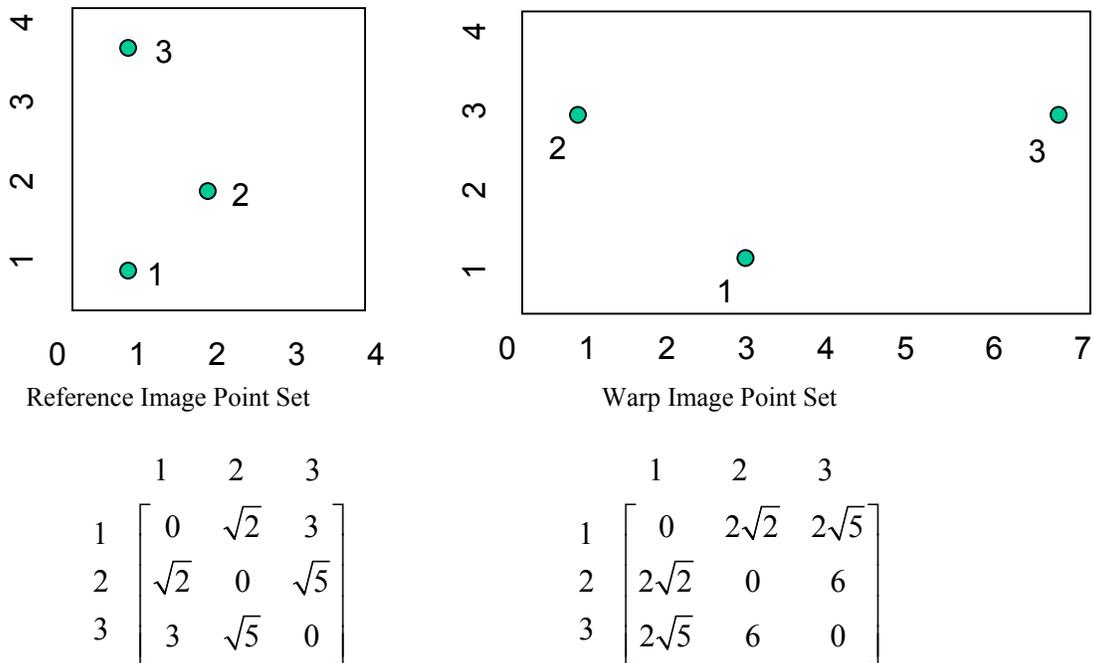


Figure 4.3.3: Determining Scale through equivalent distance ratios to other points.

For scale determination, it is simply a matter of comparing the ratio of point distances, from the reference image, to the ratio of matching point distances from the warp image:

$$(4.1) \text{ Dist Ratios: } \frac{dist_{ref(1 to 2)}}{dist_{ref(1 to 3)}} = \left(\frac{\sqrt{2}}{3} \right) \cong \frac{dist_{warp(2 to 1)}}{dist_{warp(2 to 3)}} = \left(\frac{2\sqrt{2}}{6} \right)$$

Point Angle Comparison: Angle matching can be accomplished, along with distance, as an additional discriminator for point matching. With this technique, it is necessary to compare vertices. So, 3 points will be used to define each angle of interest, and can be computed through the use of the following formula:

$$(4.2) \quad \theta[i, j, k] = \arccos \left(\frac{a^2 - b^2 - c^2}{-2 * b * c} \right)$$

Due to the 3 point per angle requirement, it is necessary to perform a 3-D matrix comparison of angles to determine point matches. This process is outlined in figure 4.3.4 and is similar to the distance matrix approach, except that now each point has a plane of angles associated with it instead of just a row and column of distances.

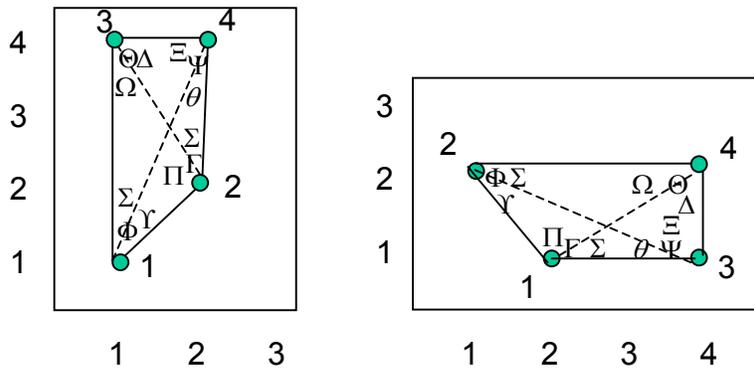


Figure 4.3.4.a: Identifying Angles for potential point matching.

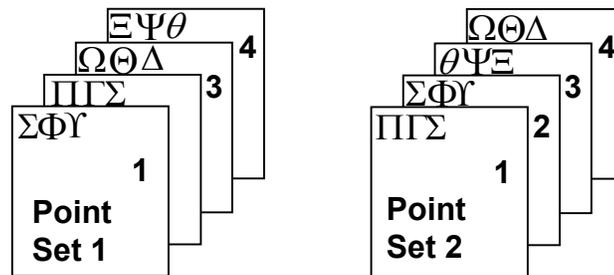


Figure 4.3.4.b: Each Point is isolated into planes containing angles for every vertex.

Point Set 1	Point Set 2	# of Matching Angles
<p>①</p> <p>1 2 3 4</p> $1 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \Phi & \Upsilon \\ 0 & \Phi & 0 & \Sigma \\ 0 & \Upsilon & \Sigma & 0 \end{bmatrix}$	<p>①</p> <p>1 2 3 4</p> $1 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & \Gamma & \Pi \\ 0 & \Gamma & 0 & \Sigma \\ 0 & \Pi & \Sigma & 0 \end{bmatrix}$	2
<p>②</p> <p>1 2 3 4</p> $1 \begin{bmatrix} 0 & 0 & \Pi & \Gamma \\ 0 & 0 & 0 & 0 \\ \Pi & 0 & 0 & \Sigma \\ \Gamma & 0 & \Sigma & 0 \end{bmatrix}$	<p>②</p> <p>1 2 3 4</p> $1 \begin{bmatrix} 0 & 0 & \Upsilon & \Phi \\ 0 & 0 & 0 & 0 \\ \Upsilon & 0 & 0 & \Sigma \\ \Phi & 0 & \Sigma & 0 \end{bmatrix}$	6
<p>③</p> <p>1 2 3 4</p> $1 \begin{bmatrix} 0 & \Omega & 0 & \Theta \\ \Omega & 0 & 0 & \Delta \\ 0 & 0 & 0 & 0 \\ \Theta & \Delta & 0 & 0 \end{bmatrix}$	<p>③</p> <p>1 2 3 4</p> $1 \begin{bmatrix} 0 & \theta & 0 & \Psi \\ \theta & 0 & 0 & \Xi \\ 0 & 0 & 0 & 0 \\ \Psi & \Xi & 0 & 0 \end{bmatrix}$	0
<p>④</p> <p>1 2 3 4</p> $1 \begin{bmatrix} 0 & \theta & \Xi & 0 \\ \theta & 0 & \Psi & 0 \\ \Xi & \Psi & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	<p>④</p> <p>1 2 3 4</p> $1 \begin{bmatrix} 0 & \Omega & \Delta & 0 \\ \Omega & 0 & \Theta & 0 \\ \Delta & \Theta & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	0

Figure 4.3.4.c: Discriminating Point Matches through Angle match count.

Although this method is useful as an additional criteria for determining matching points, the additional requirements of the 3-Dimensional analysis negatively impacts processing time. However, since this method is used only after the distance comparison technique has already identified potential matches, the point sets are often much smaller and take considerably less processing time. In fact, the angle matching criteria is also implemented after the LoG Maxima comparison that follows. This further prunes the point sets before the angle match comparison is utilized. The default error that is allowed for matching angles is set at $\frac{1}{4}$ degree, but can be changed via a slider bar in the LoGWaR GUI.

LoG Maxima Comparison: The LoG Maxima comparison is based on the concept that the LoG threshold extrema should have similar values before a match is allowed. A graphical example of this method is shown in figure 4.3.5.

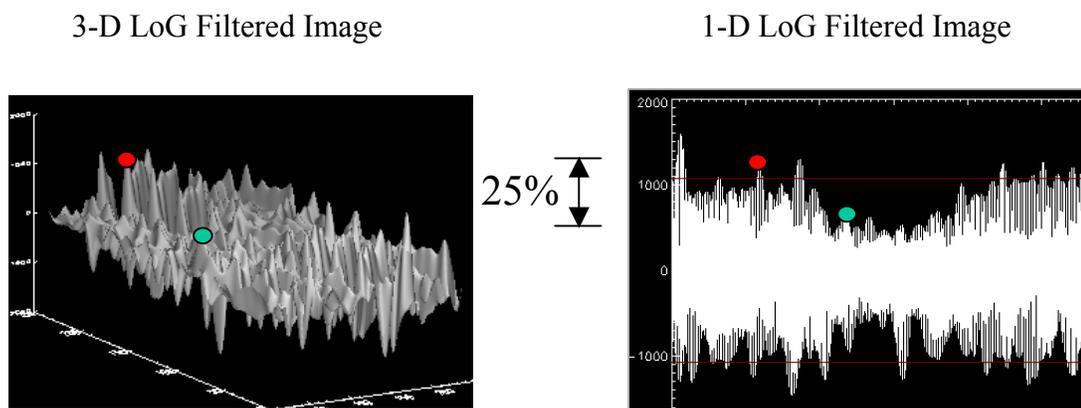


Figure 4.3.5: Discriminating Point Matches through LoG Maxima Comparison (+/- 25%).

Since the LoG values are normalized before comparison, it would be normal for related points to have a similarity within 10-25%, if not better. An additional benefit of the LoG filter is the additional discrimination of points based on positive and negative extrema values.

Match Distance Comparison: The first statistical, post-match based criteria that is sometimes used, is limited to images with little to no rotational variation. This technique is based on the concept that matches should have similar local translation offsets when

compared to a localized image transformation model. If the offsets vary, by greater than 1 standard deviation from the mean, they are candidates as outliers for ejection from the point match set. This concept is portrayed in figure 4.3.6.

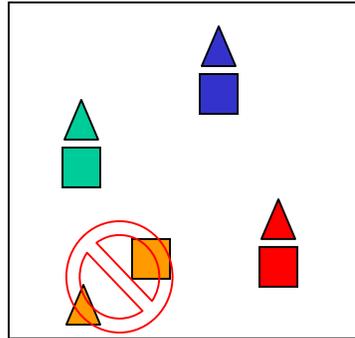


Figure 4.3.6: Post Match Comparison based on statistical analysis of match offset

In the course of a standard registration, with all of the matching criteria enabled (via GUI pushbuttons), the *Distance Matching* routing will provide a quick analysis, resulting in preliminary matches. This set of potential GCPs will then be quickly scrutinized by the *LoG Maxima* test. By the time the potential GCPs reach the *Angle Matching* routing, the subset is much smaller and can be more efficiently managed by the required 3-Dimensional analysis. Finally, most anomalous matches can be identified through localized statistical analysis of *Matching GCP Distances*, if the images are not rotated with respect to each other. Additional, post-match, global statistical analysis of RMS Distance Error of Matches will be covered in section 4.7. This is done in conjunction with determination of the registration accuracy and includes analysis of variation from polynomial transform models.

4.4 Developing the Transform Model:

Once enough points from the reference and warp images have been matched to uniquely (at minimum) relate them, it is simply a matter of solving n-equations for n-unknowns using matrix inversion or psuedo-inversion to develop the polynomial transform. “To see how GCPs are used to find the polynomial coefficients, suppose we

have located M pairs of GCPs in the distorted image and reference (image or map) coordinate systems. Assuming the global polynomial distortion model, we can write, for *each* pair m of GCPs, a polynomial equation of degree N in each variable, x and y ,

$$(4.3) \quad x_m = a_{00} + a_{10}x_{refm} + a_{01}y_{refm} + a_{11}x_{refm}y_{refm} + a_{20}x_{refm}^2 + a_{02}y_{refm}^2$$

$$(4.4) \quad y_m = b_{00} + b_{10}x_{refm} + b_{01}y_{refm} + b_{11}x_{refm}y_{refm} + b_{20}x_{refm}^2 + b_{02}y_{refm}^2$$

leading to M pairs of equations. This set of equations can be written in vector-matrix form for the x coordinates of the image as”, (Schowengerdt 1997)

$$(4.5) \quad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} = \begin{bmatrix} 1 & x_{ref1} & y_{ref1} & x_{ref1}y_{ref1} & x_{ref1}^2 & y_{ref1}^2 \\ 1 & x_{ref2} & y_{ref2} & x_{ref2}y_{ref2} & x_{ref2}^2 & y_{ref2}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{refM} & y_{refM} & x_{refM}y_{refM} & x_{refM}^2 & y_{refM}^2 \end{bmatrix} \begin{bmatrix} a_{00} \\ a_{10} \\ a_{01} \\ a_{11} \\ a_{20} \\ a_{02} \end{bmatrix}$$

or,

$$(4.6) \quad X = WA$$

$$(4.7) \quad Y = WB$$

and utilizing the matrix inverse for uniquely-determined solutions,

$$(4.8) \quad A = W^{-1}X$$

$$(4.9) \quad B = W^{-1}Y$$

or utilizing the psuedo-inverse for over-determined solutions,

$$(4.10) \quad A' = (W^T W)^{-1} W^T X$$

$$(4.11) \quad B' = (W^T W)^{-1} W^T Y$$

4.5 Image Warping/Resampling

Implementing the polynomial transform from the preceding section is simply a matter of determining the new x_{warp} and y_{warp} geometric location within the warp plane and then transferring the sampled reference image grayscale value into this new image space. This is often done pixel-by-pixel from the transformed image plane, by sampling the grayscale value (nearest neighbor/bicubic) from the original image (figure 4.5.1). For example, warp image pixel (1,1) would be processed through the inverse polynomial transform to determine the location in the original image to sample the grayscale value from. Although this may seem backward compared to the intuitive method of transferring pixel information from the original image to the warped image, it avoids pixel “dropout” in the output.

Original Image Location (0.75,0.81) Transformed Image Pixel (1,1)

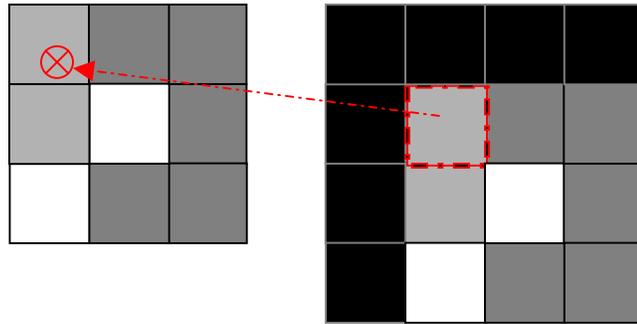


Figure 4.5.1: Sample grayscale value of nearest original image pixel using inverse transform.

As mentioned in section 4.1, it may be necessary to institute special precautions when warping multispectral data. It may be better to utilize “nearest neighbor” resampling, instead of bilinear or bicubic, to maintain radiometric integrity of the data. However, with a Pan image, the use of bicubic convolution will provide smoother resampling since we do not have to worry about maintaining the true spectral nature of the data.

4.6 Automation Techniques for isolating GCPs

Now the to the heart of the problem, how can the LoG thresholding process, that generates GCPs, be automated, to take full advantage of the automatic point matching techniques? When utilizing the LoGWaR program, it is easy for the user to determine an appropriate threshold level through visual inspection of the isolated LoG maxima. When the 'reference' and 'warp' maxima appear to have similar content and are of sufficient number and distribution throughout the image; it is simply a matter of initiating the point matching algorithms to relate the images. Three automated techniques have been developed and incorporated into the LoGWaR program to complement the manual threshold process outlined above.

Image Wide, Preset Threshold Level, Preset # of GCPs: The fastest automation technique involves thresholding the LoG filtered image at a preset, user-defined level (i.e. 50%). Once accomplished, the resulting extrema pixels are limited to a user-defined number of GCPs (i.e. 50 pts). This number does not necessarily equate to the number of final matches, but instead limits the maximum size of the point set, that will be used for comparison by the point matching algorithms. This number should be large enough to ensure an adequate number of matching points are produced from the point matching routines, but, not so large as to overly tax processing capability. Although this is heavily dependent on processing capabilities of the computer, 50 to 75 points seem to work well per ROI (the current limit is around 100 points per region).

If the threshold is set too low, many additional GCPs are selected than are required. This problem is easily solved by first sorting the resulting extrema pixels based on their normalized LoG values. Once this is accomplished, the point set is truncated to the highest LoG values based on the user-defined GCP limit. In this way, it is possible to extract and compare similar locations in both images since the highest LoG extrema represent regions with the greatest rates-of-variation in both images.

The strength of this technique is also its greatest weakness. The preset threshold level allows the process to quickly identify GCPs, unfortunately, the threshold level could be too high to adequately extract sufficient GCPs to relate both images. This inherent weakness of the preset threshold is rectified in the following technique.

Image Wide, Adaptive Threshold, Preset # of GCPs: The only difference between this technique and the previous one is in the adaptive thresholding of the LoG extrema. Instead of thresholding the LoG filtered image at a predetermined, user defined level, the threshold level is determined completely by the number of initial points requested.

To accomplish this task robustly, it is necessary to monotonically decrease the threshold level until the requisite number of points has been achieved. Once the adaptive thresholding has reached this level, there are often more points extracted than necessary. For this reason, it is often necessary to sort based on LoG value and truncate to the user specified number of points.

Subregion, Adaptive Threshold, Preset # of GCPs: This technique is a powerful tool for automated large image registration. However, it does require a general image orientation before implementation is possible. This is because both subregions must represent similar areas of the image before automated analysis can commence. The requirement for general image orientation is explored in more detail in section 5.1.

This technique combines aspects of the previous techniques and implements them on image chips (~512 to 2048 work well), instead of the entire image (figure 4.6.1). Once the individual subregions have been analyzed and local GCPs have been determined, they are compiled into an image-wide set of control points. It is easy to generate several hundred GCPs in this manner for large images. With this large amount of GCPs, statistical analysis of RMS Distance Error becomes meaningful and it is possible to reduce the error in the transform model by comparing each match against the model generated from all the GCPs. The analysis of global RMSDE statistics will be analyzed more extensively in Section 4.7.

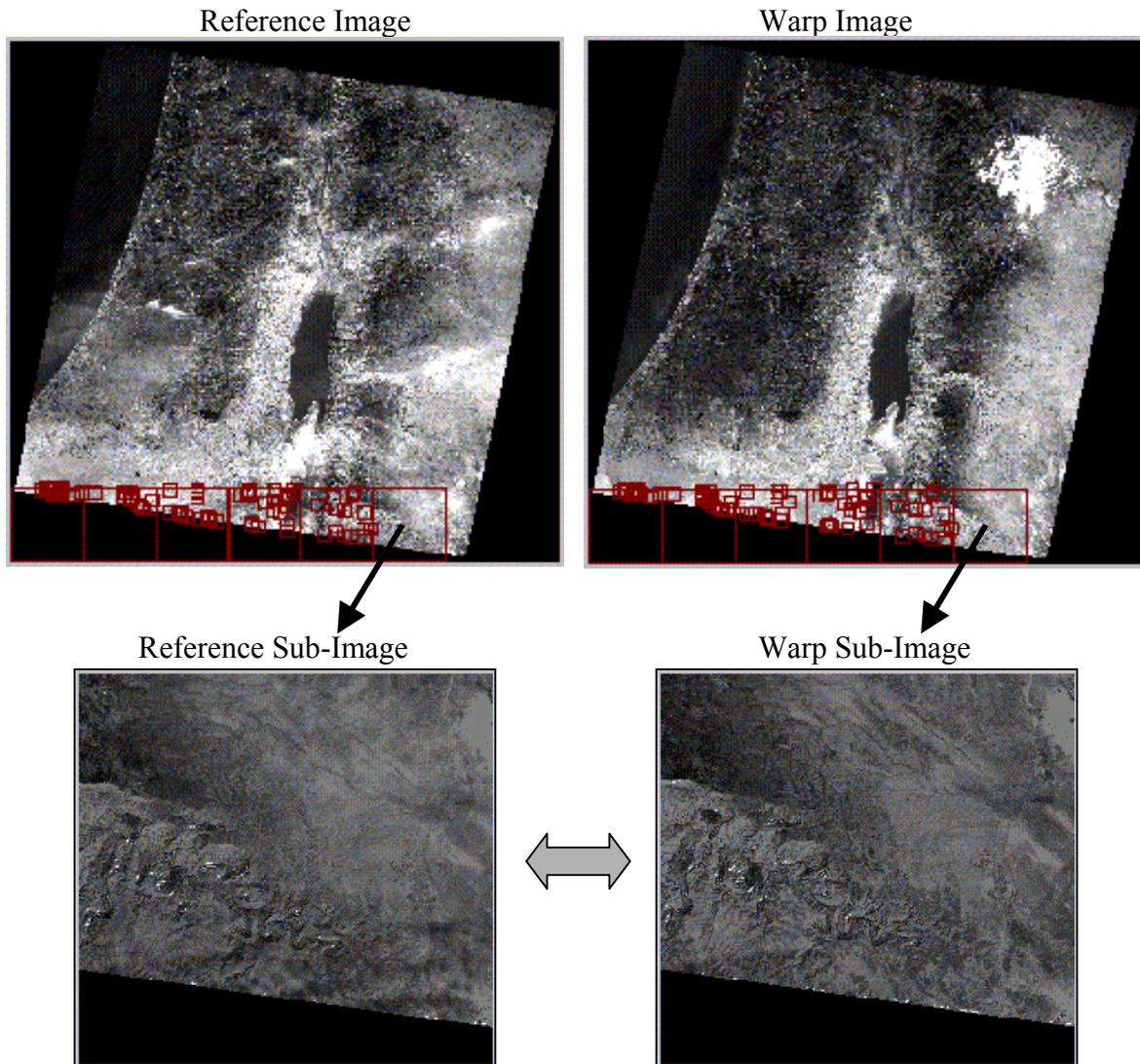


Figure 4.6.1: Subregion Analysis utilized to compare similar areas of the original images.

Subband Analysis: In order to enable the previous subregion analysis technique, it may be necessary to first perform a subband registration, to obtain the general image orientation required. The subband registration technique allows image wide analysis by utilizing datasets of reduced resolution to minimize processing requirements. This ‘reduced dataset’ could simply be a resampled version of the original images at a lower resolution (figure 4.6.2), or could be a ‘scale subband’ resulting from a wavelet decimation sequence. This is often a useful technique since LoG filtering is currently inefficient on datasets over 2k x 2k. The utility of subband analysis will be explored in further detail in section 5.1.

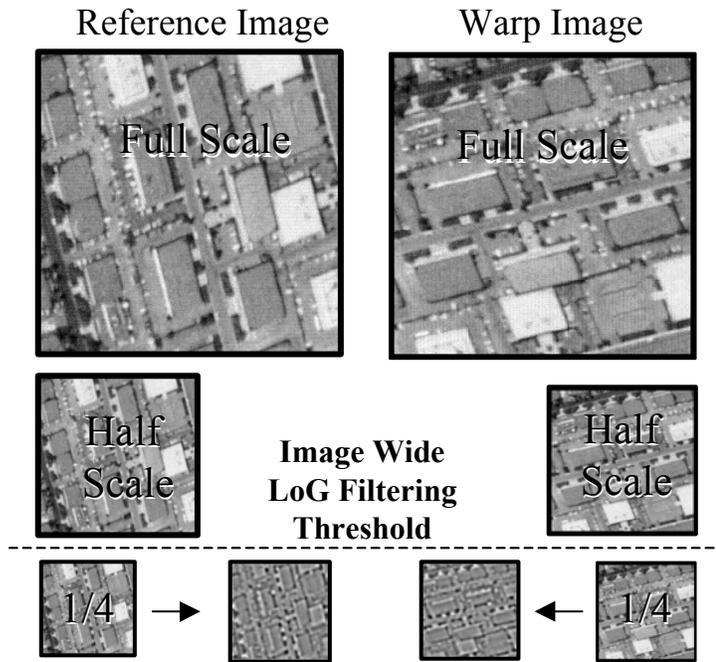


Figure 4.6.2: Reduce the resolution to enable image-wide LoG Filtering.

4.7 Determining the Registration Accuracy

One of the more challenging aspects of image registration is formulating criteria to determine the ‘goodness’ of how well the registration was accomplished. Two general methods focus on grayscale comparison and RMS Distance Error (RMSDE) analysis. Three techniques, which incorporate both of these methods, have been implemented to judge the accuracy of the registration with the LoGWaR program. Figure 4.7.1 shows an example image set, where the Warp image is a rotated, shifted and cropped subset of the Reference image.

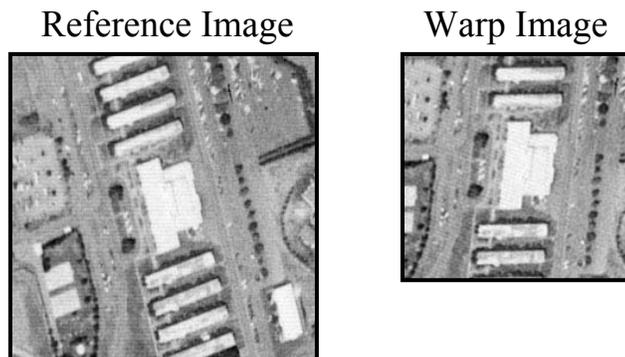


Figure 4.7.1: Sample dataset utilized for image registration.

Pixel Averaging: A useful, if somewhat qualitative, test on registration accuracy is through visual inspection of the overlaid and averaged images. This technique is common in image ‘stacking’ and can be utilized to increase the S/N of an image by averaging out the noise. Through visual inspection, it is easy to determine an inaccurate registration through any blurring that is apparent in the composite image. An accurate registration will appear clear with well defined edge detail. A comparison of a poor registration and a good registration are portrayed in figure 4.7.2.

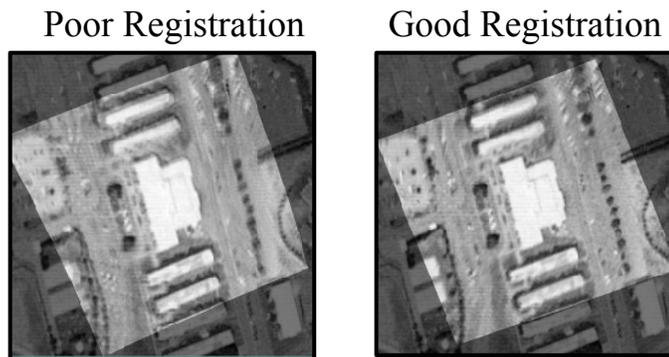


Figure 4.7.2: Comparing registration accuracy through visual inspection of Image Stacking.

Absolute Mean Variance: This metric combines aspects of both qualitative and quantitative analysis to help determine registration accuracy. The visual aspects of this metric can be likened to a ‘difference image’, where a perfect registration would be completely black. Utilizing the sample registrations from the previous example, it is easy to see why this technique is useful for determining registration accuracy.

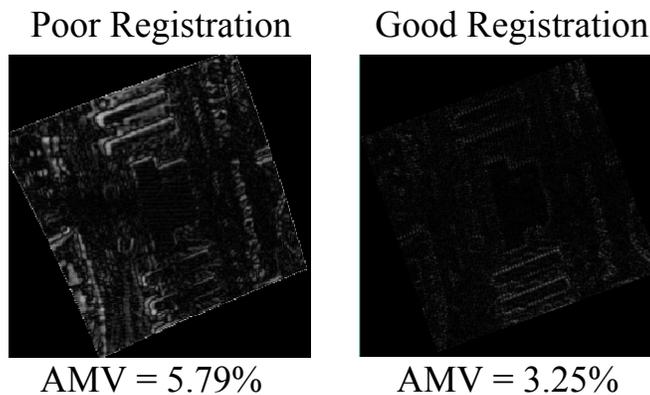


Figure 4.7.3: Comparing registration accuracy through visual inspection of Difference Image.

The quantitative metric involved with this method is based on computing the Absolute Mean Variance (AMV) between the two registered images. This process involves the following steps:

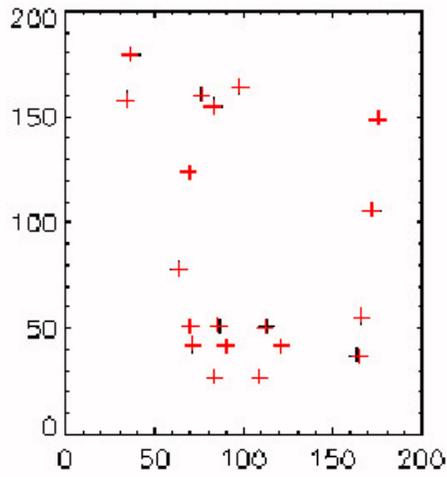
- Determine overlap of Reference Image and Warp Image
- Determine Digital Count Difference per pixel over entire scene
- Average over # of pixels overlap
- Divide by Digital Count Range of Images
- This Metric delivers a Percent AMV between two Images

Deviation from a Polynomial Model: The final metric used to determine registration accuracy, RMS Distance Error, is one of the more common techniques utilized in remote sensing. In fact, the RMSDE is used by ENVI to judge deviation of matches from the prescribed polynomial model to judge registration accuracy. Discriminating bad matches based on deviation from affine/polynomial models is shown in figures 4.7.4. By analyzing the error associated with each matched point from the polynomial model of choice, it is possible to reject bad matches.

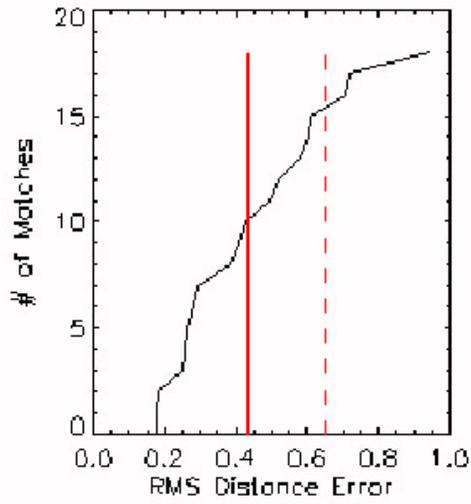
One way to do this is through analysis of the standard deviation from the RMSDE. Any matches that deviate significantly from the mean (greater than 1 standard deviation) can then be removed. The following table contains the raw data utilized by the graphs in figure 4.7.4. If an additional iteration was required to derive a transform with even less error, the matches below the 'cut line' would be removed due to their deviation from the model.

This can be done iteratively to determine a statistical solution that is of low enough error to satisfy the accuracy of registration needed. Figure 4.7.4 shows this process, which utilizes an iterative statistical solution to arrive at an RMSDE of 0.26. In this case, the iterative pruning of match points delivers a registration with almost a quarter-of-a-pixel accuracy.

X-Y Plot of Matches & Predictions

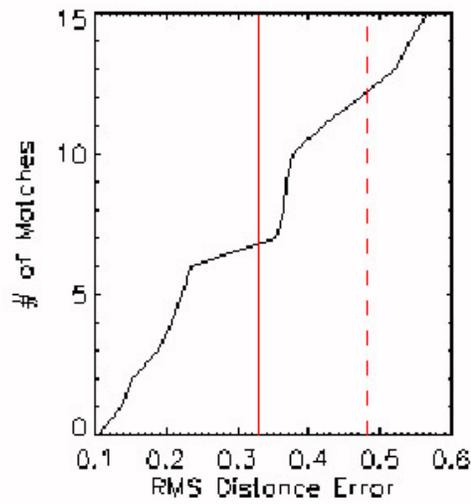
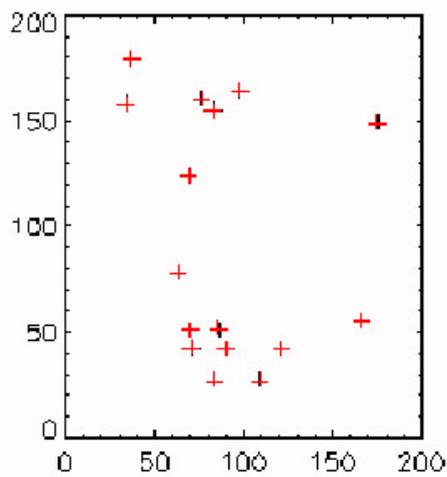


RMSDE Plots

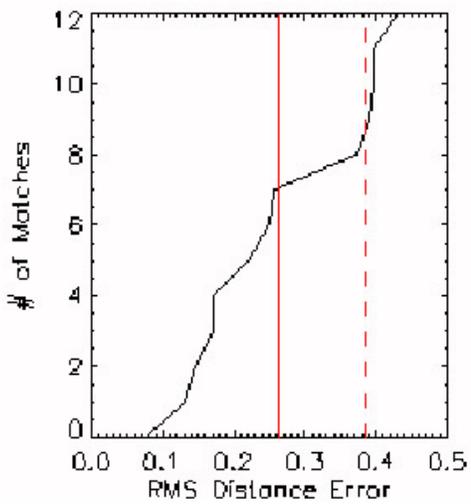
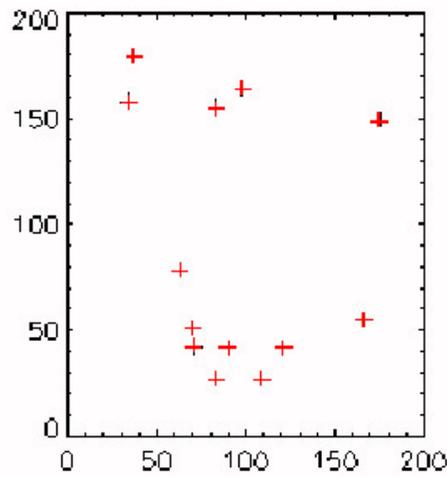


RMSDE

0.44 pix



0.33 pix



0.26 pix

Figure 4.7.4: Comparison of LoGWar matches with polynomial predictions and RMSDE plots.

Table 4.7.1: Removing Bad Matches by comparing Match Error to (RMSDE+1STD).

<i>Ref Coords [x,y],</i>		<i>Warp Coords [x,y],</i>		<i>Pred Coords [x,y],</i>		<i>Coord Error [x,y],</i>		<i>Match Error</i>
180.000	202.000	175.000	149.000	174.959	148.935	-0.0406952	-0.0647125	0.0764448
159.000	65.0000	109.000	27.0000	109.128	27.0283	0.128311	0.0282650	0.131387
203.000	111.000	166.000	55.0000	166.056	55.1356	0.0563507	0.135620	0.146861
165.000	83.0000	121.000	42.0000	120.864	41.8970	-0.136017	-0.103004	0.170618
91.0000	176.000	83.0000	155.000	82.9932	154.829	-0.00677490	-0.170898	0.171033
102.000	190.000	98.0000	164.000	98.0763	164.206	0.0762863	0.205521	0.219222
39.0000	183.000	37.0000	179.000	36.7956	179.142	-0.204445	0.142365	0.249129
44.0000	162.000	34.0000	158.000	34.1693	157.810	0.169254	-0.190292	0.254673
114.000	74.0000	70.0000	51.0000	69.6760	51.1823	-0.324020	0.182308	0.371787
118.000	66.0000	71.0000	42.0000	70.7240	42.2742	-0.275970	0.274208	0.389037
135.000	56.0000	83.0000	27.0000	83.3916	26.9391	0.391579	-0.0609474	0.396293
99.0000	97.0000	63.0000	78.0000	63.3964	77.9842	0.396427	-0.0157852	0.396741
136.000	72.0000	90.0000	42.0000	89.7698	41.6373	-0.230232	-0.362656	0.429565

Overall RMSDE Mean = 0.261753
 Overall RMSDE SDev = 0.121065

$$BadMatches = match_error > (RMSDE + 1STD)$$

$$BadMatches = match_error > (0.261753 + 0.121065)$$

$$BadMatches = match_error > (0.382818)$$

Another statistical technique that LoGWaR incorporates is an approach that iteratively rejects the matched point with greatest RMSDE. The model is recomputed for every removed match until the RMSDE is below a user-defined level. In this way, it is possible to attain a model of the desired RMSDE that is statistically based and is easily incorporated into an automatic image registration process. This technique is used for several of the test cases in Chapter 6 to obtain subpixel accuracy for the transform model and can be referenced for a more detailed explanation.

4.8 Predictive Transforms

In Chapter 5, this research will delve into the ability to cascade several transforms into one single, composite transform. One of the uses of this composite transform is to register images at a lower resolution and then predict the transform at the original scale. When utilizing this predictive transform capability, it is often necessary to have very low RMSDE (representing high levels of subpixel registration accuracy). To predict the transform of an image at 4 times the resolution, the transform will have a model with 4 times the RMSD error. For the example in figure 4.7.4, the 0.26 RMSDE becomes 1.04, or slightly greater than 1 pixel RMSD error.

Although reducing the model error is a useful tool, it is possible to eliminate too many matching GCPs. When this happens, the registration accuracy will deteriorate. Bernstein has shown that the registration error between related images decreases as the number of GCPs increases, however, the quality of GCP accuracy may decrease as their number increases (Bernstein 1987). So, there is a limit to the improvement that can be obtained by reducing the error by eliminating 'good matches' that deviate slightly from the model. This is possibly due to the added benefits that can be gained when utilizing large sets of GCPs for the psuedo-inverse solution when solving for over determined transform expressions. Regardless, determining the optimum balance between number of GCPs and model error is ripe for further research!

The Composite Transform

Often, when attempting to register images, especially large datasets, it is necessary to perform more than one transformation on the images before automated techniques can be employed. This is due to the images requiring a general-orientation relationship before subregion analysis can be employed. Because of this, it is often essential to perform basic manipulations such as scale and/or rotation on an image-wide basis in advance of utilizing automated routines on related subregions (figure 5.1). In some cases, multiple transformations may be the only way to implement automation into the registration process since a datasets may be too large to work with at full resolution.

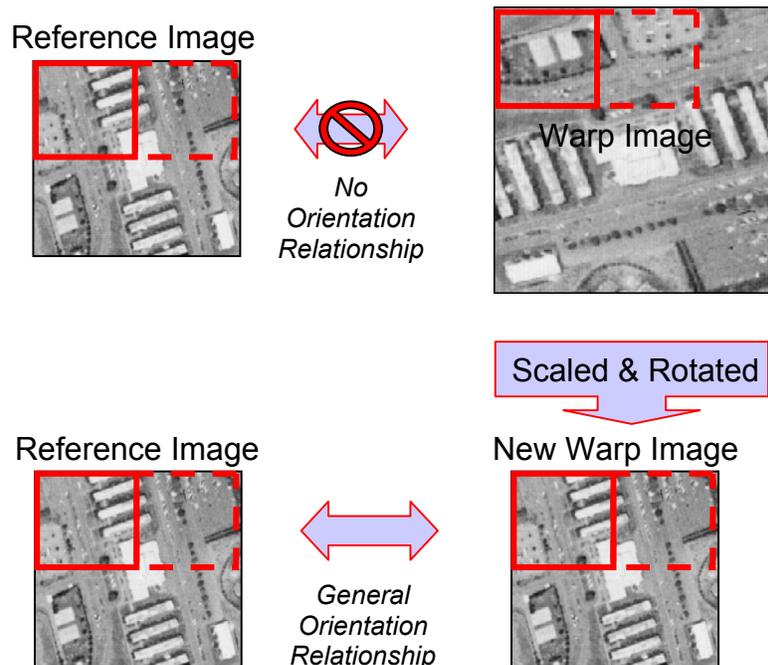


Figure 5.1: Subregion analysis requires a general image orientation relationship.

Unfortunately, each transformation would slightly degrade the data due to interpolation of the grayscale values when resampling (although nearest-neighbor resampling does not suffer from grayscale degradation, it often induces spatial artifacts

on edge detail). Obviously, these degradations are cumulative and can be detrimental to some remote sensing applications. Regardless of the type of sampling utilized and the purpose of registered products, it would be beneficial to consolidate the chain of transforms into a single mathematical model if possible.

In order to adequately manage this requirement, a mechanism should be employed that can efficiently encode multiple transformations at each stage of the registration process and be representative of the general transformation at any given time. According to Wolberg, multiple transforms can be collapsed into a single *composite transformation* by taking advantage of the unique commutative property of the affine transformation (Wolberg, 1990).

5.1 Relating Affines to Polynomial Transformations

Equations 3.3 and 3.4 show the expanded forms of the polynomial equations that can be used to relate two images. The affine model is a subset of this more general formulation that can account for the familiar effects of rotation, scale, translation (RST) and shear when relating the coordinate systems of two images. The affine model includes only the first 3 terms in both the x and y expressions and thus requires six coefficients to adequately define the relationship. The fourth term represents a change in perspective and when taken into account with the affine coefficients represents the 1st Order Polynomial Model. Additional terms in the Polynomial expressions represent

$$\begin{array}{c}
 x = a_0 + a_1x' + a_2y' + a_3x'y' + a_4x'^2 + a_5y'^2 \\
 y = b_0 + b_1y' + b_2x' + b_3y'x' + b_4y'^2 + b_5x'^2 \\
 \underbrace{\hspace{10em}} \\
 \text{Affine Transforms} \\
 \underbrace{\hspace{10em}} \\
 \text{1st Order Polynomial} \\
 \underbrace{\hspace{10em}} \\
 \text{2nd Order Polynomial}
 \end{array}$$

Figure 5.1.1: Relating the Affine Transform to the Polynomial Expression

higher orders and thus a more complex relationship between the images. Dissecting equations 3.3 and 3.4 we can identify the portions of the general polynomial equations that relate to the affine, 1st Order and 2nd order expressions as shown in figure 5.1.1.

5.2 The 3 x 3 Affine Representation

Section 3.1.1 alludes to the general utility of matrix formulations in image transformations; this section will expand upon that concept in some detail. Since the product of affine transformations is also affine, they can be used to perform a general orientation of a set of points relative to an arbitrary coordinate system (Wolberg, 1990). This is a unique property of affine relationships, since higher order polynomials generally cannot be combined with a commutative operation. Developing a composite transform for affines relies heavily on this property, which can be easily implemented in a simple 3 x 3 matrix such as the forward mapping function expressed in equation 5.1:

$$(5.1) \quad \begin{bmatrix} x' & y' & 1 \end{bmatrix} = \begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} a_2 & b_2 & 0 \\ a_1 & b_1 & 0 \\ a_0 & b_0 & 1 \end{bmatrix}$$

It should be noted that a 2 x 3 matrix contains enough elements to define an affine relationship (6 coefficients for 6 unknowns), however the symmetry of a 3 x 3 is useful for commutation when cascading elements of the composite transform. Example representations of the Affine warps, as defined by Wolberg, in the 3 x 3 matrix notation follow (figure 5.2.1).

SHIFT:	ROTATION:	SCALE:	SHEAR:
$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ T_x & T_y & 1 \end{bmatrix}$	$\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & H_x & 0 \\ H_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Figure 5.2.1: Relating the Affine Transforms to the 3 x 3 Matrix

If an affine transform is deemed adequate to describe the relationship, the six coefficients may be derived by specifying the coordinate correspondence of three noncollinear points in both images (Wolberg, pg 50, 1990). This is demonstrated in equation 5.2 where x' and y' represent the noncollinear points in the “warp image”, x and y represent the related points in the “reference image” and a_N and b_N are the unknown mapping coefficients.

$$(5.2) \quad \begin{bmatrix} x'_1 & y'_1 & 1 \\ x'_2 & y'_2 & 1 \\ x'_3 & y'_3 & 1 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix} \begin{bmatrix} a_2 & b_2 & 0 \\ a_1 & b_1 & 0 \\ a_0 & b_0 & 1 \end{bmatrix}$$

With the selection of additional GCP matches, the affine problem becomes over-defined. By solving for the coefficients using the ‘psuedo-inverse’ solution to the least squares problem it is possible to estimate the best fit of the GCPs to the affine model (equations 5.3-5.6). This solution can generate a sub-pixel registration of the datasets if the GCPs are accurate to the affine model and well distributed.

$$(5.3) \quad \begin{bmatrix} x'_1 & y'_1 & 1 \\ x'_2 & y'_2 & 1 \\ x'_3 & y'_3 & 1 \\ \vdots & \vdots & \vdots \\ x'_N & y'_N & 1 \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & 1 \end{bmatrix} \begin{bmatrix} a_2 & b_2 & 0 \\ a_1 & b_1 & 0 \\ a_0 & b_0 & 1 \end{bmatrix}$$

$$(5.4) \quad U = WA$$

$$(5.5) \quad A = U^{-1}W$$

$$(5.6) \quad A = (U^T U)^{-1} U^T W$$

Due to the dimensions of the GCP matrices, the psuedo-inverse solution (eq 5.6) is utilized to solve the linear least-squares problem since non-square matrices have no inverse.

It should be noted that with the addition of the third column, in the matrix, all the 1st Order relationships could be solved for including perspective as noted in the following normalized ($c_0=1$), general relationship:

$$(5.3) \quad \begin{bmatrix} x'w' & y'w' & w' \end{bmatrix} = \begin{bmatrix} x & y & w \end{bmatrix} \begin{bmatrix} a_2 & b_2 & c_2 \\ a_1 & b_1 & c_1 \\ a_0 & b_0 & 1 \end{bmatrix}$$

For perspective mapping projections, coefficients c_1 and c_2 are nonzero and can be related to the 1st order polynomial expression in figure 5.1.1.

Fortunately, many requirements for remote sensing registration can be accommodated with an affine or 1st Order model to relating the datasets. This is especially true if the data is orthorectified. Minimally, a general orientation or geo-reference can be achieved with the composite affine in order to relate the images spatially (Schowengerdt, 1997). If an affine model cannot obtain the relationship needed, then a higher order model can be utilized in conjunction to obtain the precision necessary. Thus minimizing the degradations due to interpolation to two separate operations. Although, comprehensive validation has not been performed on the combination of a composite affine and a 1st Order, results have been promising. It is quite possible that one combination is mathematically sound, whereas additional commutations are not. If this proves to be the case, then it is possible to solve for most remote sensing global aberrations with one composite transform; since perspective can be incorporated (if only once) into the composite model.

There is a great benefit in utilizing the composite affine to obtain an approximate registration of two datasets. The ease and flexibility it provides in combining basic geometric manipulations and the ability to precisely log, in concise mathematical notation, any changes that have been made to the data make it a powerful tool in the quest for automated registration. Additionally, the use of affine transformations for approximate correction of satellite sensor distortions has been noted in literature (Schowengerdt 1997).

Although point matching techniques have not been developed here to address shear and perspective, both of these deformations can be incorporated into the 3 x 3 model to relate datasets. Occasionally, with large images, the subregion technique can be utilized to locally relate datasets that have relatively small global shear, perspective, and even earth curvature (2nd Order) distortions. This is due to local regions exhibiting affine relationships, even though higher order global distortions exist. Once related in this manner, the LoGWaR software facilitates RMSDE analysis and warping of data at the 1st and 2nd Order via a user slider bar at the lower right hand corner. Additionally, if the warp dataset has not been corrected for sensor viewing geometries (orthorectified), it is often trivial for the user to select four dispersed GCPs to provide a 1st Order relationship between it and the reference image. By warping a dataset with this transform, it is then possible to re-register the warped image utilizing automation techniques to precisely relate the images at the subpixel level. Both transforms can then be incorporated into one composite that can be used to warp the original. This technique has shown great utility when working with non-orthorectified data.

5.3 Creating the Composite Transform

By utilizing the commutative property of the 3 x 3 affine representation, it is possible to combine several individual affine transforms into one composite transform. This consolidation of matrices provides processing efficiency and can be utilized to avoid multiple resamplings (Schowengerdt, 1997, pg 334). Equation 5.4 represents the cascading of affines ($A_1A_2A_3$) through a commutative process into a single composite transform(M_{comp}).

$$(5.4) \quad M_{comp} = A_1A_2A_3$$

Utilizing Wolberg's 3 x 3 representation of individual affine transforms, the following formulation depicts the cascading of an image relationship containing translation, rotation and shear into a composite matrix (figure 5.3.1). The convenience of this notation is

realized in the easy multiplication of matrices within many computing environments and the minimal overhead of a single 3 x 3 matrix.

As mentioned earlier in this chapter, it is often useful (especially with automated processes) to first determine the general orientation relationship between datasets. Once this is accomplished, automated subregion analysis can be utilized to produce many GCP with good distribution throughout the image for precise (subpixel) registration. By utilizing this cascading approach, it is possible to combine the general orientation affine with the precise subregion affine to produce a single solution that best relates the two datasets.

$$M_{comp} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ T_x & T_y & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$M_{comp} = \begin{bmatrix} S_x \cos\theta & S_y \sin\theta & 0 \\ -S_x \sin\theta & S_y \cos\theta & 0 \\ S_x(T_x \cos\theta - T_y \sin\theta) & S_y(T_x \sin\theta + T_y \cos\theta) & 1 \end{bmatrix}$$

Figure 5.3.1: Creating a Composite Affine Transform from 3 separate Affines.

5.4 Utilizing the 3 x 3 Affine with Wavelets for Predictive Transforms

An interesting application of the 3 x 3 formulation involves manipulating the last element in the matrix (coefficient c_0). This element represents the overall scale relationship between the two images and so can be easily manipulated to conveniently keep track of any scale changes from the base image. So, if an image is decimated once, using the partial FWT, the composite transform would equate to the product of the *current affine* relationship and a *scale modifier* matrix with coefficient c_0 equal to 2 (if reconstituted with the FWT⁻¹, then a_{33} would be equal to 1/2):

$$M_{comp} = \begin{matrix} \text{Current Affine} & \text{Scale Modifier} \\ \begin{bmatrix} a_2 & b_2 & 0 \\ a_1 & b_1 & 0 \\ a_0 & b_0 & 1 \end{bmatrix} & \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \end{matrix}$$

Figure 5.4.1: Manipulating the global scale coefficient in the Composite Affine Transform.

This result represents the new relationship between the two images at half the scale of the original warp image. The process is similar for any scale relationship change between the two images.

One of the most important implications of utilizing this technique is that predictive transformations can be developed utilizing much lower resolution version of the original images! So, when working with large images, it is possible to reduce the size of the datasets, register at the lower resolution and then predict the transform required to register the images at their original scale. This can even be accomplished with images much too large to load into virtual memory and manipulate at their native resolution.

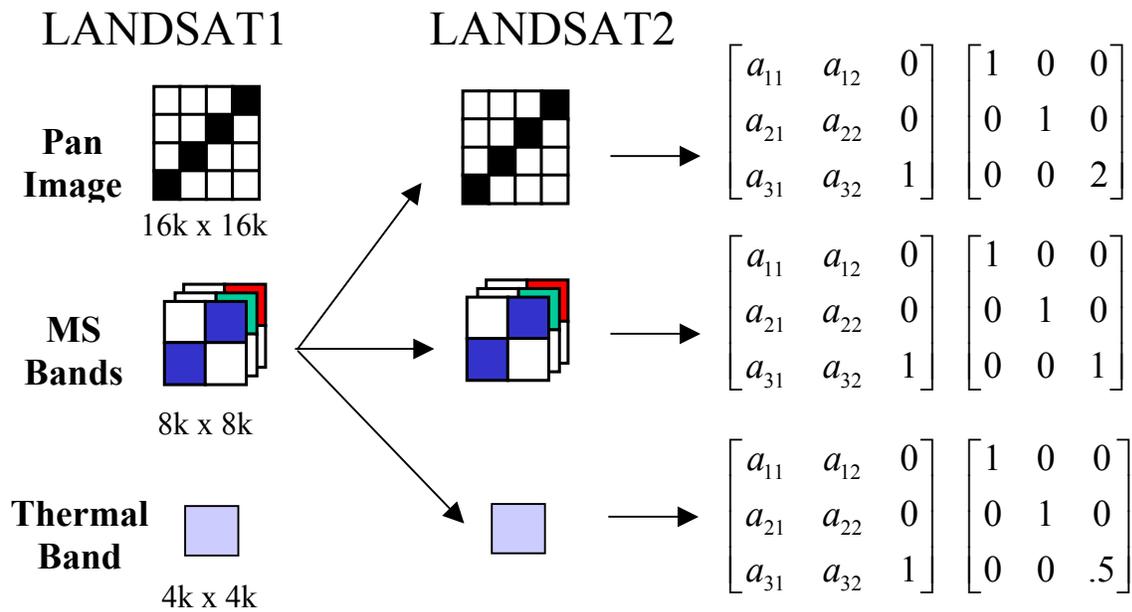


Figure 5.4.2: Utilizing the Affine Transform Matrix to Relate Multi-Scale image sets.

Figure 5.4.2 demonstrates this capability as it is applied to a pair of LANDSAT datasets that have been related by only registering a single band in the MS and predicting the transforms for the additional MS, Pan, and Thermal bands.

Although registration precision decreases in relationship to the reduced scale that the registration takes place at, it is still possible to have a good prediction of the registration at the native resolution if the subpixel accuracy of the registration is sufficient. Once again, the precision of the registration required, will often be predicated on the type of product or additional processing that is required of the registered datasets.

The obvious benefits of utilizing this approach are processing speed and more critically overcoming the virtual memory limitations of processing many of today's enormous datasets. An additional benefit of implementing scaling changes in the wavelet domain is that an iterative solution could be implemented that would allow registration at a low resolution "scale" subband level and allow the ability to predict transforms at various resolutions. This would allow a user to view the registered datasets at various resolutions by predicting the transform required when transitioning (zooming) from one resolution to the next in the wavelet "resolution pyramid" and then re-registering at the higher resolution for a more precise result. In this fashion, an analyst could immediately reap the benefits of a predicted transform at the new resolution and while dwelling there a more precise relationship could be developed as a background process.

In this manner, transitioning between various resolutions can be accomplished while still maintaining a basic spatial relationship between the two datasets. This can be accomplished quickly since only a simple 3×3 matrix multiplication and the transformation of the image at the new resolution (LL subband level) is required for the prediction, not an entirely new registration. Also, for purely automated registration of large datasets, it would be essential to register images at a low resolution to establish the initial relationship before subregion analysis could be utilized to establish a more precise relationship at the original resolution (figure 5.1).

REGISTRATION TEST CASES

In order to demonstrate the capabilities of the LoGWaR process, it will be tested with three datasets. The first dataset will be used as a control test and will be the same image that has been copied and pre-warped (with known geometric distortions) and then sent through the registration process to test the results. A second test for this research will be two large (8k x 8k) LANDSAT images and will demonstrate the capability for multiresolution registration from the same sensor. The final tertiary test case will include a multispectral image (CITIPIX) and a hyperspectral cube (HYMAP), demonstrating the capabilities to register images from two dissimilar sensors (an MS framing sensor and a HS linescanner).

6.1 Control Test: Same Image – shifted and rotated.

The image utilized for this test is a color Space Imaging shot of the Capitol Building in Washington D.C., converted to grayscale. The reference image is 300 x 300 pixels in dimension while the warp image has been cropped to 250 x 250 pixels (25 to 274) and later rotated by 135 degrees. Since the cropping procedure and the rotation will both induce shift into the image-to-image relationship, they be tested separately.

Shift: The cropping procedure extracts the central portion of the reference image (figure 6.1.2). This causes a translation relationship between the two images of 25 pixels in both the x & y . The ability to use a statistical model to reject the matches of greatest error and improve the accuracy of the registration is evident in figure 6.1.1. The following results demonstrate LoGWaR's capability to determine the shift and show the utility of both the RMSDE & AMV:

```

RMSDE computed for      1st Degree Polynomial.
Overall Root Mean Square Distance Error (RMSDE) = 0.000000
Overall RMSDE Standard Deviation from Mean =      0.000000

Affine Transform:  1.00000      1.44169e-006      0.000000
                  3.15602e-007      1.00000      0.000000
                  -25.0000      -25.0001      1.00000

Registration Metric (Abs Mean Var) = 0.00208408
    
```

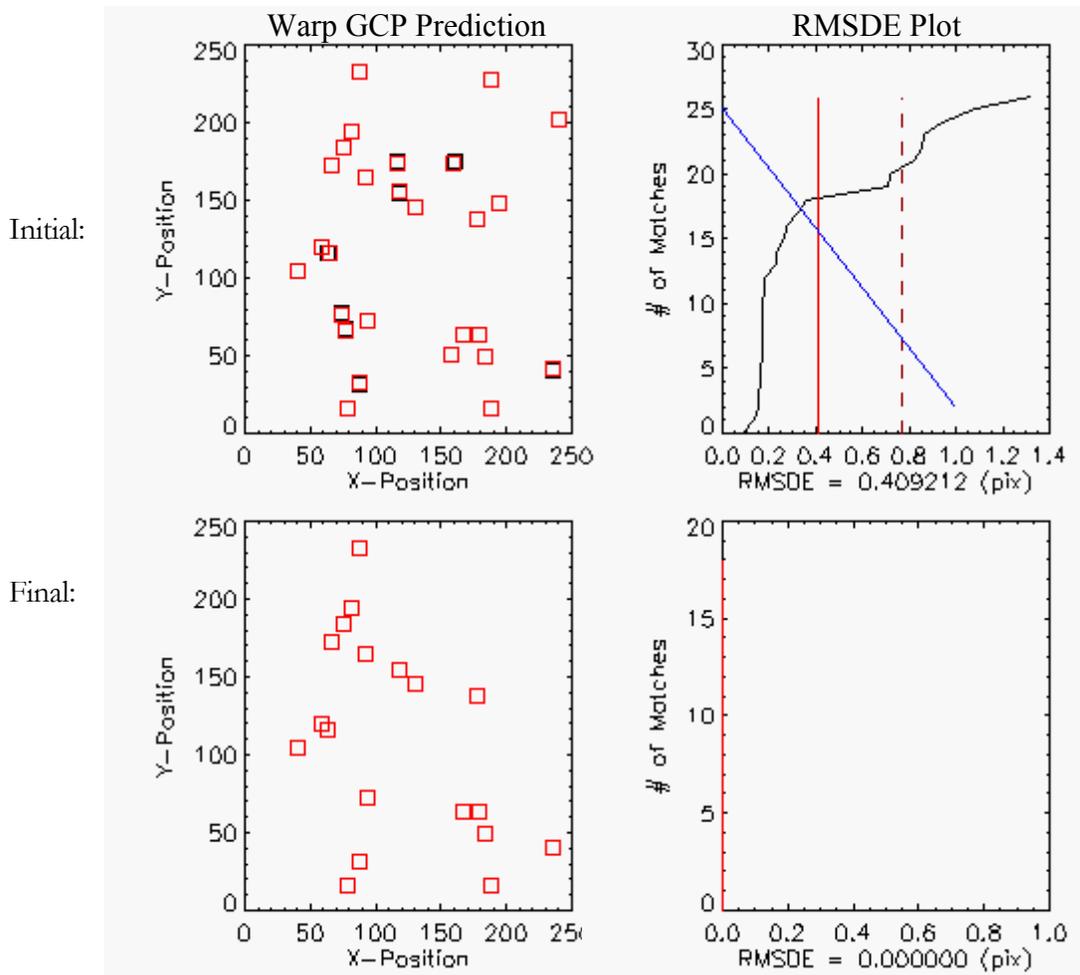


Figure 6.1.1: Utilizing RMSDE analysis to increase the accuracy of the registration model (Initial to Final).
 Image Legend: Black square boxes represent the Warp GCP locations compared to the Red Predicted GCPs.
 Plot Legend: Black Line is cumulative RMSDE, Blue Line is Histogram, Red Line is Mean, and dashed is 1STD.

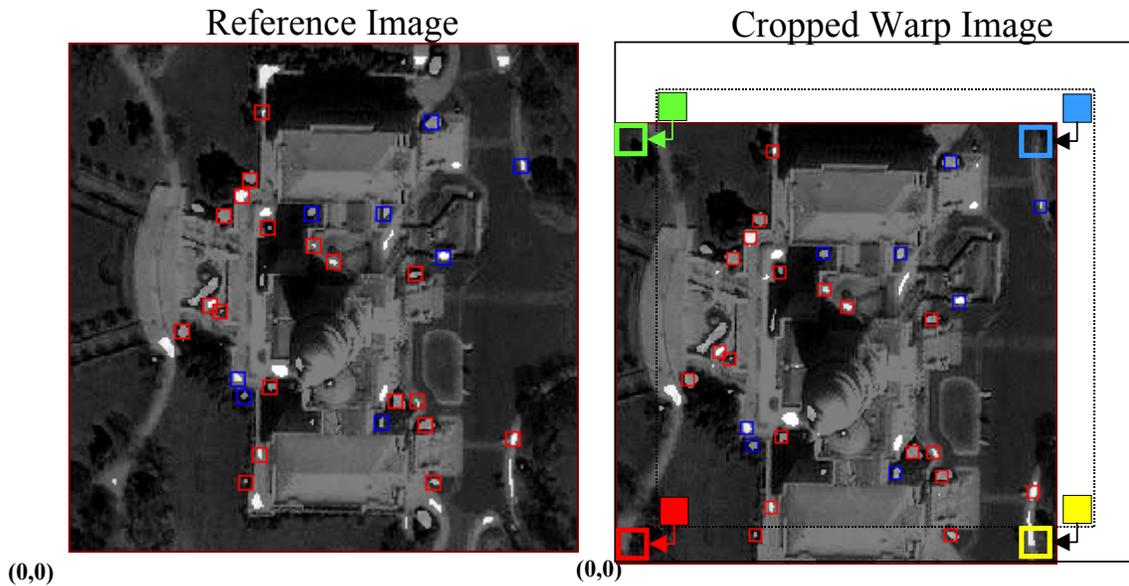


Figure 6.1.2: The crop has induced a -25 pixel shift in the x & y for the inverse transform sampling.



Figure 6.1.3: Confirming the registration results through visual inspection of overlaid and AMV images.

Figure 6.1.3 shows how the overlaid images and the Absolute Mean Variance can be used as visual evidence of a good registration. The lack of any definable information in the AMV image is due to the accuracy of the registration and the fact that the grayscale values for the related areas are exactly the same.

Rotate: Now the warp image is an exact replica of the reference image except that it has been rotated 90 degrees clockwise. The following results are non-intuitive, but, accurate non-the-less. Although the following results show the correct -90 degree rotation, the inherent shift of 299 pixels in the y direction is demonstrated in figure 6.1.5. For the image to be rotated by 90 degrees, it is necessary to rotate each pixel about the origin and then translate in the y -direction by 299 pixels for inverse transform sampling. The LoGWaR results for the transform follow:

```

RMSDE computed for      1st Degree Polynomial.

Overall Root Mean Square Distance Error (RMSDE) =  0.000000
Overall RMSDE Standard Deviation from Mean =  0.000000

Affine Transform: -3.71925e-015  -1.00000  0.000000
                  1.00000        1.81799e-015  0.000000
                  2.02505e-013  299.000    1.00000

Registration Metric (Abs Mean Var) =  0.0139355

```

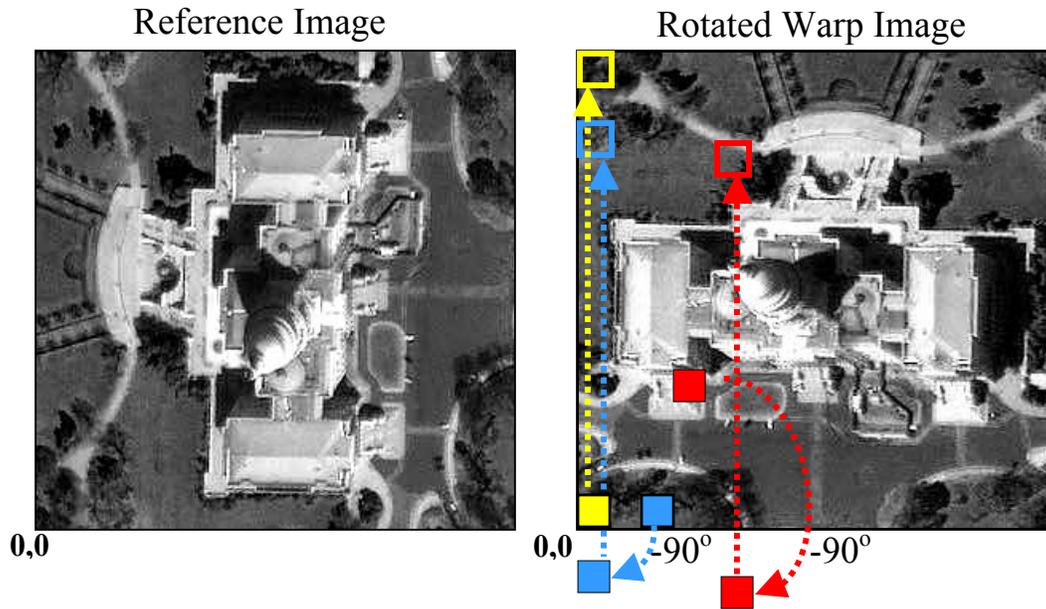


Figure 6.1.4: Use the transform⁻¹ for pixel sampling: including a shift and rotation about the origin.

Since the initial registration had very low AMV and RMSDE, additional statistical analysis of the matched points for improvement to the model is not necessary. The RMSDE plots are visible in figure 6.1.4 and the AMV is displayed in 6.1.6.

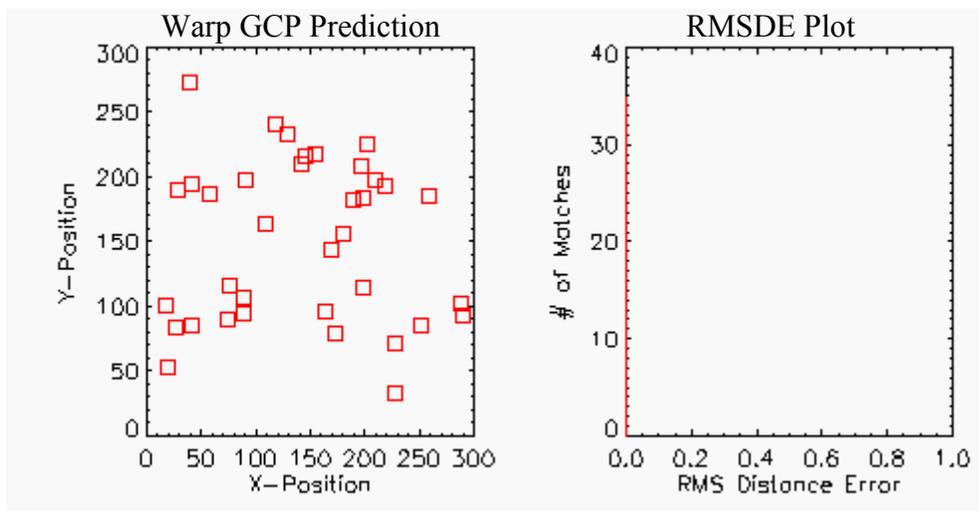


Figure 6.1.5: Utilizing RMSDE analysis to judge the accuracy of the registration model.

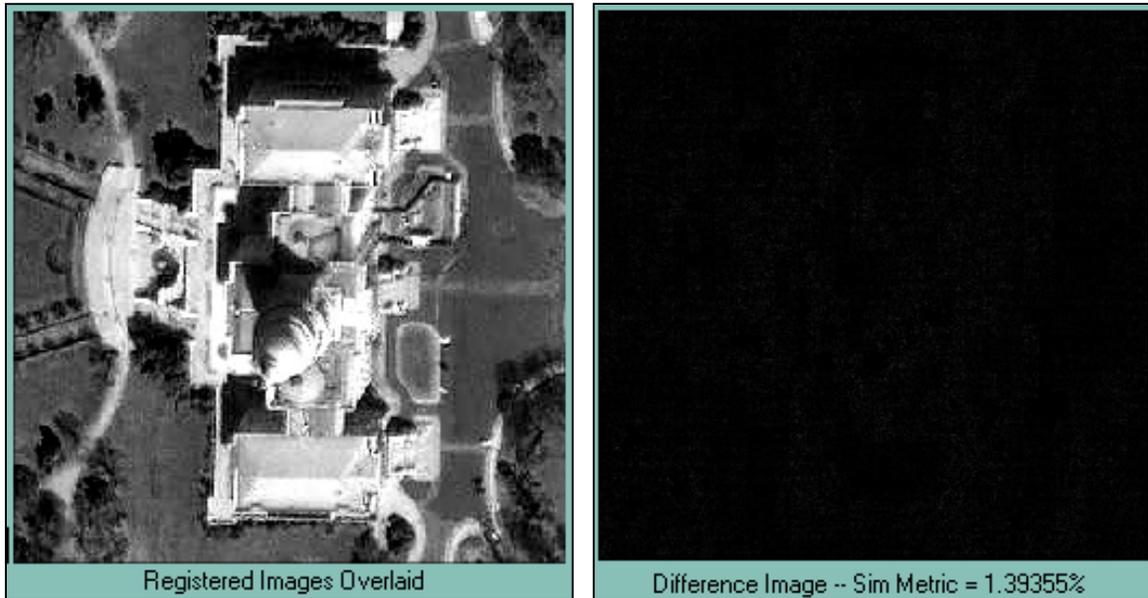


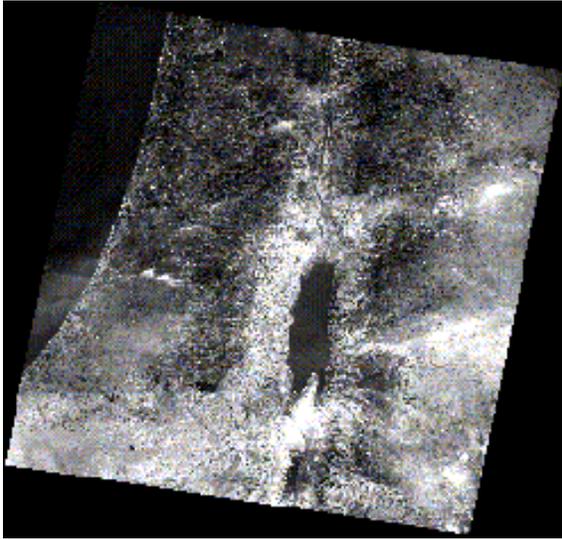
Figure 6.1.6: Confirming the registration results with visual inspection of overlaid and AMV images.

As can be seen from the results, of both the translated and rotated images, the LoGWar algorithms accurately determined the transformation needed to register the images to an almost perfect statistical accuracy. The extremely low values of the AMV and the visual inspection techniques also corroborate these results.

6.2 Multispectral/Multiresolution Registration with LANDSAT.

The registration of LANDSAT datasets offers an opportunity to test LoGWar's capabilities on large multispectral datasets (8k-16k) with multiple resolutions for the thermal, MS, and panchromatic bands. This will illustrate some of the challenges in performing accurate registrations with automated techniques. The datasets for this test were taken on January 14th and February 15th of 2000, and cover the region of Jericho, Israel (figure 6.2.1). The LANDSAT datasets were ordered as 'G-1' products and exhibited a good image orientation relationship. However, the data was not registered well enough to perform the pixel-to-pixel comparison necessary for change detection analysis. Additionally, the relatively low-resolution LANDSAT datasets often prove difficult to perform accurate supervised registrations, due to the difficulty in isolating precisely related GCPs.

Reference Image (7747x7291)



Warp Image (7750x7250)

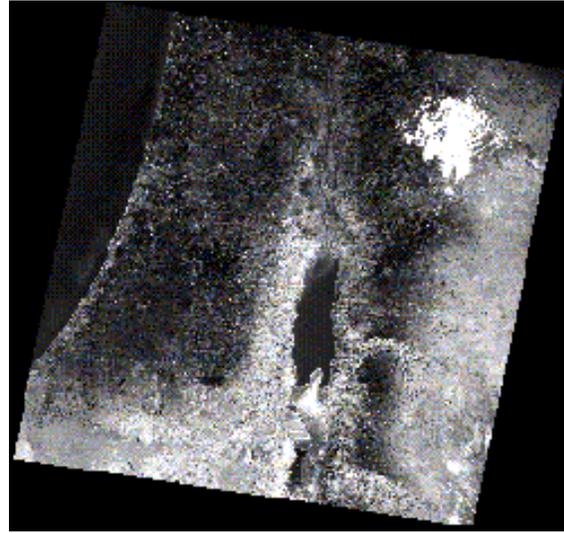


Figure 6.2.1: The test datasets of Jericho, Israel were taken in January and February of 2000.

Due to virtual memory restrictions in IDL and processing speed, there are two approaches that can be taken with LoGWaR software when attempting to register images of this size. Since the entire image cannot be filtered with the LoG filter, it is essential to either process the image in subregions or in subbands. Both of these processes will be analyzed in the context of this example.

Subregion Registration: Since these datasets exhibit a good image orientation relationship, it is possible to directly relate subregions within the images. Subregion analysis utilizes the same techniques as image-wide analysis, but is applied over successive areas of both images (figure 6.2.2). The subregion results are eventually combined to form the transform for the entire 'warp' dataset. If the user has defined a focused ROI, by boxing an area with the mouse, the *Subregion Tool* will subdivide that ROI for analysis, ignoring regions outside of the highlighted area. The subregion size is determined in the *User Preference, File* pull-down menu, within LoGWaR (default = 512). Any residual portions of the image, at the edges, will be treated as additional subregions of smaller size. A 512x512 subregion can take from 20-45 seconds based on the criteria utilized for point matching. The *Subregion Tool* automatically compensates for offset from the origin based on the user defined ROI size, thus maintaining the proper relationship between GCPs derived in each subregion relative to the original image.

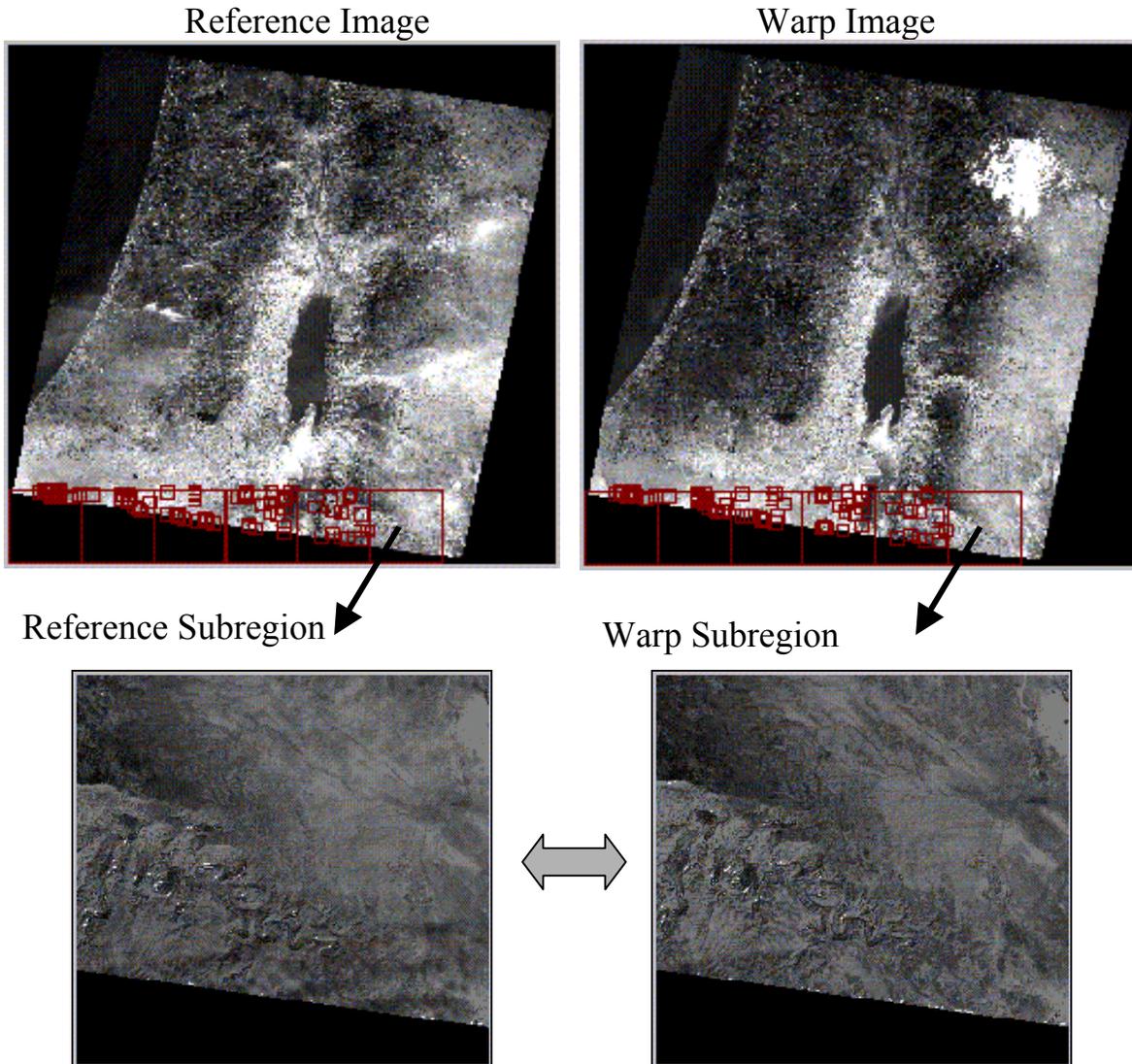


Figure 6.2.2: Subregion analysis utilized to compare similar areas of the original images.

The subregions collect local GCPs, until the entire image (or supervised focus area) has been filtered and analyzed. Once the entire image has been processed, it may be necessary to perform a check on the matched points for false matches, especially in the presence of water or clouds . It is then possible remove bad matches using an ROI tool or through statistical analysis. The resulting GCPs can be viewed in figure 6.2.3. The automated nature of the LoGWaR software can produce several hundred GCPs. The accuracy of registration has been shown to increase with the number of GCPs (Bernstein, 1987), however, the chances for bad matches to be incorporated into the transform model also increases.

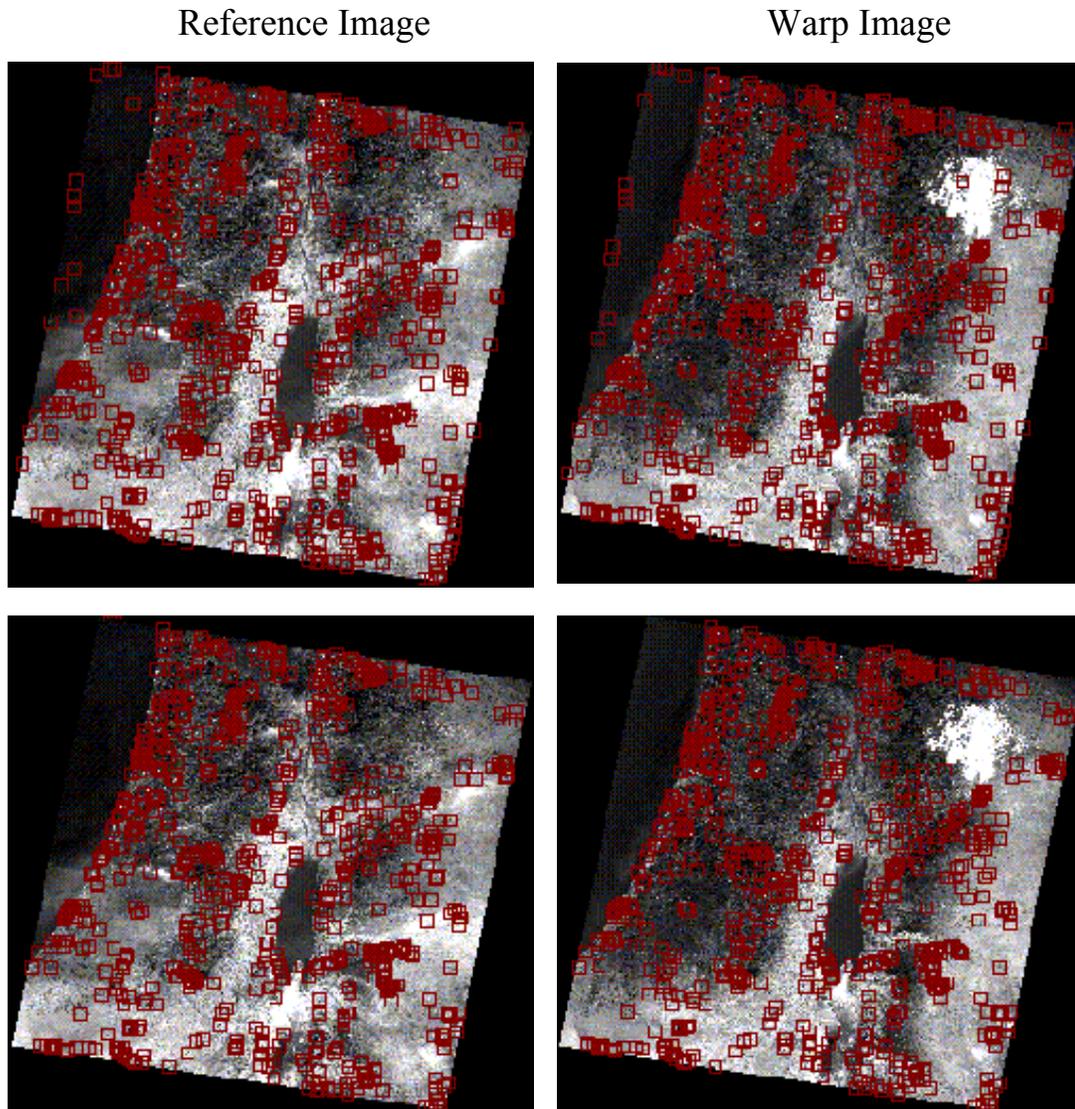


Figure 6.2.3: ROI tools and statistical analysis can be utilized to discriminate and reject bad matches.

A useful aspect of the subregion registration process is that it limits the error associated with a bad match to the dimensions of the subregion that is enforced. Thus, a 512×512 subregion will limit a bad match error to $\sqrt{512^2 + 512^2}$. This is much smaller than a bad match obtained with an image-wide technique that is limited to $\sqrt{7747^2 + 7291^2}$. This helps mitigate the impact introduced by a bad match into the model.

Statistical analysis of the GCPs enables iterative rejection of matches with the highest error. This continues until a subpixel accuracy of matches to a transform model is achieved.

(figure 6.2.4). The initial model was obtained using 492 GCPs. This was further refined by removal of matches > 1 STD from the mean RMSDE, which resulted in 444 final GCPs:

RMSDE computed for 1 Degree Polynomial.
 Overall Root Mean Square Distance Error (RMSDE) = 0.693404
 Overall RMSDE Standard Deviation from Mean = 0.0774224

Saving Affine Transform:
 0.999990 1.77472e-005 0.000000
 -3.59437e-006 0.999992 0.000000
 -6.36321 9.41334 1.00000
 Registration Metric (Abs Mean Var) = 0.113350

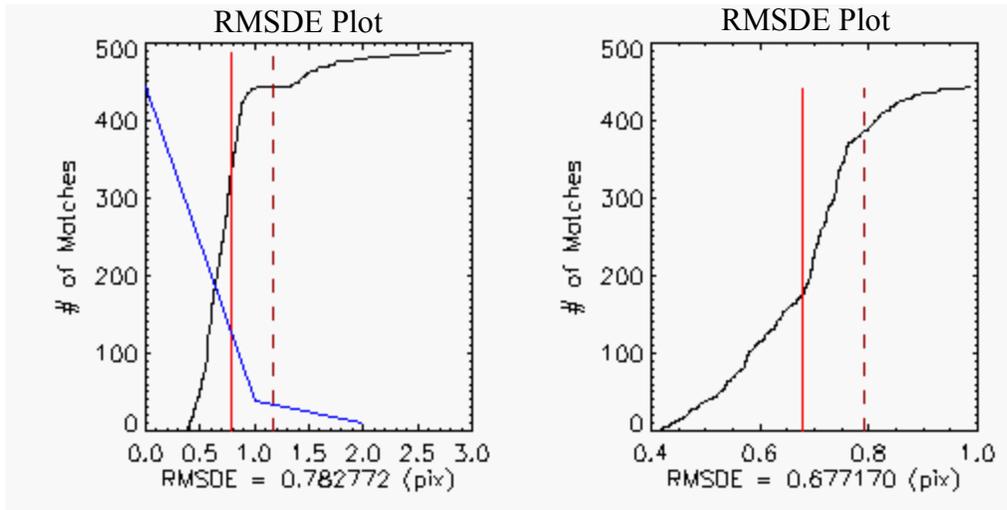


Figure 6.2.4: Utilizing RMSDE analysis to judge & improve the accuracy of the registration model.

The results of the registration were quite good, with excellent edge alignment across the image. The overlaid image clarity and the AMV image (11.335%) corroborate this result:

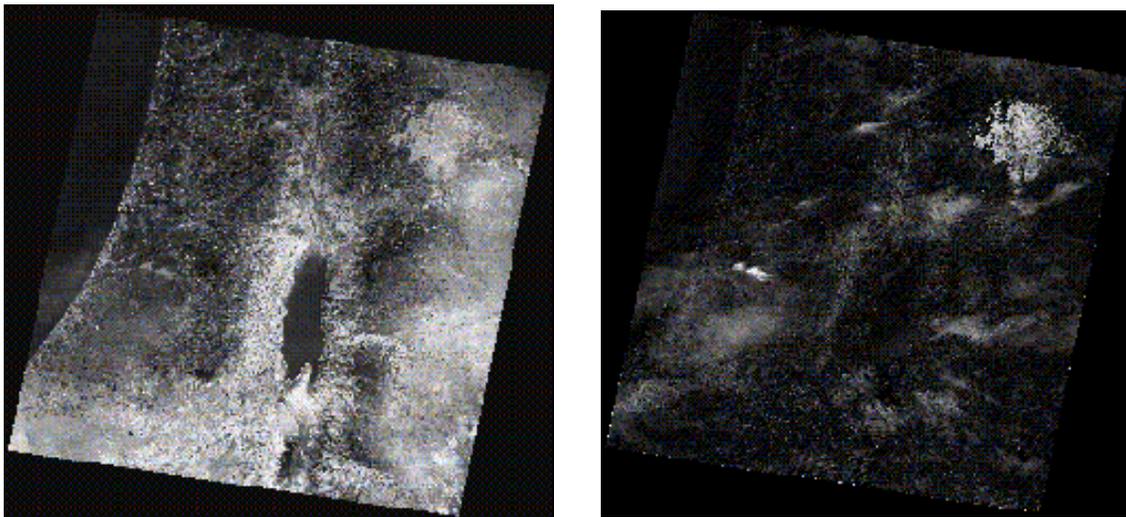
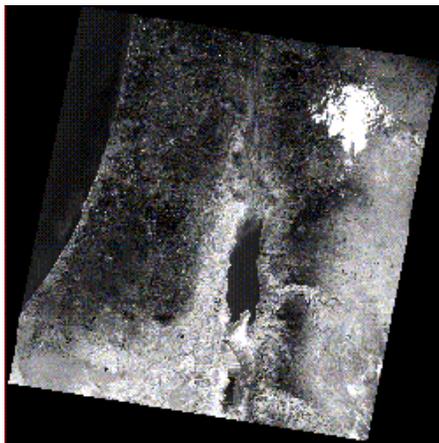


Figure 6.2.5 Confirming the registration results with visual inspection of overlaid and AMV images.

Subband Registration: A useful alternative to subregion registration is the ability to perform image-wide analysis at a reduced scale of the original. In fact, this may be the only alternative when attempting to attain an orientation relationship (section 5.1), on images that exceed the LoG filtering size limitation. The ability to register images at a lower resolution and predict the transform of the original image is a powerful technique. It facilitates efficient processing, reduces virtual memory requirements and allows the registration of very large images.

For this example, the Jericho dataset will again be utilized to demonstrate the capabilities of subband registration and predictive transformation. Two options are now available, simply scaling the image to a lower resolution or decimating the image with the wavelet transform. Scaling the image will achieve the same registration results as the wavelet decimation, but, will sacrifice the high-frequency spatial information in the image. However, wavelet decimation stores the spatial detail and allows for later use in applications such as sharpening. The reduction of the original image resolution utilizing both techniques is demonstrated in 6.2.6.

Scaled Warp Image (1937x1822)



Decimated Warp Image (2048x2048)



Figure 6.2.6 Access subbands through scaling or wavelet decimation (note dyadic dimensions).

Now that the images have been reduced in scale to a size that allows image-wide LoG filtering, it can be processed for registration with the LoGWaR program. If the user defined point limit (50 pts) does not provide enough matches, the ROI tool can be utilized to identify specific regions to extract additional matches from the image (figure 6.2.7).

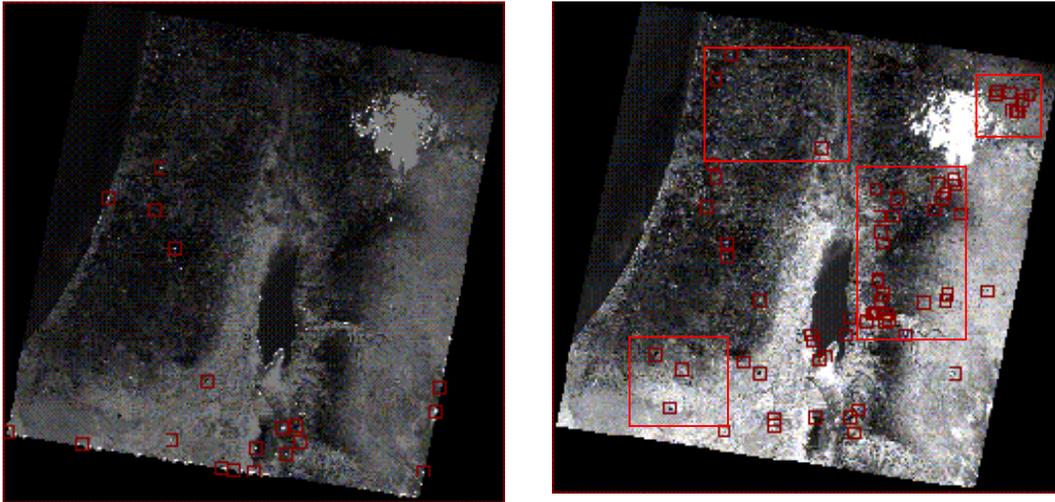


Figure 6.2.7 Utilizing the ROI tool to isolate specific areas for registration.

These matches are then be utilized to relate the reference and warp image at the reduced scale. If this is accomplished with a high enough degree of precision, the prediction can produce results similar to the transform at the original resolution. To test this hypothesis, the reduced scale registration model will be adjusted through RMSDE analysis of the matches. Three models will be utilized for the prediction: a) All Matches from reduced scale, b) All matches minus matches greater than 1STD of the Mean RMSDE, and c) Only those matches with less RMSDE than 20% of a pixel. The plots of the matches and the resulting RMSDE can be seen in figure 6.2.9. Once the required RMSDE has been achieved, those matches are utilized to derive an affine model to relate the low-resolution images. Now it is possible to predict the affine that will relate the original resolution images through the commutative property of the affine, which is discussed in Chapter 5. This is accomplished by multiplying the affine relationship at the low-resolution by the scale modifier as depicted in figure 6.2.9. Once this has been accomplished, the resulting affine models will be compared to the original affine transform which was acquired at the full resolution. Since the original affine provided a high quality registration, it will be utilized to judge the accuracy of the predicted transforms.

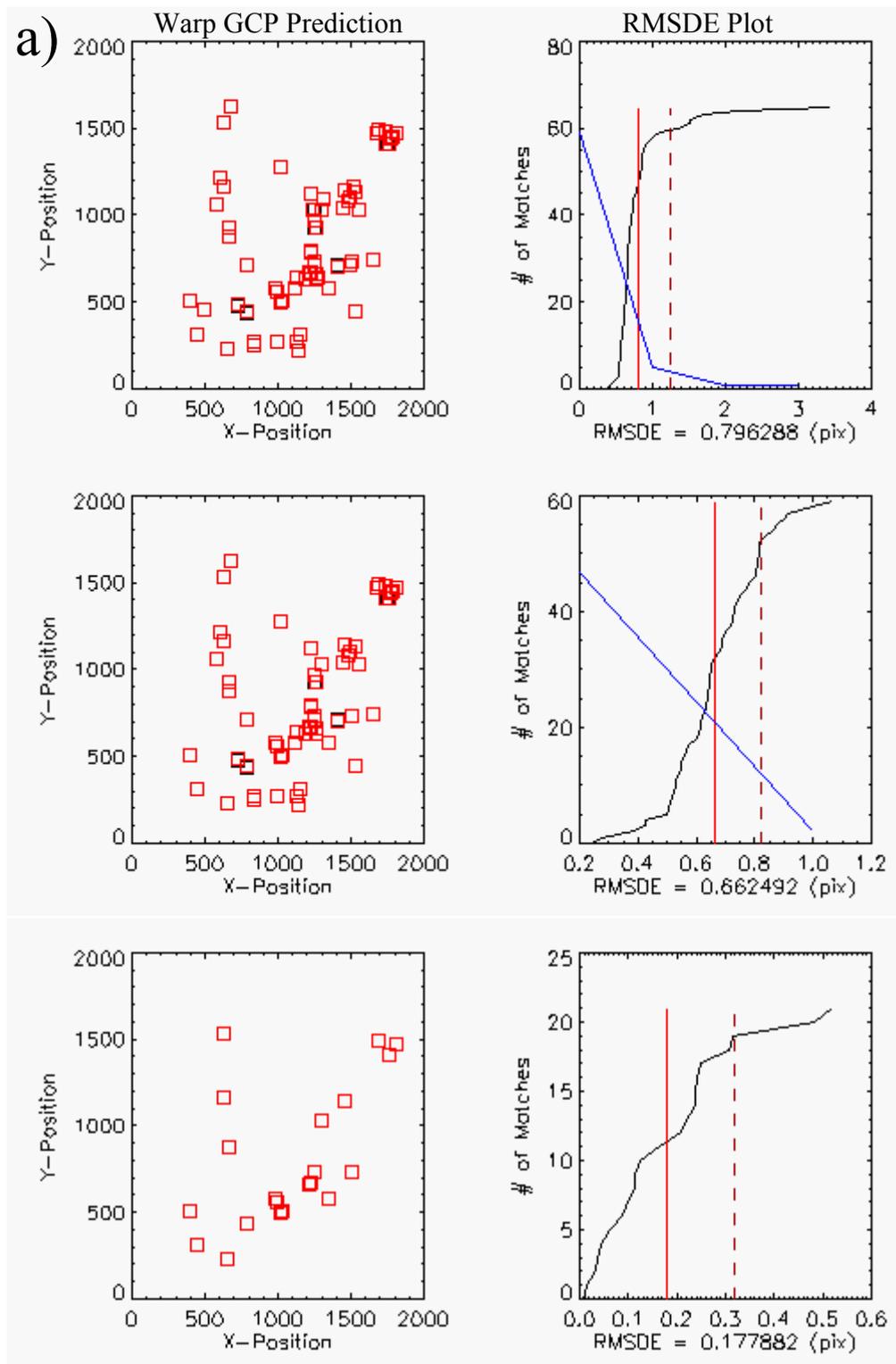


Figure 6.2.8: Utilizing RMSDE analysis to remove matches with greatest error and improve model.

$$a). \quad pred_affine = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0.999694 & 0.000211713 & 0 \\ 0.000458104 & 0.999943 & 0 \\ -1.35829 & 2.29618 & 1 \end{bmatrix}$$

$$pred_affine = \begin{bmatrix} 0.999694 & 0.000212 & 0 \\ 0.000458 & 0.999943 & 0 \\ -5.43316 & 9.18472 & 4 \end{bmatrix} \quad \mathbf{RMSDE=0.57878}$$

$$b). \quad pred_affine = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0.999747 & 0.000305711 & 0 \\ 0.000414231 & 1.00014 & 0 \\ -1.45117 & 2.00238 & 1 \end{bmatrix}$$

$$pred_affine = \begin{bmatrix} 0.999747 & 0.000306 & 0 \\ 0.000414 & 1.00014 & 0 \\ -5.80468 & 8.00952 & 4 \end{bmatrix} \quad \mathbf{RMSDE=0.98062}$$

$$c). \quad pred_affine = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1.00039 & 0.000466588 & 0 \\ 0.000289899 & 1.00047 & 0 \\ -1.85605 & 1.31172 & 1 \end{bmatrix}$$

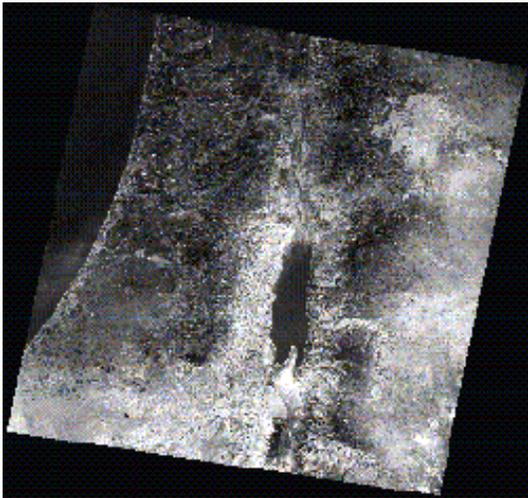
$$pred_affine = \begin{bmatrix} 1.00036 & 0.000467 & 0 \\ 0.00029 & 1.00047 & 0 \\ -7.4242 & 5.24688 & 4 \end{bmatrix} \quad \mathbf{RMSDE=1.55218}$$

$$original_affine = \begin{bmatrix} 0.999990 & 0.0000177 & 0 \\ -0.00000359 & 0.999992 & 0 \\ -6.36321 & 9.41334 & 4 \end{bmatrix}$$

Figure 6.2.9: Predicting the affine model by inducing a scale modifier; the RMSDE tests a pixel at the origin for each model in figure 6.2.8 (a, b,& c) compared to the results of those obtained at the original resolution.

From the preceding results it would seem that predicting a good transform relies more heavily on the number of matches than on the precision of the matches used to derive the affine model. It would at first seem realistic to presume that the 0.7 pixel RMSDE that was attained at the original resolution could best be predicted by a transform model at $\frac{1}{4}$ the resolution if it has 4 times the precision (i.e. RMSDE = 0.2 pix.). However this does not seem to be the case. This test seems to suggest that the predictive model that best describes the transform for the original resolution may be the one that retains as many matches as possible (excluding outliers) and may not always be the one with the greatest precision. Visual results seemed quite good when utilizing the model from 6.2.9.a. The original resolution overlaid and AMV images can be viewed in figure 6.2.10.

Overlaid Image (7747x7291)



AMV Image = 12.3019%

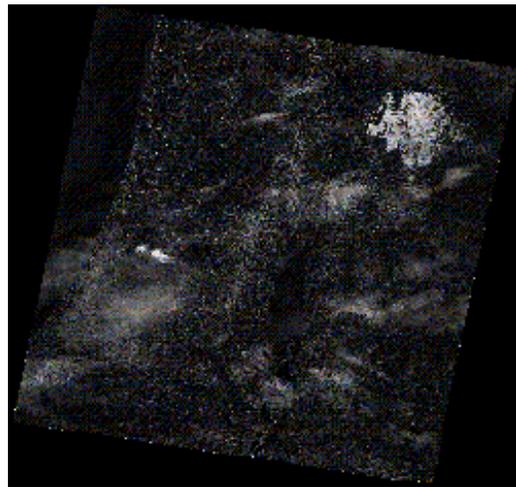


Figure 6.2.10 Confirming the registration results with visual inspection of overlaid and AMV images.

Multiresolution Agility: An added benefit of utilizing the affine model for predictive transforms, especially for multiresolution datasets like LANDSAT, is the ability to develop one transform model and utilize it for all of the image elements. This means that an affine model can be developed with the high-resolution panchromatic band and that same model, can be utilized to predict the transforms for the MS bands and the thermal band and vice-versa (figure 6.2.11). This can take much frustration out of registering an entire LANDSAT dataset to another image and now all elements of the warped dataset have the same relationship which is codified into a 3x3 matrix.

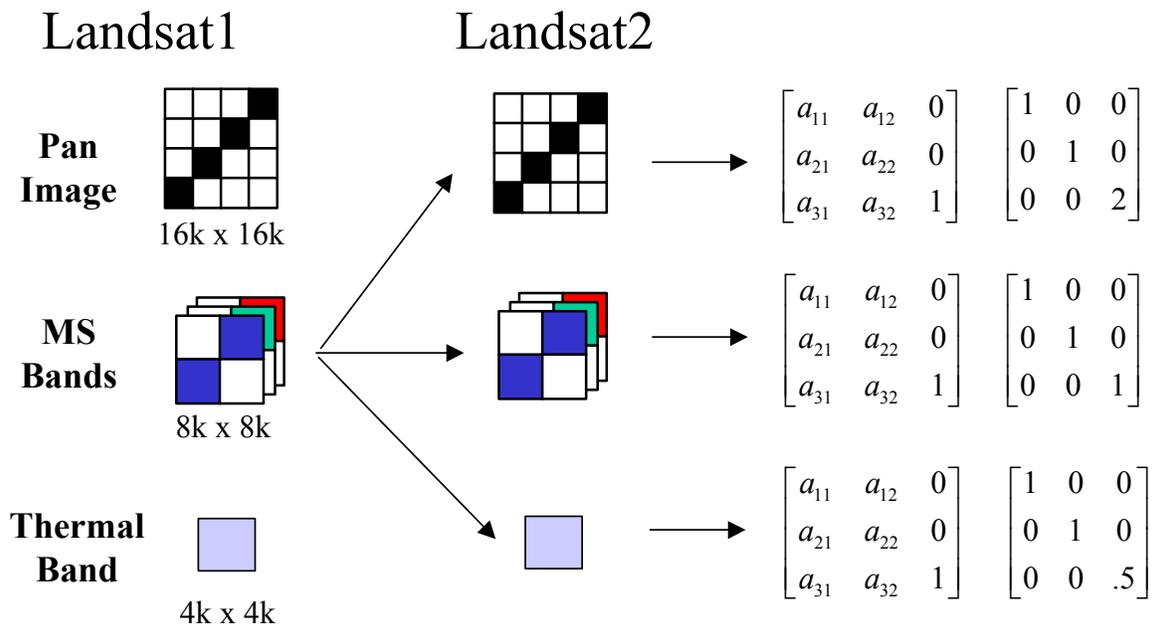


Figure 6.2.10 Affine models demonstrate great agility in multiscale transformations.

6.3 Multisensor/Multiresolution Registration with CITIPIX and HYMAP.

The Mobile, Alabama harbor area site, which contains both CITIPIX RGB (6 inch-HM_2001_USA_1022_0344) and HYMAP HS data (3 meter-hy20010511F01R12S00), collected over the same region. The CITIPIX ROI was extracted at twenty times the resolution of the HYMAP ROI.

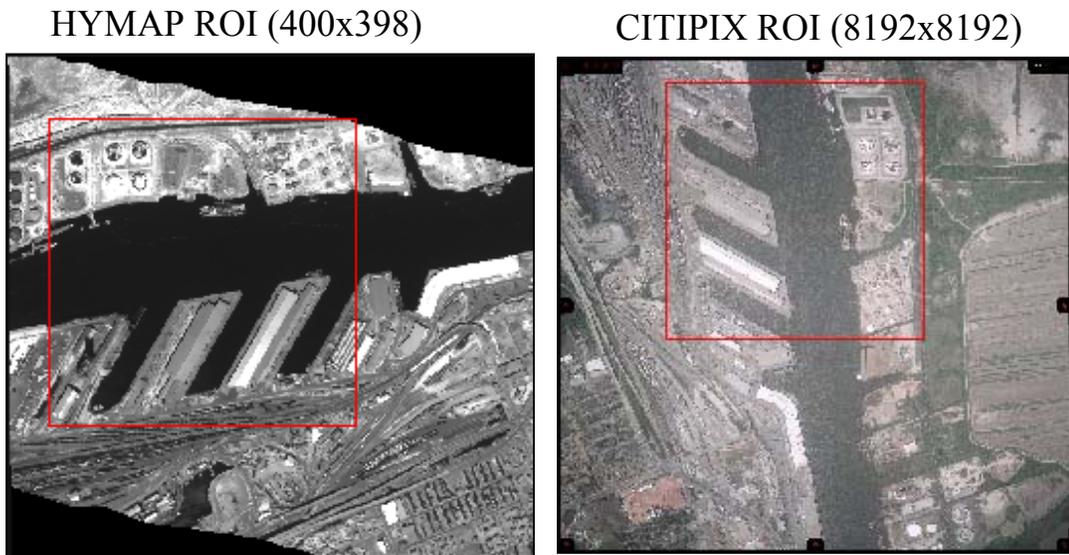


Figure 6.3.1: ROIs taken from the HYMAP HS dataset (band 0.6491) and CITIPIX RGB.

Note that ROIs have been extracted with similar image features (figure 6.3.1). The dyadic size of the CITIPIX image will facilitate wavelet sharpening in Section 7.1, once the datasets have been registered. In order to preprocess the datasets, the CITIPIX image will be downsampled to 1/20th its native size and rotated 270° (figure 6.3.2). The LoGWaR software allows for this supervised manipulation and logs image changes with affine transforms that can later be automatically combined into a composite transform by the program (figure 6.3.3).

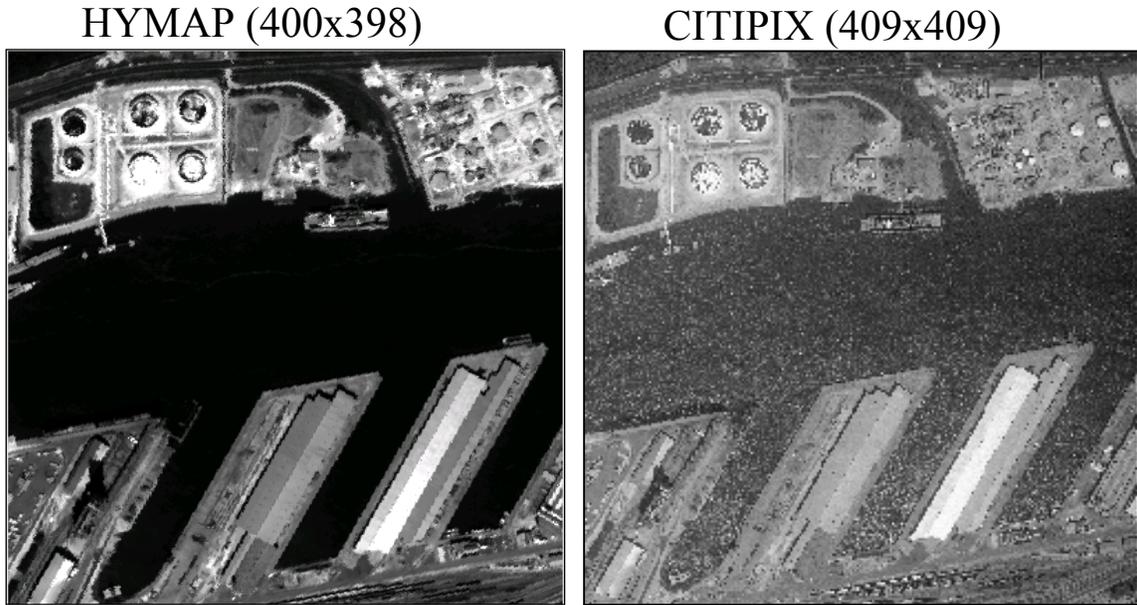


Figure 6.3.2: The scaled and rotated CITIPIX image is now prepared for subregion analysis.

Composite Transform Matrix:			
1.00000	0.00000	0.00000	
0.00000	1.00000	0.00000	Current
0.00000	0.00000	1.00000	Affine
0.00000	-1.00000	0.00000	
1.00000	0.00000	0.00000	Rotation
0.00000	8191.00	1.00000	
1.00000	0.00000	0.00000	
0.00000	1.00000	0.00000	Scale
0.00000	0.00000	20.00000	

Figure 6.3.3: Storing image manipulations and current affine transformation into Composite Matrix.

Now that the images have an orientation relationship, subregion analysis can be utilized to extract match points. Since the images have different characteristics, due to their spectral bands, sensor format (line scanner vs framing), and resolution, the LoGwaR program was optimized via the GUI interface. Some of the user defined options, located at the bottom of the LoGwaR GUI (figure 6.3.4), can be changed to optimize the extraction of GCPs. For this test, the Maxima Similarity button was unchecked because of the potential for very large differences in the LoG image values. Also, since the images have a good orientation relationship, the Match Distance button was checked to perform localized statistical analysis on the potential matched points .

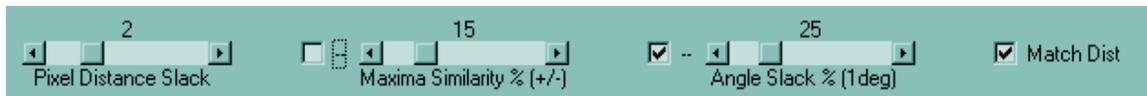


Figure 6.3.4: LoGwaR’s interface options for point matching, located on bottom of GUI.

Another LoGwaR capability, demonstrated below, is the subregion analysis that can be accomplished within a ROI (figure 6.3.5). This technique simply ‘walks’ from the lower left corner, across rows and up columns, using the user defined subregion size (205pix), staying within the prescribed ROI. This technique is useful when the images cover dissimilar features and cropping is not desired. It can be seen that when the Match Distance button is utilized, very few anomalous matches are extracted. Caution must be utilized since images cannot have rotational dissimilarity for it to operate effectively!

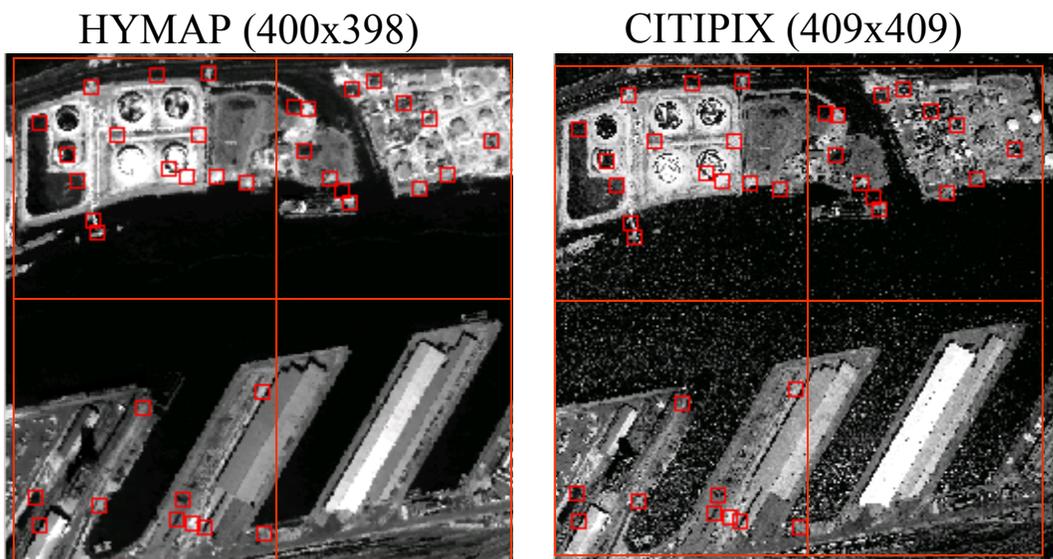


Figure 6.3.5: Matches resulting from LoGwaR’s subregion analysis technique (205x205area).

Since the lower left corner did not produce any matches with this configuration (using the subregion technique), a new ROI was created for focused extraction of GCPs. With the addition of matches from this region (figure 6.3.6), a desirable quantity and distribution of GCPs have been produced for global statistical analysis to commence.

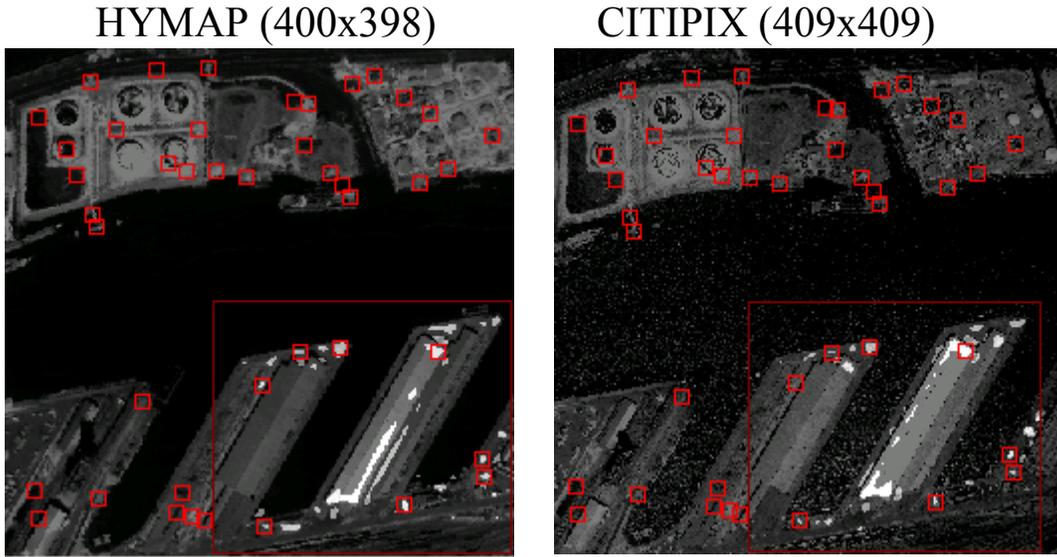


Figure 6.3.6: An additional ROI added due to sparse content in lower left corner.

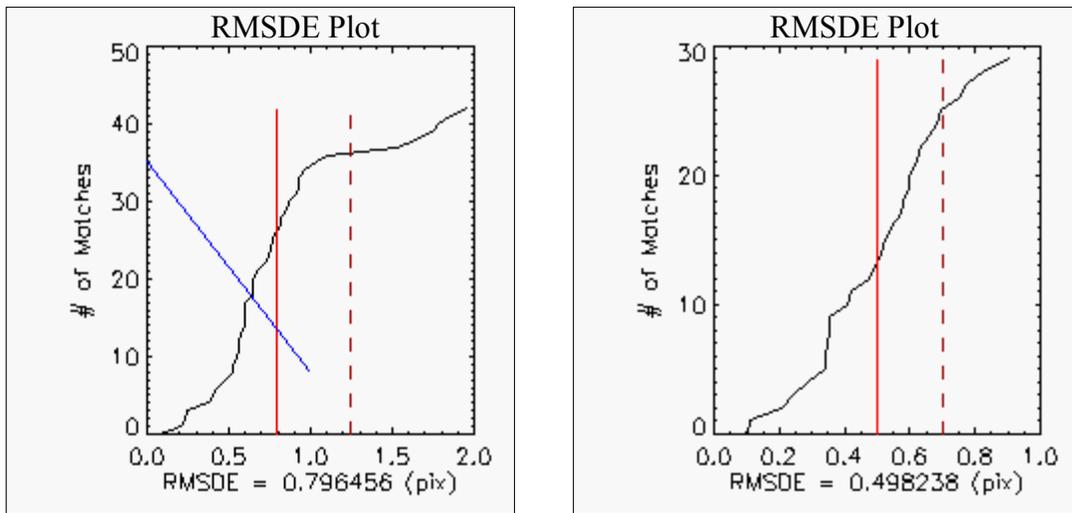


Figure 6.3.7: Utilizing RMSDE analysis to judge the accuracy of the registration model.

As can be seen from the RMSDE analysis (figure 6.3.7), the matches derived with the subregion technique and localized statistics enabled, has initially achieved subpixel accuracy. It should be noted that all of the initial matches are within the Distance Match specification (2 pixels error) that is used to determine potential matches. Therefore, the

features that have been matched are probably related but some may deviate from the model due to imprecise GCPs because of quantization, sampling, parallax, elevation, etc. In the previous examples, bad matches were iteratively rejected from the model if they deviated from the mean by more than one standard deviation. For this example, an alternative technique will be demonstrated.

Since LoGWaR allows the user to enter the precision of the model desired, it is possible to iteratively remove the match with the largest error (for the given polynomial degree) and then recomputed the model based on the remaining matches. This process continues until either the desired precision is reached or there are not enough points left to transform the image at the desire polynomial degree. The following list displays the resulting RMSDE Means generated for this iterative process. This technique is useful to achieve the registration accuracy needed (user defined in this example to be less than half a pixel) for predictive warping, since the error increases with the scale of the prediction:

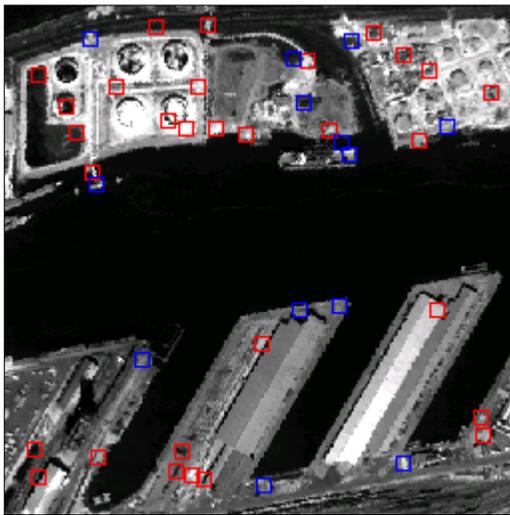
```

Number of Matches Loaded =      43
RMSDE Mean = 0.796456
Degree of Polynomial used Test = 1
RMSDE Mean = 0.740595
RMSDE Mean = 0.712932
RMSDE Mean = 0.691343
RMSDE Mean = 0.664629
RMSDE Mean = 0.628197
RMSDE Mean = 0.604795
RMSDE Mean = 0.589627
RMSDE Mean = 0.575741
RMSDE Mean = 0.558026
RMSDE Mean = 0.546082
RMSDE Mean = 0.530231
RMSDE Mean = 0.515644
RMSDE Mean = 0.498238
.....
RMSDE SDev = 0.203998
# of Good Matches =      30

```

Here the number of matches went from 43 to 30, whereas the RMSDE went from ~0.8 pixels to below 0.5 pixels, respectively. This technique is slightly different in that it allows the user to uniquely identify the precision of the transform model desired and can allow completely autonomous statistical analysis of the match points.

HYMAP (400x398)



CITIPIX (409x409)

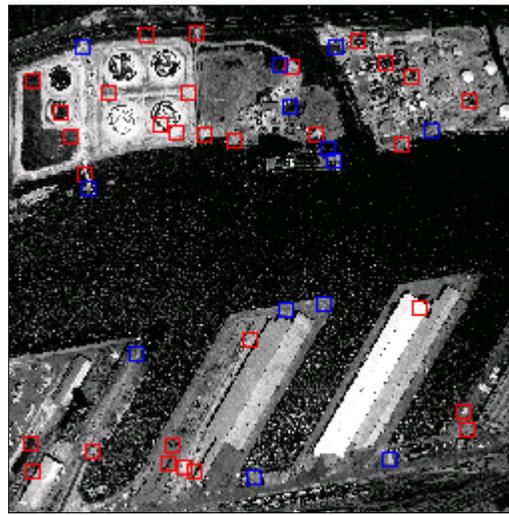
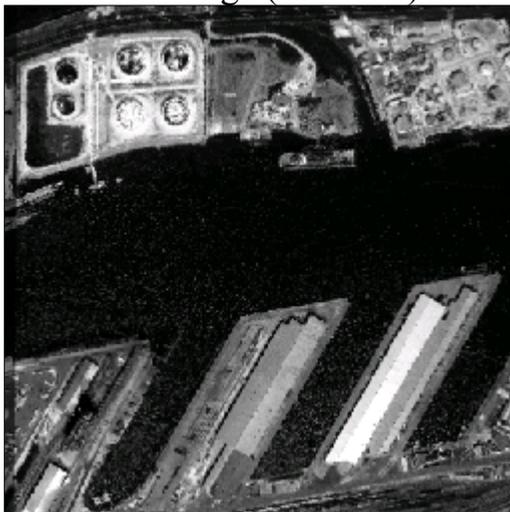


Figure 6.3.8: Point Matches remaining from LoGWaR's iterative statistical analysis of RMSDE <5 pix

The final match locations are displayed on the images in figure 6.3.8, where the blue icons represent the matches removed from the model. The resulting transform can be visually analyzed via the overlaid and AMV (9.55713%) images in figure 6.3.9.

Overlaid Image (400x398)



AMV Image = 9.55713%



Figure 6.3.9 Confirming the registration results with visual inspection of overlaid and AMV images.

Although the AMV seems relatively high compared to previous 'good' registrations, it should be recalled that the datasets are quite dissimilar in grayscale...especially in the water regions (even though a histogram matching operation

was performed for registration). These results will be utilized in section 7.1 to illustrate the ease of implementing wavelet sharpening once the images are well registered. This transform model can now be utilized to warp the original CITIPIX image chip at full resolution (8192x8192) if the affine chain of manipulations has been maintained for the composite transform.

In order to perform the transform on the original image, it is necessary to update the composite transform discussed in figure 6.3.4. To do this, the latest affine transform (developed at the reduced resolution) must be incorporated into the composite matrix.. Since the image rotation, scale reduction in the wavelet domain, and registration transform have all been accomplished in an affine format, they can be combined with the commutative technique discussed in Chapter 5. This process is automatically logged within the LoGWaR program (except for the pull-down menu update of the current affine) and is displayed below:

Composite Transform Matrix:			
0.996564	-0.00121325	0.000000	
-0.00476494	0.987933	0.000000	Current
-5.64522	7.10843	1.00000	Affine
0.000000	-1.00000	0.000000	
1.00000	0.000000	0.000000	Rotation
0.000000	8191.00	1.00000	
1.00000	0.000000	0.000000	
0.000000	1.00000	0.000000	Scale
0.000000	0.000000	20.00000	

Figure 6.3.10 LoGWaR automatically keeps track of image manipulations with affine matrices.

First the scale factor, from wavelet decimation or downsampling, is accounted for through commutation and then the rotation. It should be noted that the order of this operation is important, due to any wavelet padding that was incorporated to obtain dyadic dimensions. The order of the matrix commutations must be in the inverse order as that applied in the LoGWaR operations to obtain the correct results. The composite affine results will be utilized to transform the original scale (8k x 8k) image (figure 6.3.11).

Once this is accomplished, the HYMAP image can be upsampled to 20 times its original dimensions, utilizing Nearest Neighbor resampling to maintain spectral integrity. Now the images are spatially related for comparison or additional processing.

Alternatively, the warped CITIPIX image could be decimated with the FWT to 512x512 and the HYMAP upsampled to 512x512 to allow for wavelet sharpening. This would entail transferring the high frequency information from CITIPIX to HYMAP and performing the inverse FWT to provide up to four levels of sharpening (reference Chapter 7 and Appendix A for more detail on wavelet sharpening).

Figure 6.3.11 The Composite Transform that can be utilized to warp the original image.

$$\begin{array}{l}
 \begin{array}{ccc}
 \textit{Rotation Affine} & \textit{Latest Affine} & \textit{Scale Affine} \\
 \textit{Comp_Trans} = \begin{bmatrix} 0.0000 & -1.0000 & 0.0000 \\ 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 8191.00 & 1.0000 \end{bmatrix} & \begin{bmatrix} 0.996564 & -0.00121325 & 0.0000 \\ -0.00476494 & 0.987933 & 0.0000 \\ -5.64522 & 7.10843 & 1.0000 \end{bmatrix} & \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 20.0000 \end{bmatrix} \\
 \textit{Comp_Trans} = \begin{bmatrix} 0.0000 & -1.0000 & 0.0000 \\ 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 8191.00 & 1.0000 \end{bmatrix} & \begin{bmatrix} 0.996564 & -0.00121325 & 0.0000 \\ -0.00476494 & 0.987933 & 0.0000 \\ -112.904 & 142.169 & 1.0000 \end{bmatrix} & \\
 \textit{Comp_Trans} = \begin{bmatrix} -0.00121325 & -0.996564 & 0.0000 \\ 0.987933 & 0.00476494 & 0.0000 \\ 142.169 & 8303.90 & 1.0000 \end{bmatrix} & &
 \end{array}
 \end{array}$$

The ability to automatically register images with LoGWaR is dependent on isolating similar edge detail within the datasets. This capability is heavily dependent on analyzing the images as similar scale/resolution. The multiresolution aspects of this research allow comparison of data at similar scale for the extraction of related spatial frequency detail. The rationale for automatic analysis of data at similar resolutions can be noted from the images in figure 6.3.12. Here the original CITIPIX is downsampled from 20x to 5x the resolution of the HYMAP data. Even at this reduced scale it is easy to see the how automated detail comparison between the two images would be difficult. However, once the images have been scaled to the same resolution, the similarity in edge detail becomes apparent.

For this reason, registration at similar resolutions is critical for automatic correlation of edges between datasets. Furthermore, the need for predictive transforms,

CITIPIX Chip (Downsampled to 5x)



CITIPIX Chip (downsampled to 1x)



HYMAP Chip (original scale)

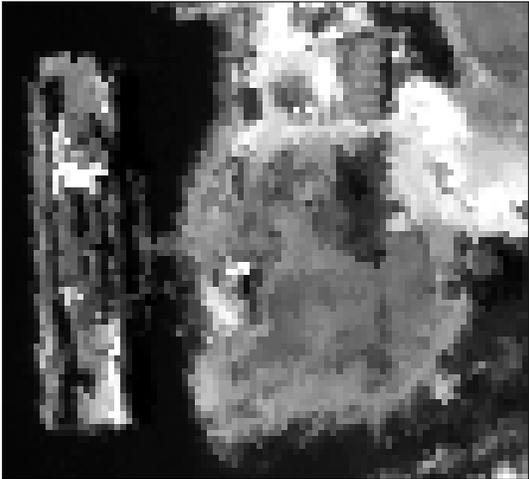


Figure 6.3.11: Comparison of CITIPIX and HYMAP data at 5x and 1x resolutions.

accomplished at lower resolutions, becomes necessary when attempting to automatically relate multiresolution datasets. This approach requires incorporation of the affine composite transform for relating the original hi-resolution image to the low-resolution image space. All of these techniques have been incorporated into the LoGWaR program, with the added benefit of automatic logging and computation of the affine matrices necessary to perform predictive transformations.

ADDITIONAL APPLICATIONS

As with any research, there is often potential for application beyond the scope of the “core” research area. This section is dedicated to highlighting some of the applications that could gain immediate benefit from the registration research presented thus far. Several analytical techniques, such as image sharpening, stacking, mosaicing, fusion, and change detection can now be applied since the datasets have been spatially related. An added benefit of the wavelet based approach for registration, is the potential for wavelet based sharpening the low-resolution spectral image plane with the high frequency detail derived from the high-resolution data. In conjunction with the sharpened product, there is now potential for pure endmember selection for spectral unmixing algorithms. Some of these applications will be now be discussed.

7.1 Wavelet Sharpening - High Frequency Detail transfer w/FWT⁻¹

In continuation of the process flow developed for the LoGWaR technique, figure 7.1.1 includes the additional steps required for Spatial Sharpening of images in the Wavelet domain. Due to competing resources (photons), remote sensing system can either provide high-resolution spatial information or detailed spectra, but not both simultaneously. By utilizing the process above (figure 7.1), it should be possible to transfer the image detail from high-resolution images to low-resolution images. A recent SPIE paper seems to corroborate this idea, "if the data are taken nearly at the same time, some cross-sensor resolution enhancement techniques are able to produce a merged image as close as possible to what would be a high spatial resolution hyperspectral image...Multiresolution Wavelet Decomposition is the most interesting tool to perform this process." (Peytavin 1996)

When utilizing wavelet analysis, special considerations must be undertaken to maintain the proper scale relationship. The dyadic requirement of the FWT mandates that these similar regions must be pixels related by a ‘power of 2’ in both the ‘x’ and ‘y’ dimension. Since the

largest common dyadic size image chip in figure 4.1.1 is 2048x2048, these areas will be utilized for registration. It is also possible to ‘zero-pad’ the image to obtain a dyadic size, but, caution must be introduced here if combined with image rotation. The reasons for this were explained in Chapter 5, when it becomes necessary to keep track of image manipulations with the composite transform. The figure below demonstrates this dyadic common area requirement:

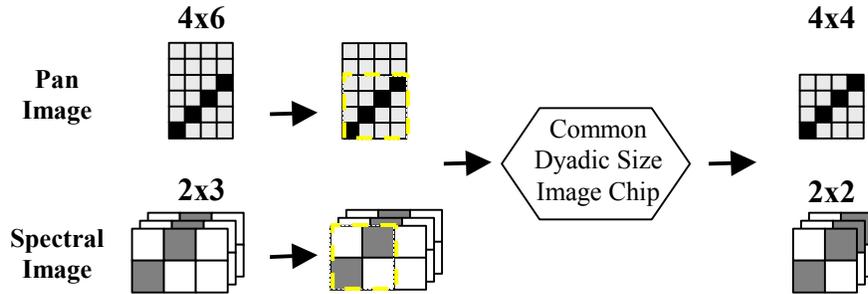


Figure 7.1.1: Extraction of common areas with dyadic dimensions.

Since it is often useful to relate the process flow with real data, a test dataset can be viewed in figure 7.1.1. The high-resolution (hi-res) image, which will be utilized for this test case is the CITIPIX (6 inch) image of Mobile, AL. The corollary hyperspectral (HS) data is a HYMAP (1 m) image that was taken of the same region. When utilizing the FWT, some unique requirements for specific image dimensions and relationships between images are required and so will require additional preparation (figure 7.1.2).

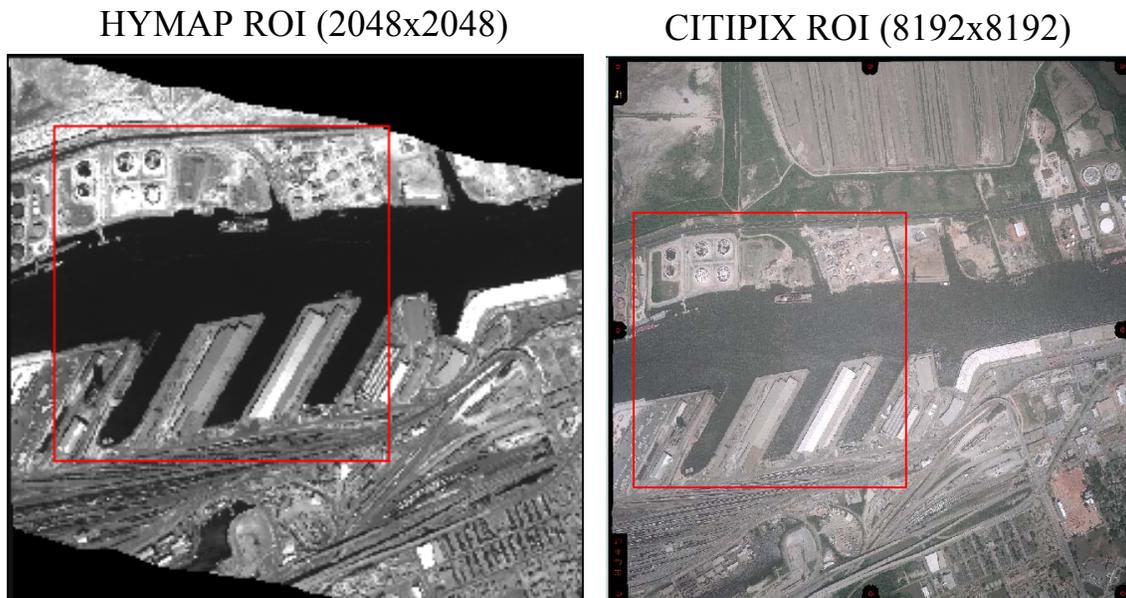


Figure 7.1.2: Mobile dataset with dyadic image chips relating CITIPIX to HYMAP.

Now the hi-res image is resampled to the closest dyadic factor of the spectral image size. For our test case using the CITIPIX (6 inch) RGB image and the HYMAP (3 meter) spectral cube, it is necessary to resample the CITIPIX data to the nearest power of two (2^n) resolution. Since there are approximately 118 inches in 3 meters, then we could decompose 118 into its dyadic constituents (59, 29.5, 14.8, 7.4, & 3.7) and find the closest to 6 inches. Since 7.4 is the closest dyadic constituent, it is necessary to resample 6 inch pixels to 7.4 inch pixels. This will require the CITIPIX image to be processed through the FWT using four iterations. The resulting “scale” image, $\varphi_4(x, y)$, will then be 16x smaller (2^4) and contain a “detail” plane, $\Psi_4(x, y)$, containing the 4 highest frequency bands that have been stripped off.

An alternative approach is to determine the scale difference between the two datasets:

$$(7.1) \text{ SCALE} = (\text{HYMAP Resolution} / \text{CITIPIX Resolution}) = (118/6) = 19.667$$

This indicates that the CITIPIX image is 19.667 times the resolution of the HYMAP dataset. Since the FWT requires a dyadic scale relationship, it is reasonable to utilize the nearest lower power of two. So, we can easily compute this value in IDL with the following formula:

$$(7.2) \text{ DYADIC_Power} = \text{FLOOR}(\text{alog}(\text{SCALE})/\text{alog}(2)) = 4$$

$$(7.3) \text{ DYADIC_Scale} = 2^{\text{DYADIC_Power}} = 2^4 = 16$$

In this example we could warp the CITIPIX image to 0.8136 scale ($16 / 19.667$) to attain the required dyadic relationship. So, the CITIPIX image can now be related to the HYMAP image after 4 decimations of the FWT.

Since the CITIPIX imagery is 3-band RGB, the spectra can tolerate modification through interpolation with minimal impact. So, either bilinear or bicubic interpolation can be utilized for resampling the data since it provides “smoother” results and will provide crisper edges for sharpening. In the case of the multispectral registration to hyperspectral, nearest neighbor resampling would be utilized when it is important to maintain the spectra.

HYMAP ROI (2048x2048)



CITIPIX ROI (8192x8192)



Figure 7.1.3: Dyadic ROI from HYMAP and CITIPIX that were registered in Section 6.3.

Once the wavelet decimated CITIPIX image (the scale plane $\varphi(x, y)$) has been registered with the appropriate spectral band of the HYMAP dataset (figure 7.1.3), it is possible to transfer the high frequency detail ($\Psi(x, y)$) that has been iteratively stripped off utilizing the FWT. This unique application involves transferring the detail plane of the CITIPIX image over to the registered spectral band of HYMAP. Now that the “scale plane” of the warp image is of the same scale, dimension, and orientation of the low-res HS image, it is possible to perform the FWT^{-1} on the “spectral band/warp detail plane” combination (figure 7.1.4). Recall that this combination is both the warped scale ($\varphi(x, y)$) and the warped detail plane ($\Psi(x, y)$) now combined into a wavelet decimated structure. It is essential to ensure the dyadic image size and scale plane relationship is maintained prior to the transfer of detail information. In this way we can elegantly add detail at increasingly greater dyadic frequencies for band sharpening analysis.

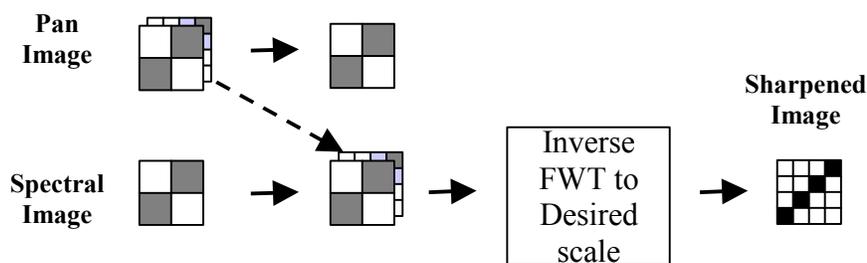


Figure 7.1.4: Transferring “Detail Plane” from a Pan to a Spectral image before FWT^{-1} .

Since the 8kx8k CITIPIX image is too large to demonstrate the sharpening results, another subregion will be utilized for analysis. Note that the sharpening can now occur over the entire 8kx8k area or on any individual subregion of interest. The following regions will be utilized for demonstrating the wavelet sharpening technique (figure 7.1.5).

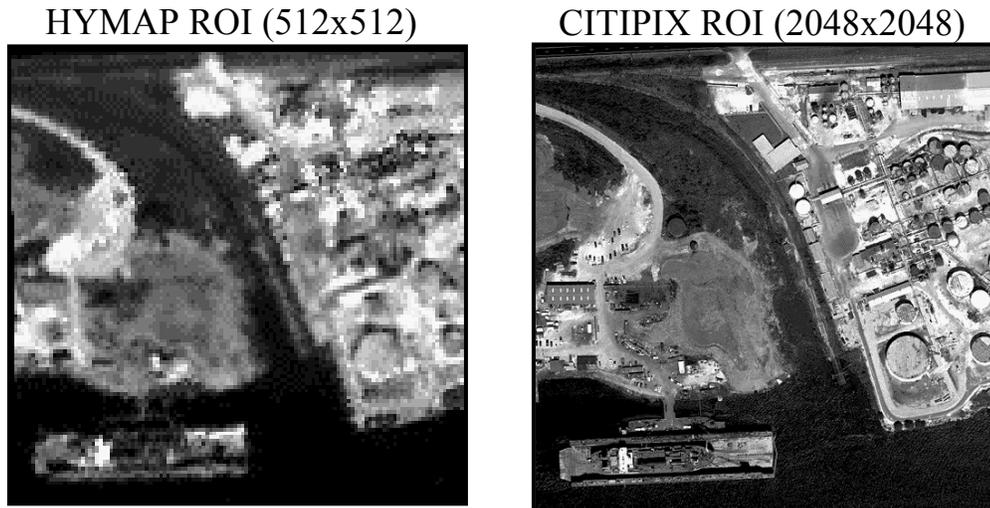


Figure 7.1.5: Dyadic sub-ROI from HYMAP and CITIPIX that will be utilized for sharpening.

Now the CITIPIX image is decimated with two iterations of the FWT to produce a scale band ($\varphi(x, y)$) with the same dimensions as the HYMAP image (512x512). This process can be viewed using Mallat's Representation, which separates the detail planes into the horizontal, vertical, and diagonal components (figure 7.1.6).

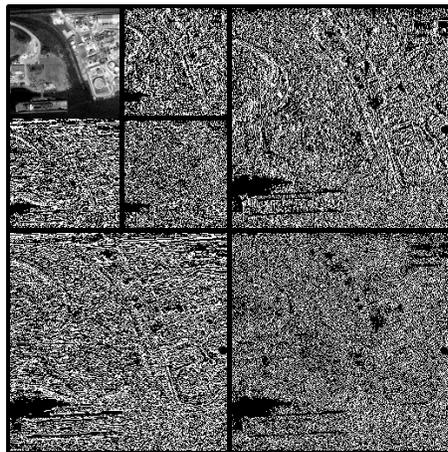


Figure 7.1.6: Two iterations of the FWT, utilized to decimation the CITIPIX image.

Now the CITIPIX subband and the HYMAP spectral band regions are at the same resolution and dimension. It is now possible for the transfer of high frequency spatial detail from the CITIPIX image ($\Psi(x, y)$) to HYMAP. This is physically accomplished by trading the CITIPIX subband with the HYMAP spectral band into the wavelet matrix (figure 7.1.7).

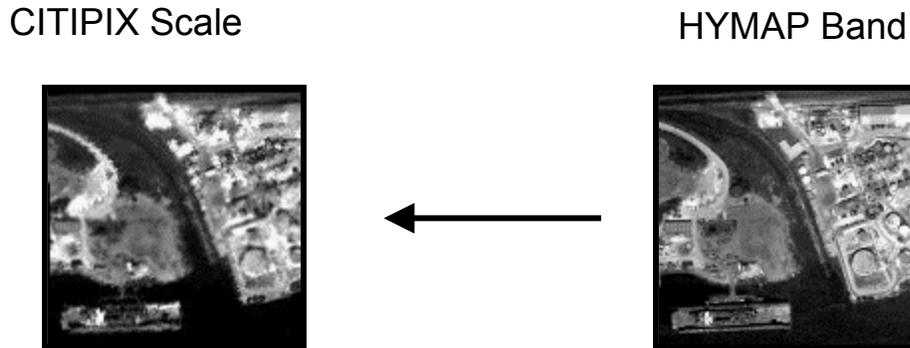


Figure 7.1.7: Switch the CITIPIX scale plane with the HYMAP band in the wavelet matrix.

With the HYMAP band now imbedded within the high frequency content of the CITIPIX wavelet matrix (figure 7.1.8) it is possible to perform an FWT^{-1} to reconstitute the detail into the HYMAP image. At each inverse FWT level, the overall radiometry of the HYMAP image is preserved in the sharpened product. This is especially useful for any additional processing such as spectral unmixing.

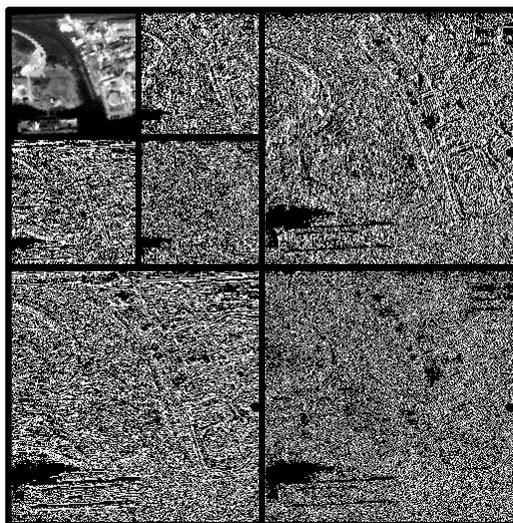


Figure 7.1.8: HYMAP band embedded within the CITIPIX detail planes.

The results of the FWT^{-1} can be viewed at the same dimensions to better illustrate the edge enhancement capabilities of the wavelet sharpening. The following four images, figure 7.1.9, represent the HYMAP band during three stages of the sharpening process (A-C) and the original CITIPIX image for comparison (D).

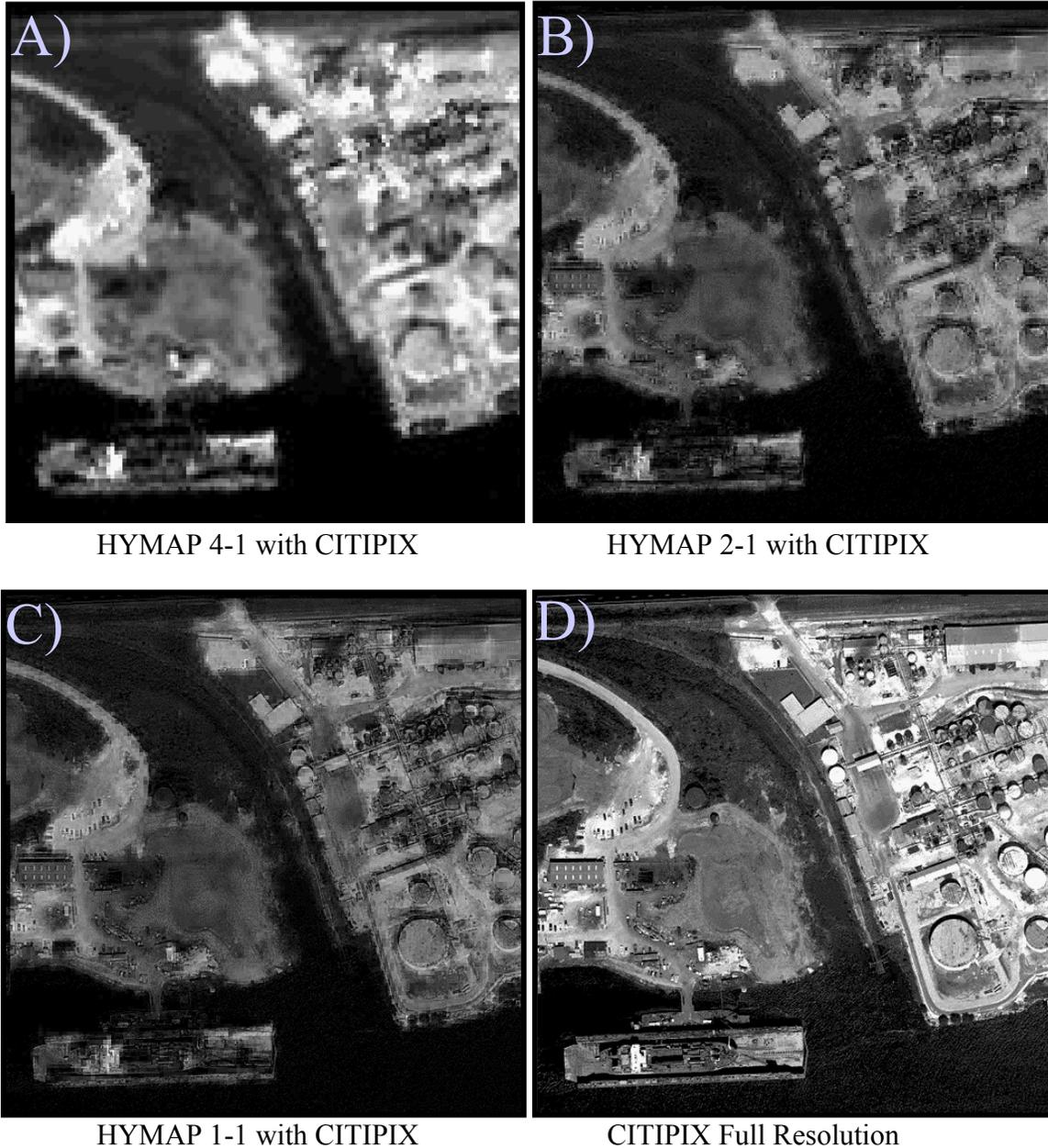


Figure 7.1.9: Results of the HYMAP sharpening (A-C) compared to the original CITIPIX area (D).

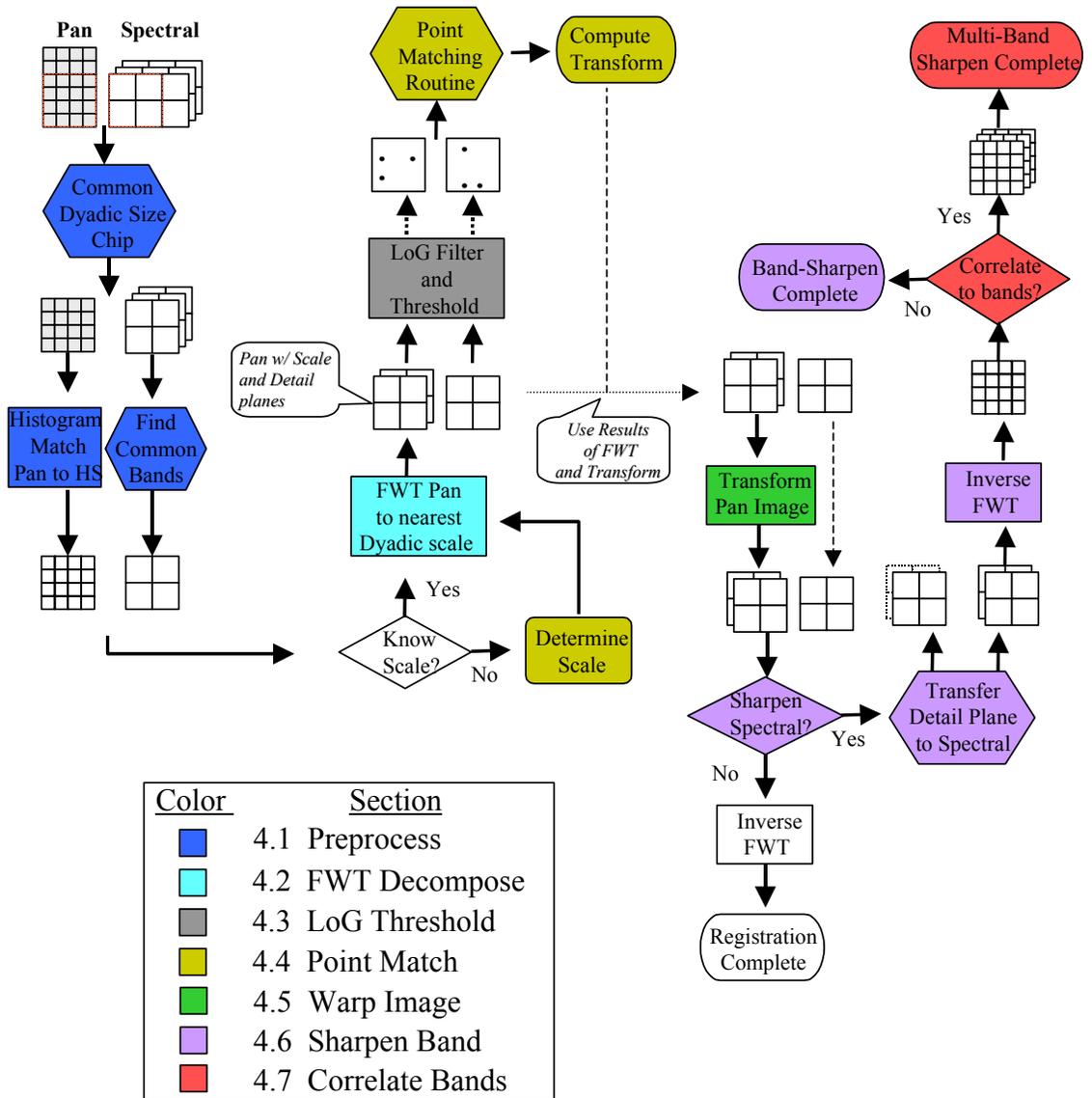
Since the FWT utilizes the Haar wavelet, some edge artifacts are evident in the sharpened product. This is due, in part, to the “blocky” nature of the Haar ‘mother wavelet’. However, there are many benefits in utilizing the Haar, compared to most other wavelets. First, the ability to maintain overall radiometric integrity (Appendix A) allows for correct application of many spectral algorithms such as unmixing. Also, since the FWT resampling is lossless, the data can be completely recovered at any time and each level maintains the same overall grayscale values as the original image. Additionally the speed of the FWT makes this technique very attractive when processing large images.

Also, it is important to note the artifacts that will be introduced as a result of relief displacement (parallax effects) as they may apply to the sharpened product. Any relief displacement (or its effects such as shadowing) between the datasets will hamper attempts at good sharpening. Because of the detrimental effects that this may have on the sharpened products, it is often necessary to obtain datasets with similar viewing geometries to produce satisfactory results. Even though the HYMAP and CITIPIX test case in Figure 7.1.8 had similar viewing geometries, this edge artifact effect can still be noticed on the windowed building above the dry-dock.

Once wavelet sharpening has been accomplished, a direct comparison/analysis of the sharpened HYMAP dataset (at various resolutions) to the original spectral data is possible. This comparison is often worthwhile, since it can add greatly to an analyst’s knowledge of a scene when both the hi-res spatial information and the spectral content can be observed together. An exemplar case, is the ability to compare spatial knowledge of a panchromatic band, spectra of an HS cube, and radiometry from a low-res thermal band in unison.

The overall process flow for sharpening images based on this technique follows in figure 7.1.10. This process has been adapted from the registration flowchart to highlight special requirements of the FWT.

Figure 7.1.10: The Multisensor Image Registration and Sharpening Process



7.2 Pure Pixel selection for Endmember Libraries

“Pure endmember” selection for spectral unmixing often requires the supervised selection of “pure endmember” classes. This is especially true for stepwise unmixing (Gross & Schott 2000). Through the sharpening process highlighted here, it is now possible to utilize the detail from the higher resolution image to identify homogeneous spectral regions within the low-resolution spectral plane. What may appear to be a ‘pure spectra’ in the low-resolution HS band may not in the high-resolution image (figure 7.2.1).

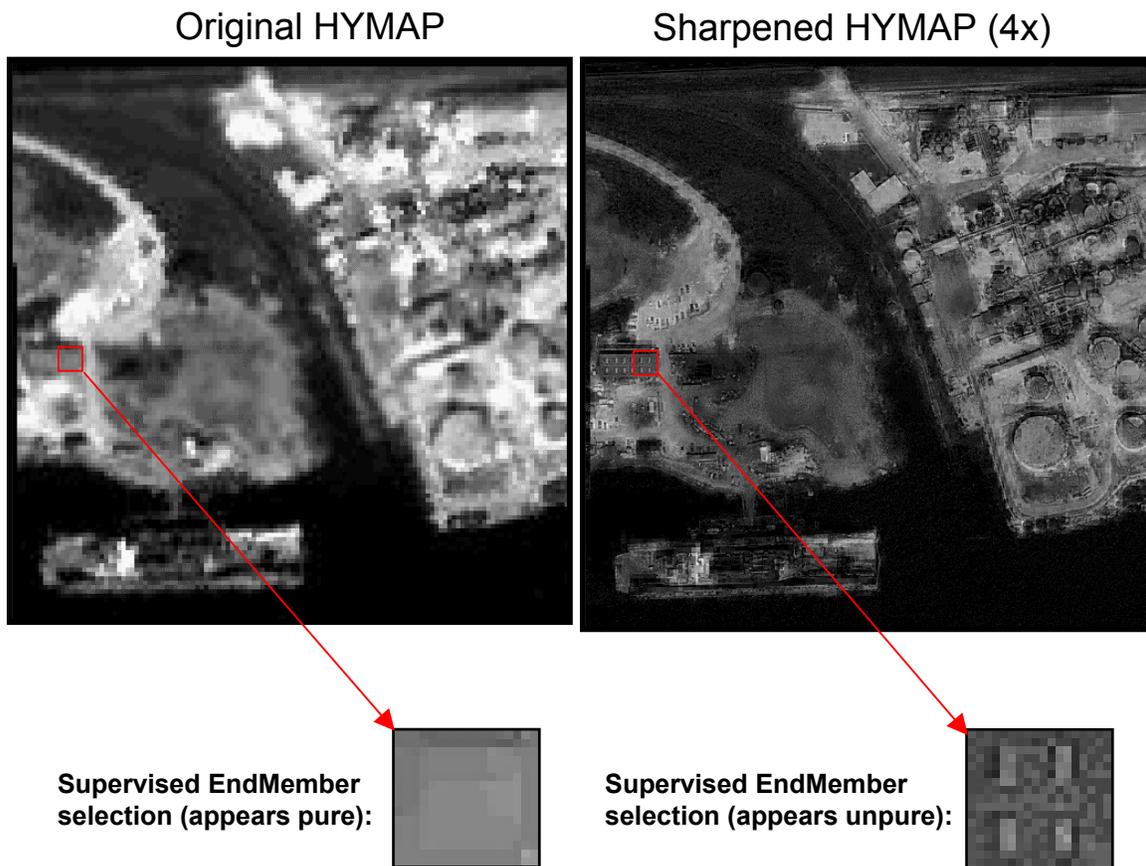


Figure 7.2.1: The ability to select ‘pure’ endmembers is often based on edge detail perception.

The utility of supervised unmixing algorithms to discriminate species, depends heavily on the ability to accurately define “pure endmembers” within a spectral data cube. In fact, without “pure endmembers”, stepwise unmixing loses much of its capability to accurately unmix spectral data (Gross & Schott 2000). Conversely, with “pure endmembers” comprising the endmember library, the algorithm can accurately unmix spectral compositions at the

subpixel level. The ability to incorporate the high-frequency detail directly into low-resolution spectral bands, greatly increases the ability to select pure pixels and provides critical support for algorithms such as stepwise unmixing.

The Following Applications have not been implemented or pursued under this effort. They are only mentioned as potential areas for further research, since they all require high degrees of spatial correlation that can be obtained with the registration techniques developed in this thesis.

7.3 Super-Resolution by combining Spatial Sharpening with Spectral Unmixing

With the ability to ‘sharpen’ images and to ‘spectrally unmix’ end-members at the subpixel level, it should be possible to combine the results of these two operations to create a super-resolution product that incorporates both properties. This hybrid product could be produced from the byproducts of sharpening and unmixing.

First, it would be essential to analyze the subpixel edge-detail (high frequency content) within a sharpened image pixel. Once this is done, the super-pixel would be divided into regions (connected components) based on the edge-detail and texture. The resulting area of each connected component could then be normalized at the pixel level to determine a percent fill factor (fraction map) for each region. Second, the ‘fraction maps’, which normally result from an unmixing operation, would be compared at the super-pixel level to the sharpened product. This comparison would provide the best estimate, as to how a super-pixel would be comprised based on spatial structure and spectral content. (figure 7.3.1).

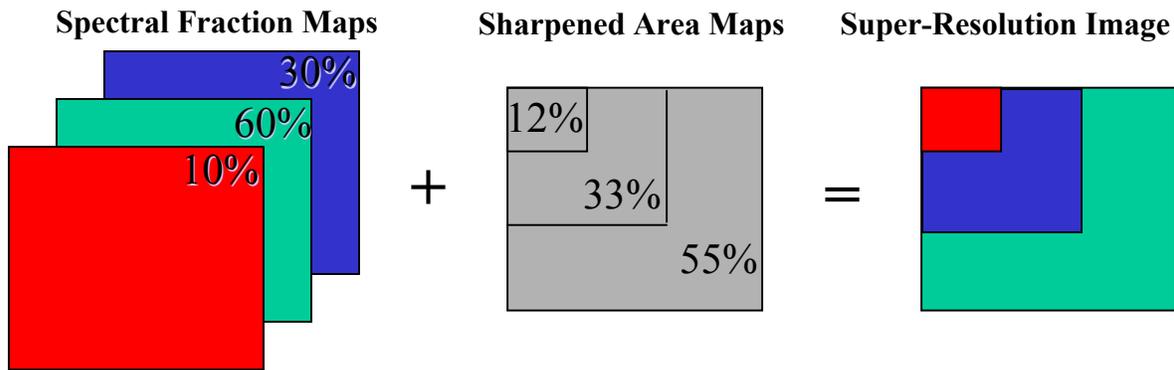


Figure 7.3.1: Combine the results of Spectral Unmixing and Sharpening to create Super-Resolution images.

7.3 Cross-Band Correlation - Detail Transfer to other spectral bands

If there is a high degree of correlation, band-to-band, between the multispectral (warp image) and hyperspectral data (reference image), then it should be possible to interpolate/extrapolate sharpening to remaining planes of the hyperspectral cube. Schott maintains that cross-band correlation can be accomplished, “if we assume that the reflectance in the two bands are approximately correlated with zero bias such that:

$$(7.3.1) \quad r_2 = Cr_1 + \varepsilon$$

where r_1 and r_2 are the reflectance values in band 1 and band 2, C is approximately a constant, and ε is the error due to the lack of perfect correlation between r_1 and r_2 .”(Schott 1997) Here, reflectance would be related to the intensity level (digital count/grayscale) of the spectral plane data.

The ability to transfer detail information to additional spectral planes, beyond the band related image planes of the high-resolution and lower resolution images, has important applications. Propagation of high frequency detail into additional wavelengths would allow additional analysis in those bands of interest. In addition, composite analysis could now be accomplished on the entire spectral cube, at the new resolution.

The ability to sharpen non-correlated bands can also be accomplished if there is a known relationship between bands of interest. For instance, the ability to sharpen a LANDSAT

thermal band with the information from the panchromatic band could be accomplished due to the known relationship between the data even though the end detail may appear quite different. This is often useful to obtain knowledge of how the radiometry of a given area compares to known geography that may be readily apparent in the panchromatic or even MS bands.

Since the LANDSAT bands already have dyadic relationships, this could easily be accomplished by bringing the Pan or MS band into the LoGWaR program as the Reference Image and the Thermal band as the Warp Image. Once these images are resident in LoGWaR, it is straightforward to perform wavelet decimation on the hi-res Pan or MS image from the Wavelet Tools pull-down menu. Since there is a known relationship between the resolutions of these bands (Pan 4x, MS 2x, Thermal 1x), it is possible to perform one or two iterations of the FWT and transfer the detail planes of the hi-res image to the thermal band. Once this is accomplished it is just as easy to perform an inverse FWT on the thermal data to provide a sharpened result.

7.5 Utilizing the AMV (difference) image for Change Detection

With the ability to register two images at the subpixel level, the difference between those two images (AMV) can be readily utilized as a tool to detect change. This 'change detection' can be utilized to detect changes in agricultural crops, environmental concerns (erosion, fire damage, etc), and even moving objects, if the resolution is good enough. In the example of the LANDSAT test case from Section 6.2, the AMV can even be utilized to determine cloud-cover changes (figure 7.5.1).

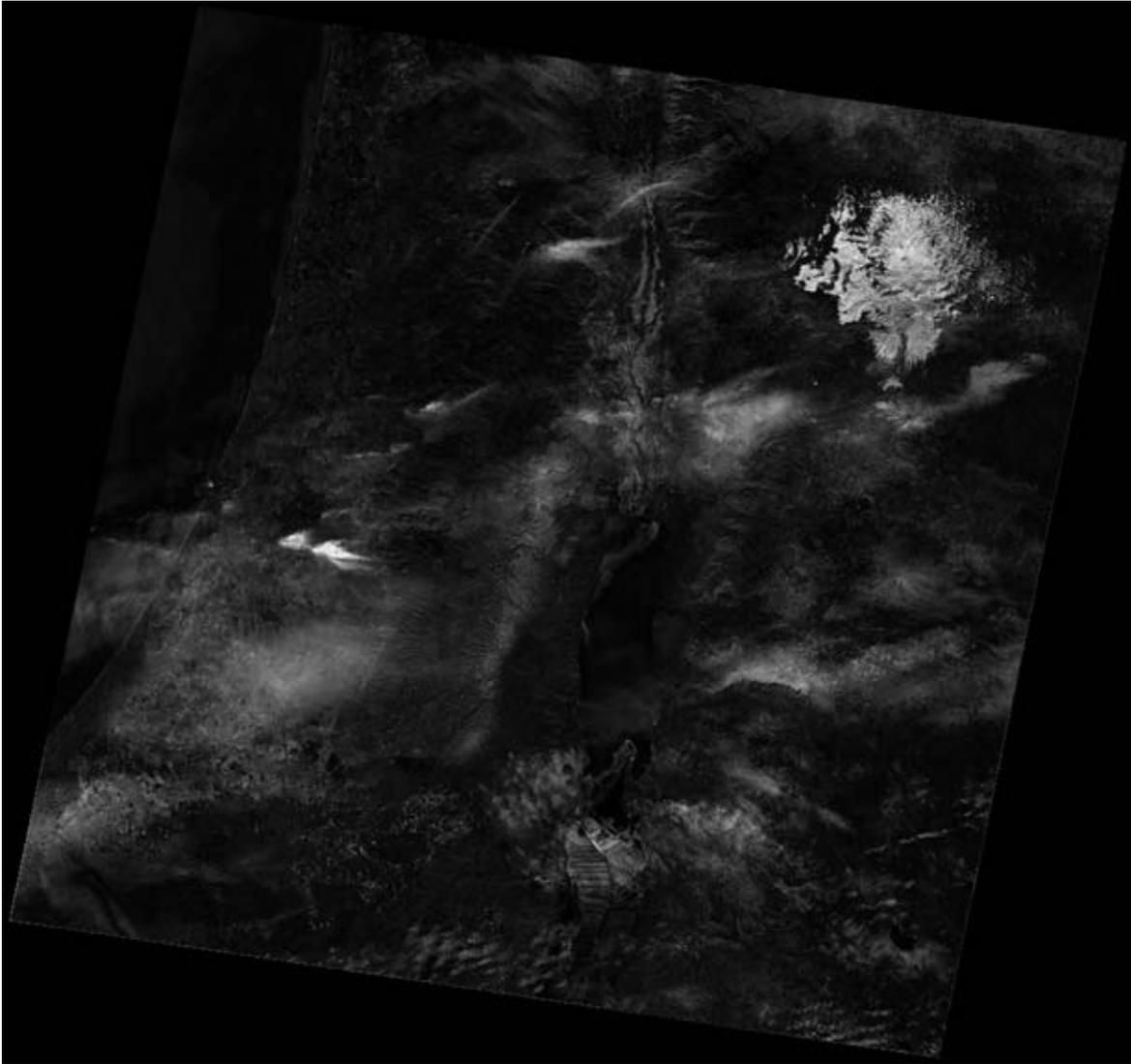


Figure 7.5.1: The AMV image can be utilized to detect change from image-to-image.

7.6 Compression of the AMV image for High Bandwidth Video Transfer

Many of today's digital video recorders offer advanced features such as motion compensation to remove the human induced jitter. This technique is basically an image registration technique that operates under very constrained limits (small amounts of rotation and translation). The ability to compute this registration in realtime allows for the motion compensation.

Since the AMV can be utilized image-to-image to capture change, it is easy to imagine how this can be of utility in image streaming applications. If the first frame of the video is maintained, it is possible to create the AMV image from each following image compared to the first. This would, in essence, be a 'change detection' video sequence that would allow for very high levels of compression due to the degree of spatial correlation in the data.

For example, run-length-encoding (RLE) is based on the premise that data can be compressed due to the repetition of grayscale values and the minimal memory required to save small difference values. In the case of the AMV, the repetitious occurrence of zero (or near zero) grayscale values, should allow for a high degree of compression. It would be imaginable to compress an entire video sequence in this manner for transmission over communications lines and then uncompress and rebuild the video sequence frame by frame. This would be useful in instances where the communications paths are severely restricted and processing time is of secondary concern. Two videos sequences would then be available for analysis, the 'difference' video and the original.

7.7 Frame Stacking for increasing Image S/N Ratio

Although this thesis will not discuss the image stacking process in detail, it is suffice to mention that this process is an excellent technique to increase image S/N by averaging out the noise. The 'overlaid images' that have been produced to visually inspect the accuracy of registration, actually average the registered images pixel-to-pixel. This simple process has been used for years in astronomy to increase the signal of 'dim' objects. The same technique can be utilized for remotely sensed images taken in low-light situations to improve the image S/N.

SUMMARY

Due to the increasing availability of both spaceborne and airborne imagery and spectral datasets now becoming available, multi-sensor image registration is becoming increasingly important. Utilizing the LoG thresholding process to automatically determine GCPs, in concert with wavelet multiresolution analysis, it is possible to automate this often time consuming process. This is possible due to the ability to compare images with similar frequency content through “Wavelet Decimation” or even standard downsampling procedures and automatically relate this edge detail utilizing point-matching techniques.

The ability to spatially relate images from different sensors allows for direct comparison and augments spectral, temporal, and spatial analysis. So, once the multi-sensor images have been registered, there are several applications that can be implemented to compare and/or fuse the information content. One application is the ability to “sharpen” the lower resolution spectral image with the detail gained from the high-resolution image for additional analysis. In this way, visual interpretation of the spectral data is more easily accomplished and additional spectral analysis is made possible. Additionally, applications such as change detection, image stacking, image mosaicing, cross-band correlation and spectral unmixing become more feasible.

In fact, many current remote sensing applications start with the assumption that two datasets are inherently registered (a fairly large assumption). This thesis attempts to add to the current body of knowledge in the area of automated image registration (section 8.1) and proposes new applications for this process.

8.1 Key Research Advancements

8.1.1 Automated GCP Selection

Critical to this research is the ability to extract similar GCPs from spatially related image sets. By utilizing the properties of the LoG filter to automatically detect and extract similar edges from two separate images, it is possible to threshold that result and obtain initial GCP sets. This technique lays the foundation for the entire thesis. The incorporation of techniques to extract GCP from both subregions and subbands using the LoG technique allows for the accurate registration of large multiresolution images with invariance to shift, rotation, and scale.

8.1.2 Automated Correlation of related GCPs using Point Matching Theory

Leveraging the ability of a relative distance matching technique to relate star fields, this thesis enables the registration of images once they have been boiled down to edge maxima point sets through the LoG thresholding process. Similar techniques have been created to compare possible GCP matches based on angle, scale, LoG Maxima value, and local statistical analysis of the match point distance.

8.1.3 Automated Statistical Analysis of RMSDE for Registration Accuracy

Once the GCPs have been extracted across the entire image, using subregion or subband techniques, statistical analysis of how well each matching GCP agrees with the overall transform model can be computed. This technique compares the RMS Distance Error of each matching point to every other point to reject outliers and to improve the accuracy of the model by removing matches with the most error. This quantitative registration metric lends itself to automation by allowing the user to define the level of error required in the model and provides a mechanism to iteratively strip off the matches that deviate most until the desired registration accuracy is attained.

8.1.4 Embedded Wavelet Structure to utilize Multiresolution Registration

By utilizing a wavelet structure for multiresolution registration it is possible to gracefully analyze images of different resolutions and register them without loss of high frequency content. By reducing the resolution of one image to the other using wavelet techniques, it is possible to utilize the LL Subband for registration similar to two image of equal resolution. Once the registration has been accomplished at the lowest common resolution, it is possible to predict the transform at any other wavelet level through manipulation of a composite affine transform. This ability allows analysts to compare the images at multiple levels of resolution (scales) and to quickly and efficient transition between these levels (zoom in and out).

8.1.5 Manipulations Automatically incorporated into Single Composite Model

With integration of composite transforms into the registration scheme, it is possible to incorporate several image manipulation automatically into a single, composite transformation model. This ability is predicated on the special commutative property of affine transforms and minimizes degradations due to numerous resamplings. This capability allows for multiresolution analysis by manipulating the scale coefficient in the affine 3x3 matrix and is a convenient way to mathematically express a series of image manipulations.

8.1.6 Predictive Transformation of hi-res Images from low-res coefficients

The multiresolution capabilities of the wavelet structure combined with the ability to easily manipulate composite affine relationships allows for the ability to register images at a reduced resolution and to predict the transform necessary to warp the original scale image. Also, by incorporating the global statistical analysis of RMSDE, it is possible to determine the expected model error at the original scale. The predictive transform technique not only allows for efficient registration of very large datasets at a reduced scale, it also provides a way to relate multiresolution image datasets.

8.1.7 Radiometrically accurate Wavelet Sharpening Technique

The development of a radiometrically accurate *Wavelet Sharpening* technique incorporates many of the earlier advancements mentioned above. The ability to relate images at the lowest common scale, determine the transform for the original high resolution image, decimate the warped image to the nearest dyadic level, transfer the high frequency detail, and finally reconstitute the low resolution image with the inverse FWT requires all of the above processes, integrated into a single application.

8.1.8 Techniques for Judging Registration Accuracy

New techniques have been developed in this research to judge the registration accuracy of the LoGWaR process. The Absolute Mean Variance (difference image) metric provides both a quantitative and qualitative measure of registration accuracy. While the RMSDE metric provides a quantitative and automated measure of how well the extracted and matched GCPs conform to a specific polynomial model. Both the overlaid image product and especially the ‘flicker test’ provide excellent qualitative mechanisms to corroborate the registration accuracy of warped images through user confirmation. All of these tests combine to give good confidence in the accuracy of the registration process developed in this thesis.

8.2 Related Research

Many of the capabilities developed in this thesis for automatically registering multisensor images have already been modified, streamlined, and incorporated into other applications for specific registration tasks. Matthew Egan and Peter Kopacz, of Eastman Kodak, have seamlessly integrated and optimized the process developed here to automatically register multiresolution datasets in preparation for multispectral sharpening using panchromatic imagery. Dr. William Reynolds, also of Eastman Kodak, has utilized GCP extraction and point matching techniques developed here as input to a zero-tree prediction and hierarchical estimation using high frequency wavelet subbands to increase registration accuracy at higher resolutions (Reynolds & Walli, 2003). Finally, Gabriel Dore of RIT and Derrick Campbell of Eastman Kodak are independently utilizing the automated registration techniques developed here to determine motion estimation vectors for video super-resolution research.

Bibliography

- Bernstein, C. A., L. N. Kanal, D. Lavin, and E. C. Olson, (1987). A geometric approach to subpixel registration accuracy. CVGIP, 40:334-360.
- Chandrasekhar, A. (1999). Point extraction and matching for registration of infrared astronomical images: xi, 88 leaves.
- Daubechies, I. (1992). Ten lectures on wavelets. Philadelphia, Pa., Society for Industrial and Applied Mathematics.
- Gonzalez, R. C. and R. E. Woods (2002). Digital image processing. Upper Saddle River, N.J., Prentice Hall.
- Gross, H. N. and J. R. Schott (2002). Application of spatial resolution enhancement and spectral mixture analysis to hyperspectral images. Rochester, NY, Rochester Institute of Technology.
- Hailstone (2001). Basic Principles of Imaging Science - Notes. Rochester Institute of Technology.
- Hubbard, B. B. (1996). The world according to wavelets. Wellesley, Ma., A K Peters.
- Mallat, S. e. (1989). "A Theory for Multiresolution Signal Decompositon: The Wavelet Representation." IEEE Transactions on Pattern Analysis and Machine Intelligence 11(7): 674-693.
- S. Park and R. Schowengerdt, 1983 "Image Reconstruction by Parametric Cubic Convolution", Computer Vision, Graphics & Image Processing 23, 256.
- Peytavin, L. (1996). "Cross-Sensor Resolution Enhancement of Hyperspectral Images using Wavelet Decomposition." SPIE Algorithms for MS & HS Imagery II(2758): 193-203.
- Rao, R. M. and A. S. Bopardikar (1998). Wavelet transforms : introduction to theory and applications. Reading, MA, Addison-Wesley.
- Reynolds, W .and Walli, K. (2003), "Wavelet-based Image Registration", SPIE Optical Science and Technology, Applications of Digital Image Processing XXVI.
- Schott, J. R. (1997). Remote sensing : the image chain approach. New York, Oxford University Press.
- Schowengerdt, R. A. (1997). Remote sensing, models, and methods for image processing. San Diego, Academic Press.
- Wolberg, G. (1990). Digital Image Warping. California, IEEE Computer Society Press.

Appendix A

“1-D” Haar Fast Wavelet Transform (FWT) Example

Adapted from notes provided during Digital Image Processing II, Dr. Harvey Rhody

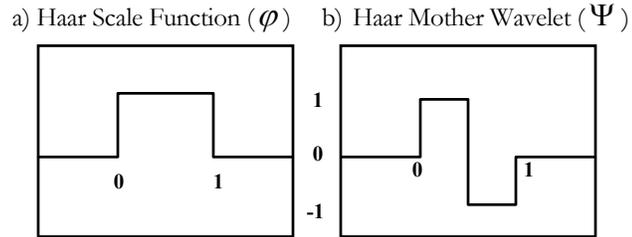


Figure A.1: The Haar Fast Wavelet Transform.

For a function $f(t)$ over the interval $[0,1)$:

$$f(t) = s_0 \varphi_{\left[0, \frac{1}{2}\right)} + s_1 \varphi_{\left[\frac{1}{2}, 1\right)}$$

$$f(t) = \frac{s_0}{2} (\varphi_{[0,1)} + \Psi_{[0,1)}) + \frac{s_1}{2} (\varphi_{[0,1)} - \Psi_{[0,1)})$$

$$f(t) = \frac{s_0 + s_1}{2} \varphi_{[0,1)} + \frac{s_0 - s_1}{2} \Psi_{[0,1)}$$

$$f(t) = c_0 \varphi_{[0,1)} + d_0 \Psi_{[0,1)}$$

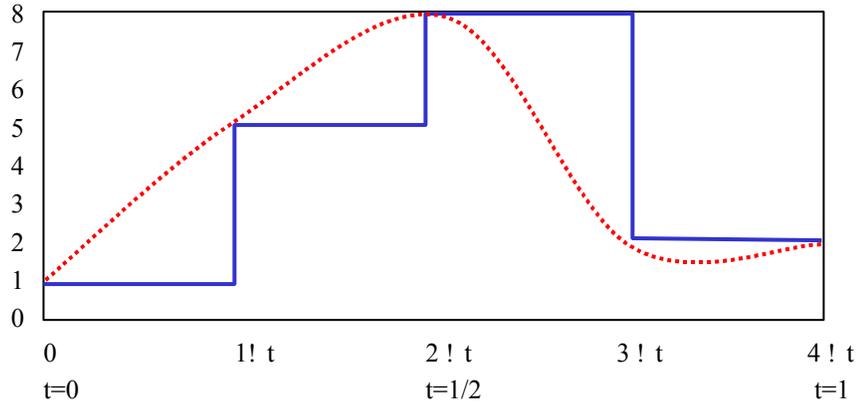
$$c_0 = \frac{1}{2}(s_0 + s_1)$$

$$d_0 = \frac{1}{2}(s_0 - s_1)$$

$$s_0 = c_0 + d_0$$

$$s_1 = c_0 - d_0$$

Figure A.2: An Example “1-D” Signal.



$$f(t) = 1\varphi_{\left[0, \frac{1}{4}\right)} + 5\varphi_{\left[\frac{1}{4}, \frac{1}{2}\right)} + 8\varphi_{\left[\frac{1}{2}, \frac{3}{4}\right)} + 2\varphi_{\left[\frac{3}{4}, 1\right)}$$

$$f(t) = \frac{1+5}{2}\varphi_{\left[0, \frac{1}{2}\right)} + \frac{1-5}{2}\Psi_{\left[0, \frac{1}{2}\right)} + \frac{8+2}{2}\varphi_{\left[\frac{1}{2}, 1\right)} + \frac{8-2}{2}\Psi_{\left[\frac{1}{2}, 1\right)}$$

$$f(t) = 3\varphi_{\left[0, \frac{1}{2}\right)} - 2\Psi_{\left[0, \frac{1}{2}\right)} + 5\varphi_{\left[\frac{1}{2}, 1\right)} + 3\Psi_{\left[\frac{1}{2}, 1\right)}$$

Utilizing the Haar FWT:

$$s = (1, 5, 8, 2) = a^{(2)}; \quad \text{Since, } n = \log_2 4 = 2$$

$$a^{(1)} = \left(\frac{1+5}{2}, \frac{8+2}{2} \right) = (3, 5)$$

$$c^{(1)} = \left(\frac{1-5}{2}, \frac{8-2}{2} \right) = (-2, 3)$$

$$s^{(1)} = (a^{(1)}; c^{(1)}) = (3, 5, -2, 3)$$

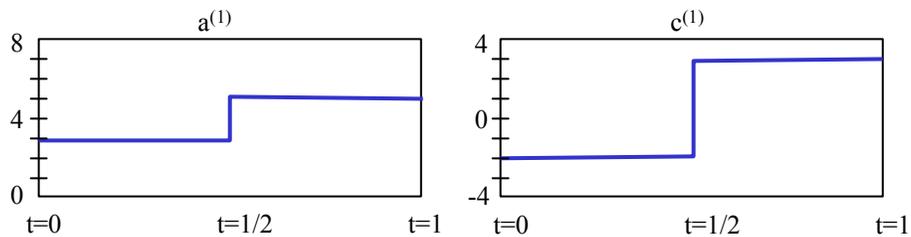


Figure A.3: The resulting Scale and Detail products, $s^{(1)}$, after first pass of FWT.

$$a^{(0)} = \left(\frac{3+5}{2} \right) = 4$$

$$c^{(0)} = \left(\frac{3-5}{2} \right) = -1$$

$$s^{(0)} = (a^{(0)}; c^{(0)}; c^{(1)}) = (4, -1, -2, 3)$$

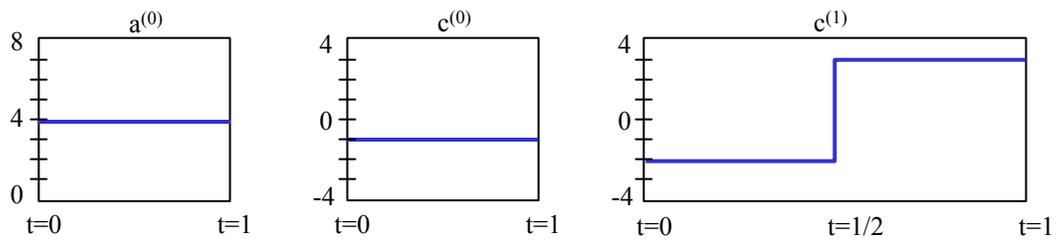


Figure A.3: The resulting Scale and Detail products, $s^{(0)}$, after second pass of FWT.

$$f(t) = 4\varphi^{(0)} - 1\Psi_0^{(0)} - 2\Psi_0^{(1)} + 3\Psi_1^{(1)}$$

“2-D” Haar FWT Example

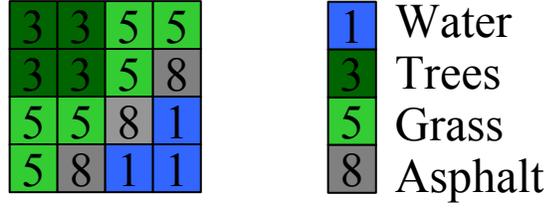


Figure A.4: Synthetic Image with Grayscale values representing scene compositions.

The FWT:

Original Image

$$\begin{array}{|c|c|c|c|} \hline 3 & 3 & 5 & 5 \\ \hline 3 & 3 & 5 & 8 \\ \hline 5 & 5 & 8 & 1 \\ \hline 5 & 8 & 1 & 1 \\ \hline \end{array} = \begin{array}{|c|c|c|c|} \hline 3 & 3 & 5 & 5 \\ \hline 3 & 3 & 5 & 8 \\ \hline 5 & 5 & 8 & 1 \\ \hline 5 & 8 & 1 & 1 \\ \hline \end{array} = \left\{ \begin{array}{l} f(x_0^1, y^0) = 3\varphi_{[0,1]} + 3\varphi_{[1,2]} + 5\varphi_{[2,3]} + 5\varphi_{[3,4]} \\ f(x_1^1, y^0) = 3\varphi_{[0,1]} + 3\varphi_{[1,2]} + 5\varphi_{[2,3]} + 8\varphi_{[3,4]} \\ f(x_2^1, y^0) = 5\varphi_{[0,1]} + 5\varphi_{[1,2]} + 8\varphi_{[2,3]} + 1\varphi_{[3,4]} \\ f(x_3^1, y^0) = 5\varphi_{[0,1]} + 8\varphi_{[1,2]} + 1\varphi_{[2,3]} + 1\varphi_{[3,4]} \end{array} \right.$$

Step 1: Decimate Image in “x” direction ($x \downarrow$).

$$f(x_0^1, y^0) = \frac{3+3}{2}\varphi_{[0,2]} + \frac{3-3}{2}\Psi_{[0,2]} + \frac{5+5}{2}\varphi_{[2,4]} + \frac{5-5}{2}\Psi_{[2,4]} = 3\varphi_{[0,2]} + 0\Psi_{[0,2]} + 5\varphi_{[2,4]} + 0\Psi_{[2,4]}$$

$$f(x_1^1, y^0) = \frac{3+3}{2}\varphi_{[0,2]} + \frac{3-3}{2}\Psi_{[0,2]} + \frac{5+8}{2}\varphi_{[2,4]} + \frac{5-8}{2}\Psi_{[2,4]} = 3\varphi_{[0,2]} + 0\Psi_{[0,2]} + 6.5\varphi_{[2,4]} - 1.5\Psi_{[2,4]}$$

$$f(x_2^1, y^0) = \frac{5+5}{2}\varphi_{[0,2]} + \frac{5-5}{2}\Psi_{[0,2]} + \frac{8+1}{2}\varphi_{[2,4]} + \frac{8-1}{2}\Psi_{[2,4]} = 5\varphi_{[0,2]} + 0\Psi_{[0,2]} + 4.5\varphi_{[2,4]} + 3.5\Psi_{[2,4]}$$

$$f(x_3^1, y^0) = \frac{5+8}{2}\varphi_{[0,2]} + \frac{5-8}{2}\Psi_{[0,2]} + \frac{1+1}{2}\varphi_{[2,4]} + \frac{1-1}{2}\Psi_{[2,4]} = 6.5\varphi_{[0,2]} - 1.5\Psi_{[0,2]} + 1\varphi_{[2,4]} + 0\Psi_{[2,4]}$$

Then for,

$$f(x_0^1, y^0): \quad a^{10} = (3, 5); \quad c^{10} = (0, 0); \quad s^{10} = (3, 5, 0, 0)$$

$$f(x_1^1, y^0): \quad a^{10} = (3, 6.5); \quad c^{10} = (0, -1.5); \quad s^{10} = (3, 6.5, 0, -1.5)$$

$$f(x_2^1, y^0): \quad a^{10} = (5, 4.5); \quad c^{10} = (0, 3.5); \quad s^{10} = (5, 4.5, 0, 3.5)$$

$$f(x_3^1, y^0): \quad a^{10} = (6.5, 1); \quad c^{10} = (-1.5, 0); \quad s^{10} = (6.5, 1, -1.5, 0)$$

$$\text{So, in image form, the scale plane } \varphi^{10} = \begin{bmatrix} 3 & 5 \\ 3 & 6.5 \\ 5 & 4.5 \\ 6.5 & 1 \end{bmatrix} \text{ \& the detail plane } \Psi^{10} = \begin{bmatrix} 0 & 0 \\ 0 & -1.5 \\ 0 & 3.5 \\ -1.5 & 0 \end{bmatrix}$$

Step 2: Now decimate the scale plane (φ^{10}) in the “y” direction (y↓).

$$\varphi^{10} = \begin{bmatrix} 3 & 5 \\ 3 & 6.5 \\ 5 & 4.5 \\ 6.5 & 1 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 3 & 5 \\ 3 & 6.5 \end{bmatrix} \\ \begin{bmatrix} 5 & 4.5 \\ 6.5 & 1 \end{bmatrix} \end{bmatrix} \equiv \left\{ \begin{array}{l} f(x, y_0) = 3\varphi_{[0,1]} + 3\varphi_{[1,2]} + 5\varphi_{[2,3]} + 6.5\varphi_{[3,4]} \\ f(x, y_1) = 5\varphi_{[0,1]} + 6.5\varphi_{[1,2]} + 4.5\varphi_{[2,3]} + 1\varphi_{[3,4]} \end{array} \right\}$$

$$\begin{aligned} f(x^1, y_0^1) &= \frac{3+3}{2}\varphi_{[0,2]} + \frac{3-3}{2}\Psi_{[0,2]} + \frac{5+6.5}{2}\varphi_{[2,4]} + \frac{5-6.5}{2}\Psi_{[2,4]} \\ &= 3\varphi_{[0,2]} + 0\Psi_{[0,2]} + 5.75\varphi_{[2,4]} - 0.75\Psi_{[2,4]} \end{aligned}$$

$$\begin{aligned} f(x^1, y_1^1) &= \frac{5+6.5}{2}\varphi_{[0,2]} + \frac{5-6.5}{2}\Psi_{[0,2]} + \frac{4.5+1}{2}\varphi_{[2,4]} + \frac{4.5-1}{2}\Psi_{[2,4]} \\ &= 5.75\varphi_{[0,2]} - 0.75\Psi_{[0,2]} + 2.75\varphi_{[2,4]} + 1.75\Psi_{[2,4]} \end{aligned}$$

Then for,

$$\begin{aligned} f(x^1, y_0^1): \quad a^{11} &= (3, 5.75); \quad c^{11} = (0, -0.75); \quad s^{11} = (3, 5.75, 0, -0.75) \\ f(x^1, y_1^1): \quad a^{11} &= (5.75, 2.75); \quad c^{11} = (-0.75, 1.75); \quad s^{11} = (5.75, 2.75, -0.75, 1.75) \end{aligned}$$

$$\text{So, the image scale plane } \varphi^{11} = \begin{bmatrix} 3 & 5.75 \\ 5.75 & 2.75 \end{bmatrix} \text{ \& the detail plane } \Psi^{11} = \begin{bmatrix} 0 & -0.75 \\ -0.75 & 1.75 \end{bmatrix}$$

This represents the 1st dyadic (power of 2) multiresolution decomposition. The scale image plane (φ^{11}) is half the resolution of the original image in both “x” and “y”.

Step 3: Now decimate the scale plane (φ^{11}) again in the “x” direction (x↓).

$$\varphi^{11} = \begin{bmatrix} 3 & 5.75 \\ 5.75 & 2.75 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 3 & 5.75 \end{bmatrix} \\ \begin{bmatrix} 5.75 & 2.75 \end{bmatrix} \end{bmatrix} \equiv \left\{ \begin{array}{l} f(x_0, y) = 3\varphi_{[0,1]} + 5.75\varphi_{[1,2]} \\ f(x_1, y) = 5.75\varphi_{[0,1]} + 2.75\varphi_{[1,2]} \end{array} \right\}$$

$$f(x_0^2, y^1) = \frac{3+5.75}{2} \varphi_{[0,2]} + \frac{3-5.75}{2} \Psi_{[0,2]} = 4.375\varphi_{[0,2]} - 1.375\Psi_{[0,2]}$$

$$f(x_1^2, y^1) = \frac{5.75+2.75}{2} \varphi_{[0,2]} + \frac{5.75-2.75}{2} \Psi_{[0,2]} = 4.25\varphi_{[0,2]} + 1.5\Psi_{[0,2]}$$

Then for,

$$f(x_0^2, y^1): a^{21} = (4.375); \quad c^{21} = (-1.375); \quad s^{21} = (4.375, -1.375)$$

$$f(x_1^2, y^1): a^{21} = (4.25); \quad c^{21} = (1.5); \quad s^{21} = (4.25, 1.5)$$

So, the image scale plane $\varphi^{21} = \begin{bmatrix} 4.375 \\ 4.25 \end{bmatrix}$ & the detail plane $\Psi^{21} = \begin{bmatrix} -1.375 \\ 1.5 \end{bmatrix}$

Step 4: Now decimate the scale plane (φ^{21}) again in the “y” direction (y↓).

$$\varphi^{21} = \begin{bmatrix} 4.375 \\ 4.25 \end{bmatrix} \equiv \left\{ f(x^2, y^2) = 4.375\varphi_{[0,1]} + 4.25\varphi_{[1,2]} \right\}$$

$$f(x^2, y^2) = \frac{4.375+4.25}{2} \varphi_{[0,2]} + \frac{4.375-4.25}{2} \Psi_{[0,2]} = 4.3125\varphi_{[0,2]} + 0.0625\Psi_{[0,2]}$$

Then for,

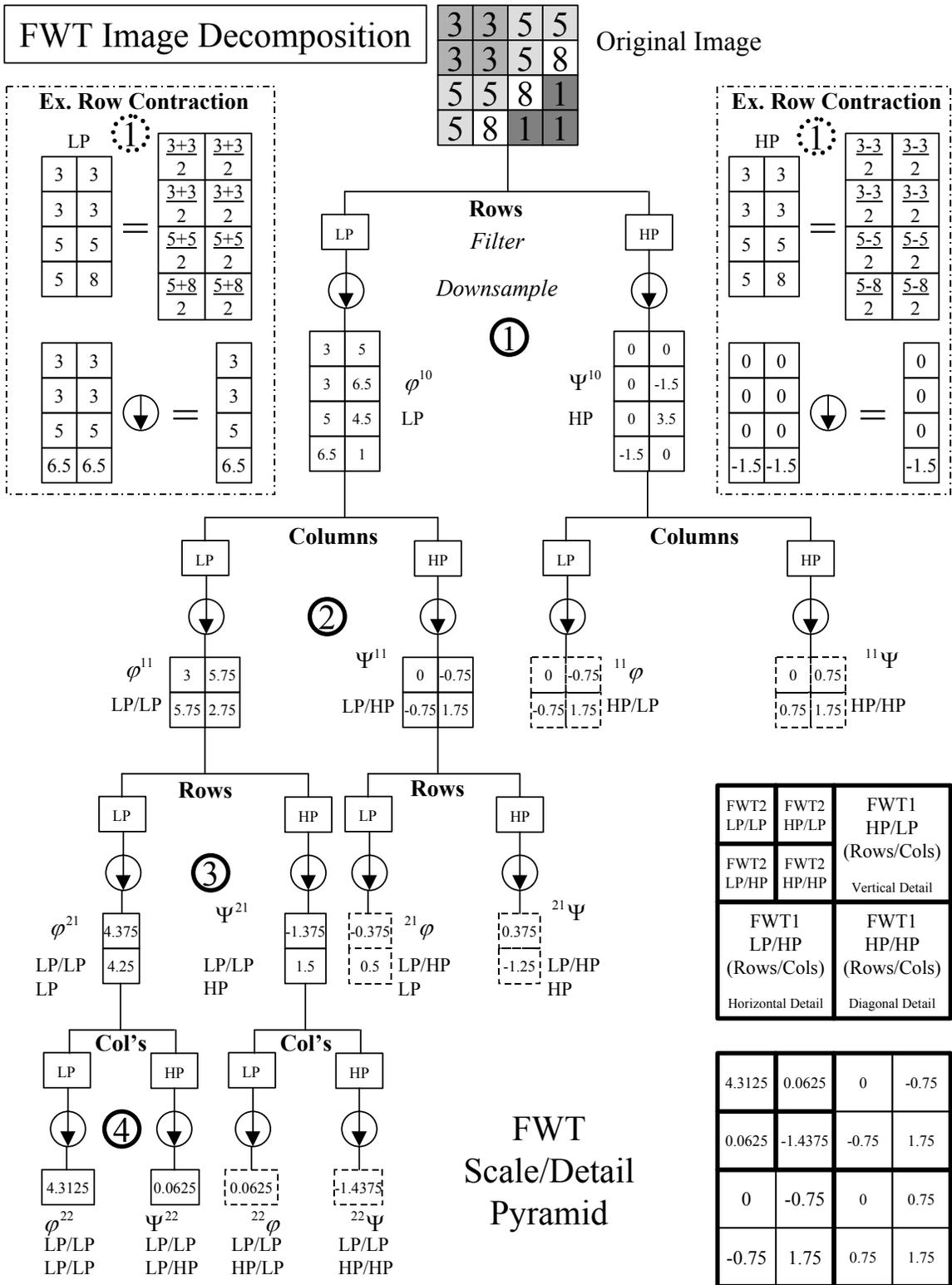
$$f(x^2, y^2): a^{22} = (4.3125); \quad c^{22} = (0.0625); \quad s^{22} = (4.3125, 0.0625)$$

So, the image scale plane $\varphi^{22} = \begin{bmatrix} 4.3125 \end{bmatrix}$ & the detail plane $\Psi^{22} = \begin{bmatrix} 0.0625 \end{bmatrix}$

At this point, the original image has been decomposed twice with the FWT and the scale plane (φ^{22}) is composed of only one “super-pixel”. This super-pixel value represents the average of the 16 pixels in the original image.

$$\varphi^{22} = \frac{3+3+5+5+3+3+5+8+5+5+8+1+5+8+1+1}{16} = \frac{69}{16} = 4.3125$$

The ability of the FWT to preserve the radiometry, on average (at each decimation), is important for many spectral applications including “unmixing”. In this way, overall radiometry is exactly preserved while high frequency information is stripped away.



FWT2 LP/LP	FWT2 HP/LP	FWT1 HP/LP (Rows/Cols)	
FWT2 LP/HP	FWT2 HP/HP	Vertical Detail	
FWT1 LP/HP (Rows/Cols)		FWT1 HP/HP (Rows/Cols)	
Horizontal Detail		Diagonal Detail	

4.3125	0.0625	0	-0.75
0.0625	-1.4375	-0.75	1.75
0	-0.75	0	0.75
-0.75	1.75	0.75	1.75

FWT
Scale/Detail
Pyramid

Figure A.5: Summary of FWT Image Decomposition (dashed boxes represent non-essential steps).

The FWT Scale/Detail Pyramid representation, as developed by Mallat (Mallat 89), is useful for two reasons. First the decomposition results in the same size matrix (4x4) as the original image, so it does not require additional storage requirements. Secondly, it allows for immediate analysis of detail in the horizontal, vertical and diagonal directions, while still producing the reduced resolution scale image (figure A.7). The only questionable requirement is the necessity for additional image processing steps (represented by the dashed boxes in figures A.5 and A.6) to obtain the vertical and diagonal detail planes.

Original Image Mallat's FWT Pyramid Representation

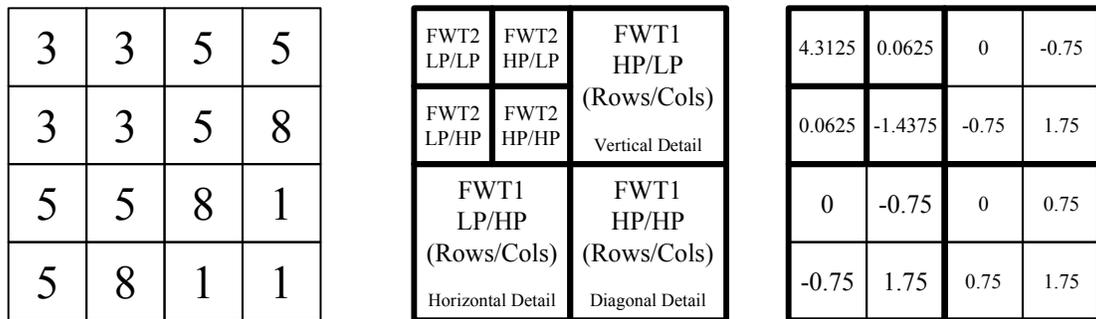


Figure A.7: Mallat's FWT Pyramid, showing that a 4x4 information matrix is retained.

One iteration of the FWT, on a real image, can be seen in Figure A.8 below:

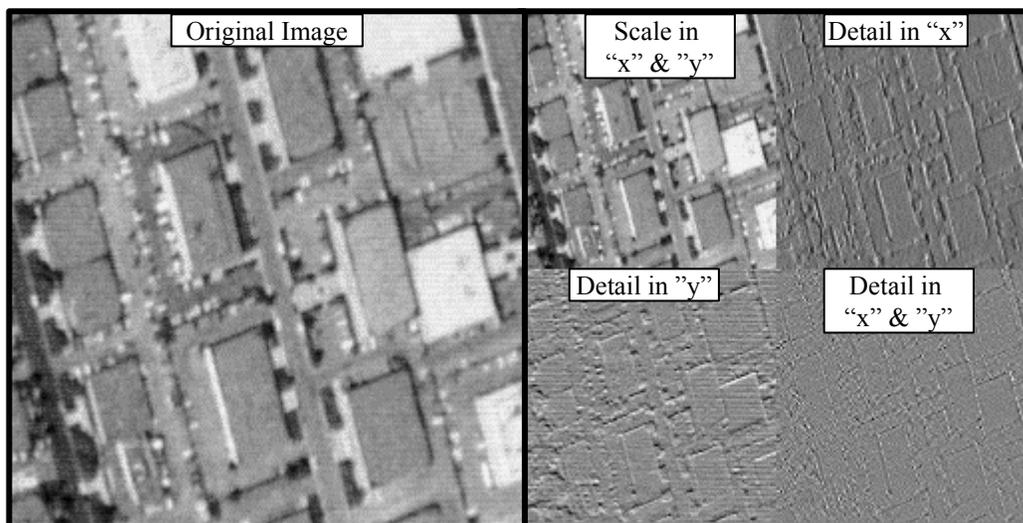


Figure A.8: Mallat's FWT Pyramid, showing that a 4x4 information matrix is retained.

Two iterations of the FWT, on the same image, can be seen in Figure A.9 below:

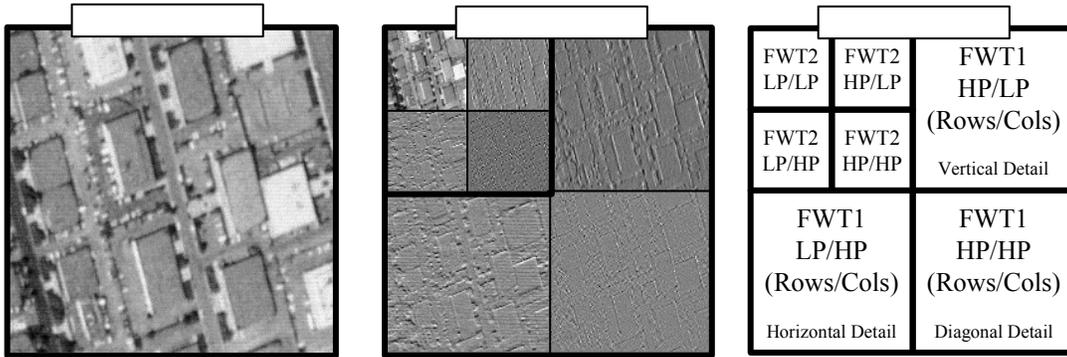


Figure A.9: Mallat's FWT Pyramid, showing the 2nd Iteration and related products.

So, each iteration of the FWT strips the image of its highest remaining frequencies, until none remain as demonstrated in Figure A.10:

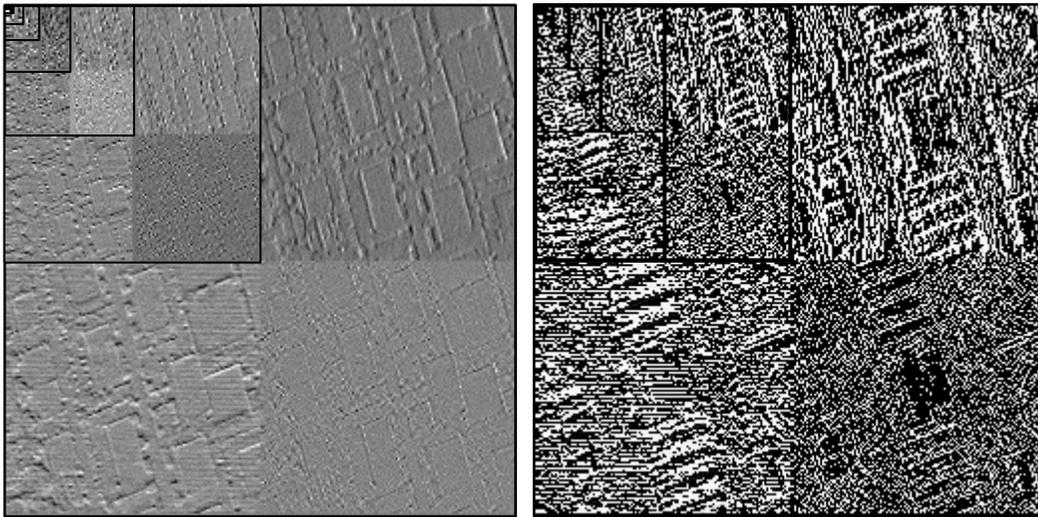


Figure A.10: Mallat's FWT Pyramid, showing a full decomposition of the original image.

A faster FWT decomposition would negate these additional steps, if the vertical and diagonal detail planes were not required. This representation may not be as useful for analysis (since the vertical detail is stretched and there is no diagonal detail represented), but the processing requirements are much less. This is because each iteration now requires 4 steps, instead of 6, saving 33% in processing for each stage (fig A.11). So, Mallat's Representation actually requires 50% more processing than what is required to orthogonally decompose the original image and still retain the ability to perfectly reconstruct it through the FWT^{-1} .

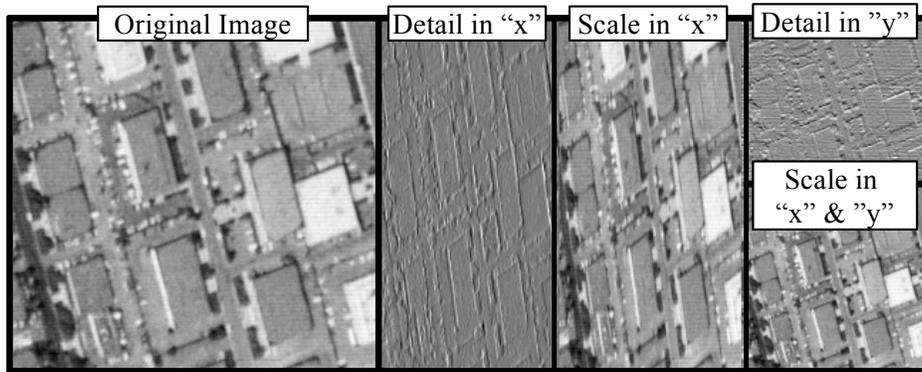


Figure A.11: FWT with minimum required processing steps (4).

The Inverse FWT (FWT⁻¹):

In order to prove that the FWT is an orthogonal decomposition, it is necessary to demonstrate the ability to precisely reconstruct the original signal from the scale plane and the detail planes. As expected, this process is very similar to the forward FWT, only in reverse:

Step 1: Reconstitute the scale plane (φ^{22}) in the “y” direction ($y\uparrow$) with Ψ^{22} .

$$\varphi^{22} = \boxed{4.3125}; \Psi^{22} = \boxed{0.0625}$$

$$\left. \begin{array}{l} \varphi_0^{21} = \varphi^{22} + \Psi^{22} \\ \varphi_1^{21} = \varphi^{22} - \Psi^{22} \end{array} \right\} \begin{array}{l} \left\{ \begin{array}{l} f(x^2, y_0^1) = a^{22} + c^{22} = 4.3125 + 0.0625 = 4.375 \\ f(x^2, y_1^1) = a^{22} - c^{22} = 4.3125 - 0.0625 = 4.25 \end{array} \right\} = \varphi^{21} \end{array}$$

$$\text{So, the image scale plane } \varphi^{21} = \boxed{\begin{array}{c} 4.375 \\ 4.25 \end{array}} \text{ \& the detail plane, stored earlier } \Psi^{21} = \boxed{\begin{array}{c} -1.375 \\ 1.5 \end{array}}$$

Step 2: Reconstitute the scale plane (φ^{21}) in the “x” direction ($x\uparrow$) with Ψ^{21} .

$$\varphi^{21} = \boxed{\begin{array}{c} 4.375 \\ 4.25 \end{array}}; \Psi^{21} = \boxed{\begin{array}{c} -1.375 \\ 1.5 \end{array}}$$

$$\left. \begin{array}{l} \varphi_0^{11} = \boxed{\varphi_{0[0,1]}^{11} \quad \varphi_{0[1,2]}^{11}} = \boxed{\varphi_0^{21} + \Psi_0^{21}; \quad \varphi_0^{21} - \Psi_0^{21}} \\ \varphi_1^{11} = \boxed{\varphi_{1[0,1]}^{11} \quad \varphi_{1[1,2]}^{11}} = \boxed{\varphi_1^{21} + \Psi_1^{21}; \quad \varphi_1^{21} - \Psi_1^{21}} \end{array} \right\} \begin{array}{l} \left\{ \begin{array}{l} f(x_0^1, y^1) = [a_0^{21} + c_0^{21}; a_0^{21} - c_0^{21}] \\ f(x_1^1, y^1) = [a_1^{21} + c_1^{21}; a_1^{21} - c_1^{21}] \end{array} \right\} = \varphi^{11} \end{array}$$

$$\varphi^{11} = \left[\begin{array}{cc} 4.375 + (-1.375) & 4.375 - (-1.375) \\ 4.25 + 1.5 & 4.25 - 1.5 \end{array} \right] = \boxed{\begin{array}{cc} 3 & 5.75 \\ 5.75 & 2.75 \end{array}} \quad (\text{new scale plane})$$

As before, the earlier stored detail plane, $\Psi^{11} = \boxed{\begin{array}{cc} 0 & -0.75 \\ -0.75 & 1.75 \end{array}}$ is used for reconstituting

in the next stage of the FWT⁻¹.

Step 3: Again reconstitute the scale plane (φ^{11}) in the “y” direction ($y\uparrow$) with Ψ^{11} .

(The formulation for $f(x^1, y^0)$ using the “a” and “c” notation is left out for clarity.)

$$\varphi^{11} = \begin{bmatrix} 3 & 5.75 \\ 5.75 & 2.75 \end{bmatrix}; \Psi^{11} = \begin{bmatrix} 0 & -0.75 \\ -0.75 & 1.75 \end{bmatrix}$$

$$\varphi^{10} = \begin{bmatrix} \varphi_{00}^{10} & \varphi_{10}^{10} \\ \varphi_{01}^{10} & \varphi_{11}^{10} \end{bmatrix} = \begin{bmatrix} \boxed{\varphi_{00}^{11} + \Psi_{00}^{11}} & \boxed{\varphi_{10}^{11} + \Psi_{10}^{11}} \\ \boxed{\varphi_{00}^{11} - \Psi_{00}^{11}} & \boxed{\varphi_{10}^{11} - \Psi_{10}^{11}} \\ \boxed{\varphi_{01}^{11} + \Psi_{01}^{11}} & \boxed{\varphi_{11}^{11} + \Psi_{11}^{11}} \\ \boxed{\varphi_{01}^{11} - \Psi_{01}^{11}} & \boxed{\varphi_{11}^{11} - \Psi_{11}^{11}} \end{bmatrix} = \begin{bmatrix} 3+0 & 5.75+(-0.75) \\ 3-0 & 5.75-(-0.75) \\ 5.75+(-0.75) & 2.75+1.75 \\ 5.75-(-0.75) & 2.75-1.75 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 3 & 6.5 \\ 5 & 4.5 \\ 6.5 & 1 \end{bmatrix}$$

$$\text{So, in image form, the scale plane } \varphi^{10} = \begin{bmatrix} 3 & 5 \\ 3 & 6.5 \\ 5 & 4.5 \\ 6.5 & 1 \end{bmatrix} \text{ \& recall that } \Psi^{10} = \begin{bmatrix} 0 & 0 \\ 0 & -1.5 \\ 0 & 3.5 \\ -1.5 & 0 \end{bmatrix}$$

Step 4: Finally, the scale plane (φ^{10}) is reconstituted in the horizontal ($x\uparrow$) with Ψ^{10} .

(Again, the formulation for $f(x^1, y^0)$ using the “a” and “c” notation is left out for clarity.)

$$\varphi^{10} = \begin{bmatrix} 3 & 5 \\ 3 & 6.5 \\ 5 & 4.5 \\ 6.5 & 1 \end{bmatrix}; \Psi^{10} = \begin{bmatrix} 0 & 0 \\ 0 & -1.5 \\ 0 & 3.5 \\ -1.5 & 0 \end{bmatrix}$$

$$\varphi^{00} = \begin{bmatrix} \boxed{\varphi_{00}^{10} \pm \Psi_{00}^{10}} & \boxed{\varphi_{10}^{10} \pm \Psi_{10}^{10}} \\ \boxed{\varphi_{01}^{10} \pm \Psi_{01}^{10}} & \boxed{\varphi_{11}^{10} \pm \Psi_{11}^{10}} \\ \boxed{\varphi_{02}^{10} \pm \Psi_{02}^{10}} & \boxed{\varphi_{12}^{10} \pm \Psi_{12}^{10}} \\ \boxed{\varphi_{03}^{10} \pm \Psi_{03}^{10}} & \boxed{\varphi_{13}^{10} \pm \Psi_{13}^{10}} \end{bmatrix} = \begin{bmatrix} 3+0 & 3-0 & 5+0 & 5-0 \\ 3+0 & 3-0 & 6.5+(-1.5) & 6.5-(-1.5) \\ 5+0 & 5-0 & 4.5+3.5 & 4.5-3.5 \\ 6.5+(-1.5) & 6.5-(-1.5) & 1+0 & 1-0 \end{bmatrix}$$

$$\varphi^{00} = \begin{bmatrix} 3 & 3 & 5 & 5 \\ 3 & 3 & 5 & 8 \\ 5 & 5 & 8 & 1 \\ 5 & 8 & 1 & 1 \end{bmatrix} \text{ So, the final scale plane is derived through the } \text{FWT}^{-1}, \text{ and a}$$

perfect reconstruction of the original image is reproduced.

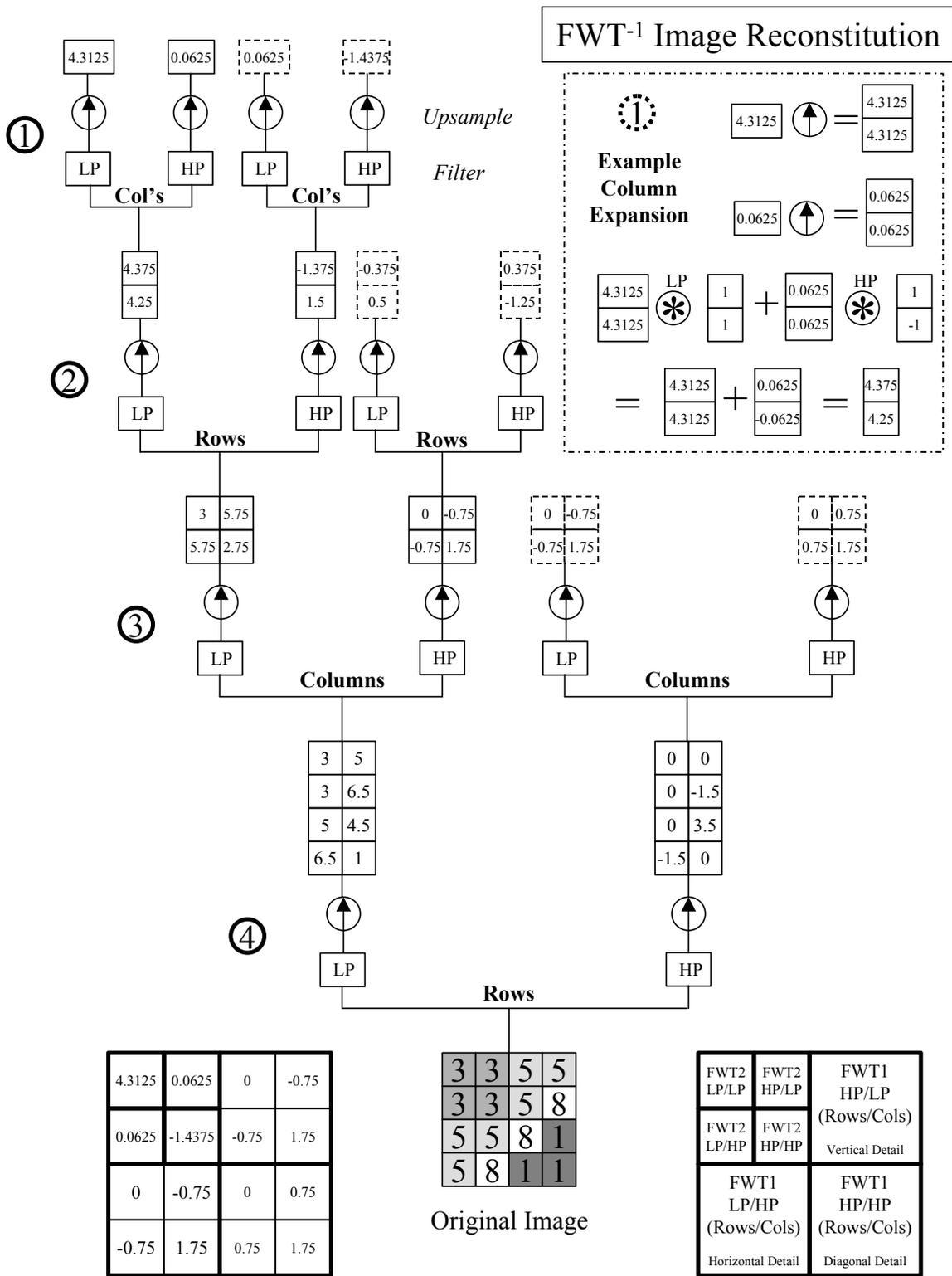


Figure A.6: Summary of FWT⁻¹ Image Reconstitution (dashed boxes represent non-essential steps).

SHARPENING

The radiometrically accurate wavelet sharpening process, proposed in section 7.1, can be compared below with Munechika's sharpening method devised in 1990.

B.1 Sharpening via Munechika's method

“One of the most straightforward methods of fusing multispectral data with higher-resolution panchromatic data relies on the assumption that there is some degree of correlation between the multispectral band and the pan band brightness values.” (Schott 1997, pg 310) Utilizing this concept, Munechika et. al., devised a method to merge multisensor data. Specifically, they tested their concept by merging SPOT 10m panchromatic images to Landsat TM 30m multispectral data. Munechika's method can be summarized in the following steps (Munechika 1990, pgs 12-13) :

- 1) The SPOT image is geometrically registered to the Landsat images.
- 2) A medium resolution panchromatic image is created from a weighted average of the Landsat TM bands 1 through 4. This synthetic image approximates the same spectral characteristics as the high resolution SPOT panchromatic channel.
- 3) The histogram of the SPOT panchromatic image is then linearly adjusted to the histogram of the synthetic TM panchromatic image. This transformation will, to the first order, account for the differing atmospheric and sensor effects between the SPOT and the Landsat TM images.
- 4) The images are then merged to create a high resolution, multiband hybrid image. The merging algorithm is:

$$(B.1) \quad DC_{Hybrid\ Multiband}(i) = DC_{SPOT\ Pan} \cdot \left(\frac{DC_{TM}(i)}{DC_{Syn\ TM\ Pan}} \right)$$

where:

$DC_{Hybrid\ Multiband}(i)$ = the digital count of the i^{th} band in the hybrid image

$DC_{SPOT\ Pan}$ = the digital count in the adjusted panchromatic SPOT image

$DC_{TM}(i)$ = the count in the i^{th} band of the original multispectral image

$DC_{Syn\ TM\ Pan}$ = the digital count in the synthetic TM panchromatic image

This method is applied on a pixel-by-pixel basis, and therefore each of the above terms also has a pixel location “(x, y)” associated with it. An example of how this technique is implemented follows:

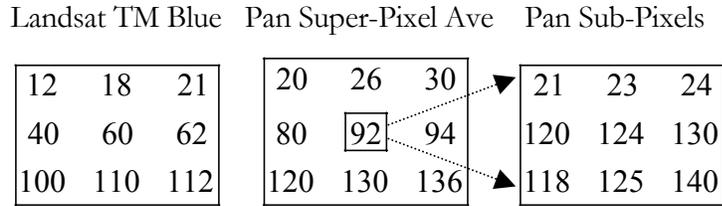


Figure B.1: Comparison of 3x3 Landsat region compared to 9x9 SPOT reduced to 3x3 Superpix.

With this raw data, it is possible to compute the $DC_{Hybrid\ Blue}$ 3x3 pixel region. This region represents the sharpened Landsat TM blue spectral band.

$$DC_{Hybrid\ Blue} = \frac{DC_{TM\ Blue}}{DC_{Pan\ Superpix}} \left[DC_{Pan\ Subpix} \right] = \frac{60}{92} \left[\begin{array}{|c|c|c|} \hline 21 & 23 & 24 \\ \hline 120 & 124 & 130 \\ \hline 118 & 125 & 140 \\ \hline \end{array} \right] = \begin{array}{|c|c|c|} \hline 14 & 15 & 16 \\ \hline 78 & 81 & 85 \\ \hline 77 & 82 & 91 \\ \hline \end{array}$$

$\overline{DC}_{Hybrid\ Blue}$ represents the average digital count value over the original space covered by the Landsat TM blue band.

$$\overline{DC}_{Hybrid\ Blue} = \frac{\sum_{i=1}^N DC_{Hybrid\ Blue}}{N} = \frac{(14 + 15 + 16 + 78 + 81 + 85 + 77 + 82 + 91)}{9} = 60$$

The fact that the results above equate to the original Landsat TM Blue pixel value, are significant. “This means that on average at the resolution of the original MS imagery, the radiometry is preserved exactly...this approach yields both radiometrically and visually improved images.” (Schott 1997, pg 310) The Munechika method is a good introduction to spatial sharpening since the method is fairly intuitive and yields radiometrically correct (on average) results.

B.2 Sharpening via Gross’s method

Gross et. al., devised an approach that relies first on the ability to spectrally unmix a dataset before application of a sharpening operation. This method attempts to sharpen the endmember fraction maps associated with unmixing as opposed to the actual raw imagery. Further detail is contained in the report referenced in the bibliography (Gross and Schott 2002). Unfortunately Gross’s method often involves underdetermined problems, since there are many more unknowns than equations, and thus is only mentioned for parties interested in further research.

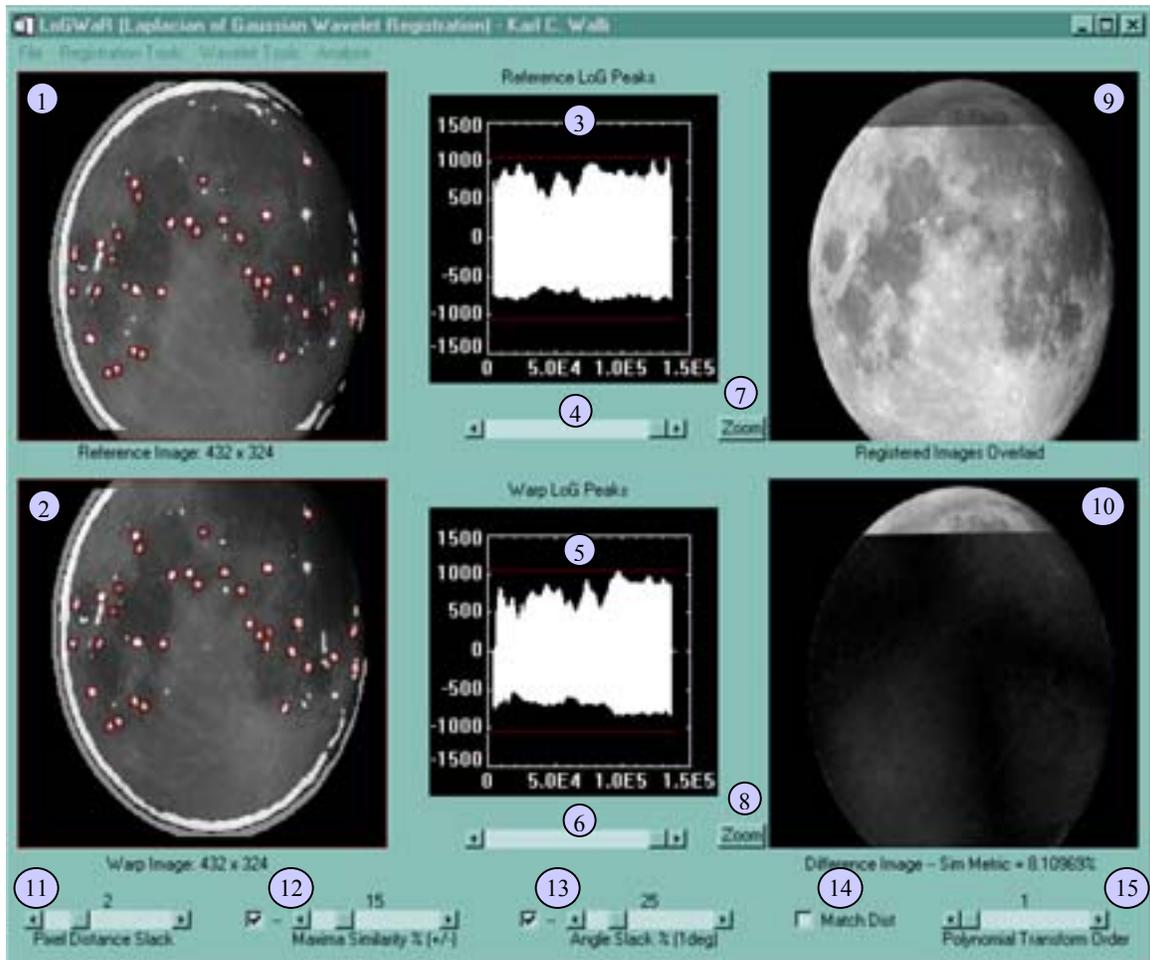
Appendix C

LoGwaR
Users Manual
Version 1.1

Karl C. Walli

July 2003

The LoGWaR GUI



LoGWaR Legend

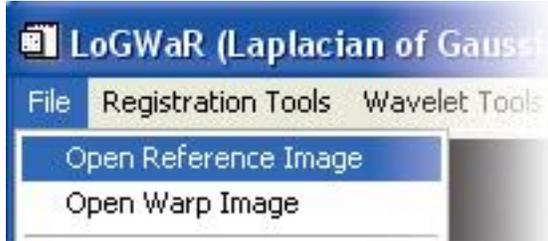
- | | |
|--------------------------------|---------------------------------------|
| 1) Reference Image | 9) Overlaid Images/Reference Zoom |
| 2) Warp Image | 10) Difference Image/Warp Zoom |
| 3) 1-D Ref LoG Plot | 11) Pixel Distance Slack Slider |
| 4) Ref Threshold Slider | 12) Maxima Similarity Slider |
| 5) 1-D Warp LoG Plot | 13) Angle Slack Slider |
| 6) Warp Threshold Slider | 14) Match Distance Button |
| 7) Reference Image Zoom Button | 15) Polynomial Transform Order Slider |
| 8) Warp Image Zoom Button | |

- 1) **Reference Image:** The Reference image display window shows a thumbnail image that is scaled to 256x256. The reference image provides the stable coordinate system to which the Warp image is transformed.
- 2) **Warp Image:** The Warp image display window shows a thumbnail image that is scaled to 256x256. The Warp image is registered to the Reference image and is then transformed to its coordinate system resulting in a Warped image.
- 3) **1-D Ref LoG Plot:** The 1-Dimensional plot of the 2-Dimensional LoG filtered Reference image. This visualization allows quick viewing of the peaks and valleys within the filtered image that will be isolated with a threshold procedure.
- 4) **Ref Threshold Slider:** This slider allows manual thresholding of the Reference image. This tool is used in conjunction with the *Manual Threshold* function found within the Analyze pull-down menu to isolate regions similar to the Warp image.
- 5) **1-D Warp LoG Plot:** The 1-Dimensional plot of the 2-Dimensional LoG filtered Warp image. This visualization allows quick viewing of the peaks and valleys within the filtered image that will be isolated with a threshold procedure.
- 6) **Warp Threshold Slider:** This slider allows manual thresholding of the Warp image. This tool is used in conjunction with the *Manual Threshold* function found within the Analyze pull-down menu to isolate regions similar to the Reference image.
- 7) **Reference Image Zoom Button:** This button provides a zoom feature of the ROI defined within the Reference image. This technique is often useful when choosing supervised GCPs. To select a GCP, simply left-click within the zoom window to the right of the button and then click the related pixel in the Warp zoom window.
- 8) **Warp Image Zoom Button:** This button provides a zoom feature of the ROI defined within the Warp image. This technique is often useful when choosing supervised GCPs. To select a GCP, simply left-click within the zoom window to the right of the button and then click the related pixel in the Reference zoom window.
- 9) **Overlaid Images/Reference Zoom:** This multifunctional window can display the resulting stacked images of a registration operation (Reference and Warped), zoom operations, or interim Reference thresholds when utilizing the *Adaptive – Subregion* procedure. The text below the window describes the current contents.

- 10) Difference Image/Warp Zoom:** This multifunctional window can display the difference image of a registration operation (Reference and Warped), zoom operations, or interim Warp image thresholds when utilizing the *Adaptive – Subregion* procedure. The text below the window describes the current contents.
- 11) Pixel Distance Slack Slider:** This slider determines the tightness of the control over relative distance matching between points within each dataset. With the default setting of 2, the point matching routine tests matching distances to within a difference of two pixels. Positive distance matches must agree to within the pixel error specified with this slider. The relative pixel matching process is the first in a series of four point matching algorithms. Each algorithm takes the results of the previous matching process to iteratively pare the point-sets down to the best matches. So, the relative distance matching technique determines the preliminary matches from which all subsequent matching techniques utilize.
- 12) Maxima Similarity Slider:** This slider analyses the similarity of the LoG filtered image values of the current matches (determined by the pixel distance matching algorithm). The similarity of the LoG value must be within plus or minus the percent identified by this slider (default is 15%).
- 13) Angle Slack Slider:** This slider determines the tightness of the control over relative angle comparison of the existing point matches. This technique analyses the angular position of matches with respect to the other points in the matched point set.
- 14) Match Distance Button:** The Match Distance Button should only be utilized on images that demonstrate good rotational agreements with each other. If this is true, then the matched points should have similar distances between the GCP pairs. If this box is checked, the matched GCPs distances will be analyzed for similarity and outliers will be rejected if they vary by more than one standard deviation from the mean.
- 15) Polynomial Transform Order Slider:** This slider can be manipulated to change the polynomial degree desired for a polynomial transform. Once the degree of the necessary polynomial is chosen, the GCPs can be utilized to determine the necessary coefficients. This toggle can also be utilized to switch the polynomial degree (1st or 2nd) utilized when computing the model error with the *Compute Matched RMS Error* function.

1. FILE Pull-down Menu Options:

Loading Images into LoGWaR:



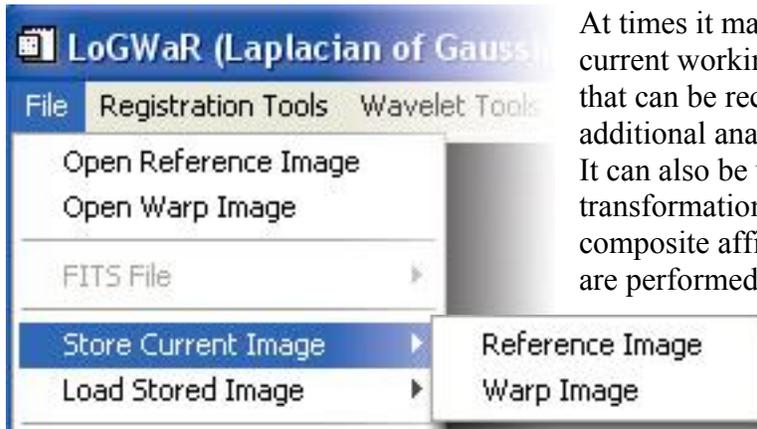
Utilizing the *FILE* pull-down menu option, it is possible to load both the Reference and Warp images into the LoGWaR environment for registration. The Reference image should be the image that will remain unaltered, while the Warp image will be the image that is to be

transformed into the Reference image geometric space. All of the standard image formats supported by IDL can be loaded in utilizing this function (JPEG, TIFF, etc.). If ENVI is available, the LoGWaR_ENVI version should be utilized to enable the full extensibility to data formats that can be opened utilizing ENVI's I/O functions.

Working with FITS Images:

The FITS I/O and additional tools can only be utilized for licenses that include IDL's Astronomy Library...

Storing the Current Working Images:



At times it may be necessary to store the current working images into a temporary file that can be recalled at a later time for additional analysis or as a safety precaution. It can also be utilized for predictive transformation when working with composite affines, since the transformations are performed on the Stored images.

Any image that is loaded will be automatically saved within the Stored image variables.

Loading Stored Images:



Loading the stored images allows for the ability to return to a previous point in the image registration process. It can also be used to load in images that were transferred in as IDL variables exported from ENVI. These images are placed in the stored image variables during execution of the LoGwaR program (prompt> logwar, reference, warp).

The LoGwaR program also allows the user to load the Warped image into the Warp image variable for additional registration analysis. This technique is

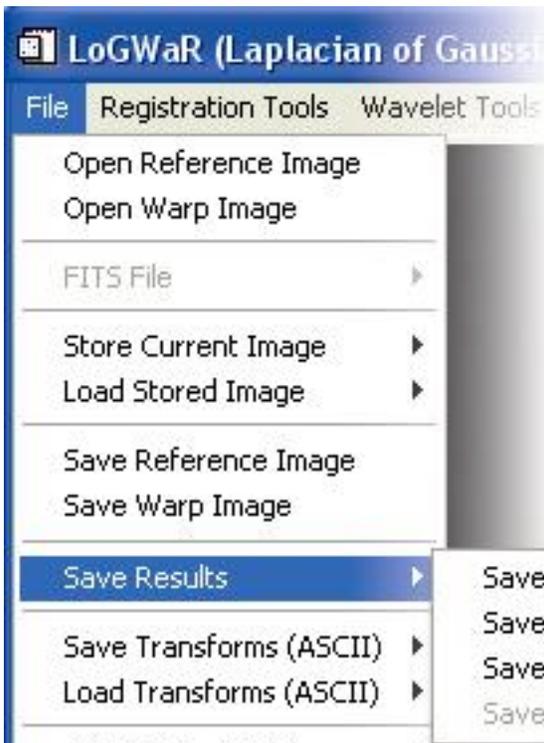
extremely useful when obtaining a rough estimate of the registration with a reduced scale image, and then utilizing the results of this operation to allow a subregion registration analysis for increased accuracy.

Saving the Current Working Images:



Occasionally it will be useful to utilize the *Registration Tools* menu to perform basic image processing manipulations on the working images (rotation, cropping) and to save the results, instead of just temporarily storing them. These functions allow the user to save the Reference or Warp working images as one of the basic image types supported in IDL.

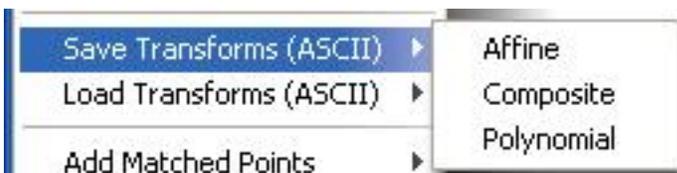
Saving the Registrion Results:



The LoGWaR program produces three useful image based results that can be saved using this function. The first, and most important, is the resulting warped image obtained from the transformation of the Warp input image to the Reference image space. This image will be the same dimensions as the Reference image and may be cropped to maintain those dimensions. Another product that can be saved is the Difference Image (Absolute Mean Variance), which represents the change in pixel value between the Reference image and the Warped Image. The final output is the Composite image or “stacked image”. This averaged image is often used

to improve the Signal to Noise of datasets by averaging-out the noise. In order to save a Multi-Band image it is necessary to save out the transformation coefficients and utilize the Thin_Warp.pro program to warp the entire dataset.

Save Transformation Coefficients:

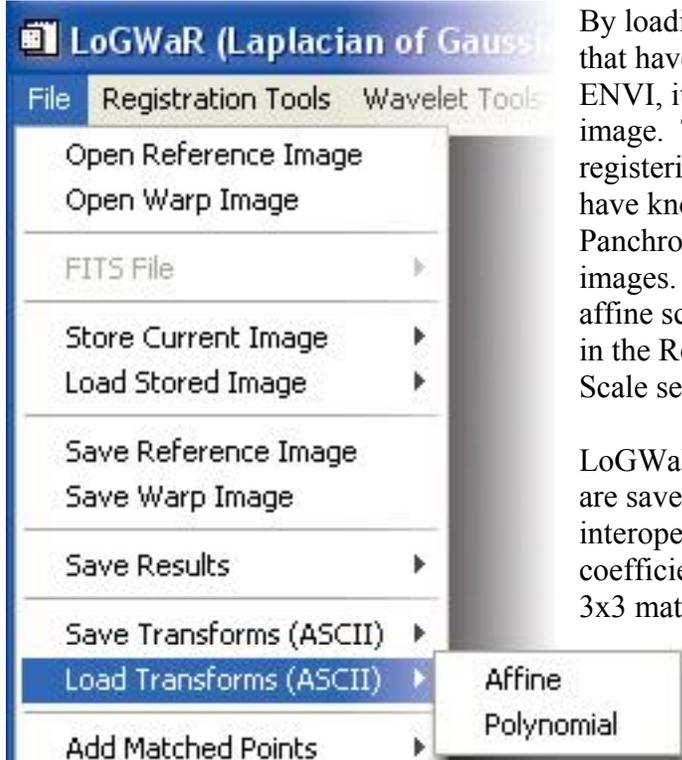


The LoGWaR program develops transformation coefficients from the matched points it extracts from the Reference and Warp images. The Affine coefficients

represent the most current affine registration results saved in a 3x3 matrix format. The Composite coefficients represents the cascade of the entire chain of image manipulations that have been combined through the *Update Composite* function and is also saved in a 3x3 matrix. The Polynomial transform is similar to the Affine except it is saved in the ENVI format and is variable in size dependent on the degree of the polynomial.

These coefficients can be utilized with the Thin_Warp.pro program to warp very large and Multi-Band images.

Load Transformation Coefficients:



By loading affine or polynomial coefficients that have been developed by LoGwaR or ENVI, it is possible to transform the Warp image. This is very useful, especially when registering datasets such as LANDSAT, that have known scale relationships between the Panchromatic, Multispectral, and Thermal images. Once loaded into memory the affine scale relationship can be manipulated in the Registration Tools pull-down in the Scale section.

LoGwaR and ENVI polynomial coefficients are saved in the same format for interoperability. However, the affine coefficients are saved by LoGwaR into a 3x3 matrix format, that can be used for the predictive scale transformation.

Adding Current Matches to the Stack:



LoGwaR matched points are treated as temporary results, until they've been accepted by the user! This

is done to allow ROI matches to be cumulatively added to increase the number of matches utilized for the image-wide registration.

Once the matches are added to the stack, it is possible to perform statistical analysis of the matches and/or utilize these matches to register the datasets.

This function allows the user to add matches that have been derived by LoGwaR or through manual selection of GCPs using point selection in the zoom window.

Adding recent matches to the stack can easily be forgotten; so take special precaution to ensure this is done, especially once 'good matches' have been derived by LoGwaR!

Saving Matched Points:



The ability to save matches either derived by LoGwaR or through supervised GCP selection is available through this function. There are two options available, the ability to save matches into a format for LoGwaR or ENVI. The reason for this is due to the way that LoGwaR loads images (from bottom-to-top format) as apposed to ENVI (from top-to-bottom format). This problem is overcome by inverting the vertical position of the matches with respect to the horizontal position.

So, to utilize matches derived by the LoGwaR program within ENVI, it is necessary to save the matches in the ENVI format. This precaution is not necessary when utilizing the LoGwaR_ENVI version, since the images are opened and saved utilizing the ENVI I/O operations.

Loading matched points into LoGwaR:



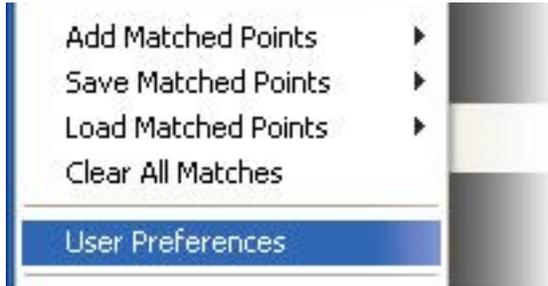
Loading matches into LoGwaR follows the same format as the *Save Matched Points* pull-down menu. Matches obtained from ENVI can be loaded in from this menu function.

Clearing all current matches:

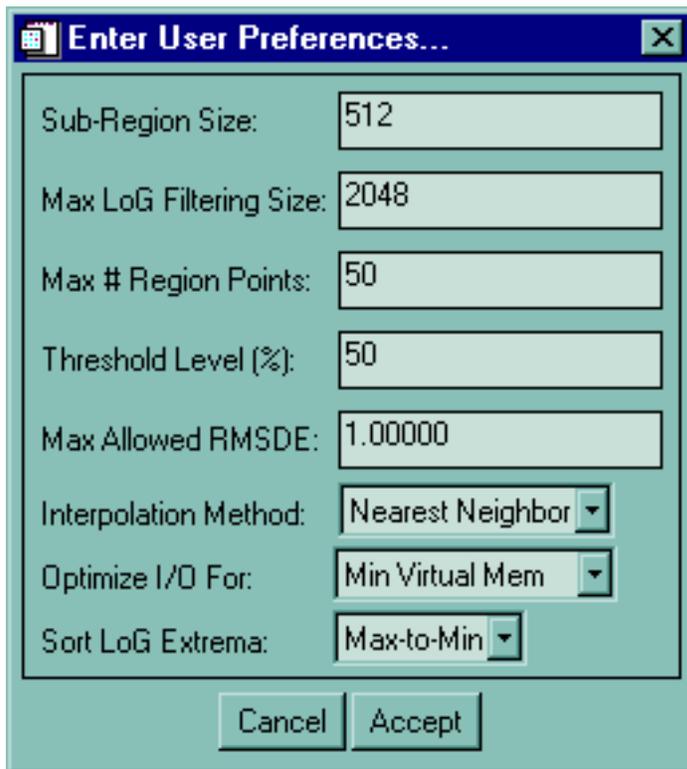


Occasionally it is desirable to clear all the current matches in order to load matches from a saved point match file or to start with a clean slate. The *Clear All Matches* menu function can be utilized to accomplish this task.

User Preferences:



LoGWaR allows several preferences to be changed to maximize flexibility and user interaction. Once the *User Preferences* pull-down menu has been chosen, a new GUI window will appear with several variables that can be changed to suit specific registration requirements.



The *Sub-Region Size* determines the horizontal and vertical dimensions of the sub-region that will be utilized for registration analysis. The size of this region depends on the number of matches desired image-wide and is limited only by processing speed and virtual memory (~ 128 to 2048 pix).

The *Max LoG Filtering Size* limits the image-wide filtering that can be accomplished. This is useful for very large images where the image size is too large for LoG convolution and where sub-region or sub-band analysis is utilized to bring the size below the maximum value.

The *Max # of Region Points* limits the number of potential

matches to the value identified in this field. The LoG maxima & minima that have been extracted through the threshold procedure are first sorted based on their rate-of-variation and then truncated to the number identified in this field. This process identifies the most well defined edge regions and limits the number to the requested amount. The peak pixels are identified from these regions and are then utilized for the input to the point matching algorithms.

The *Threshold Level (%)* identifies the LoG threshold level that will be utilized when utilizing the <Analyze> <Auto Threshold> <Preset> pull-down option. This preference is useful for registration of datasets that often require the same threshold level (FITS images are often at 1%).

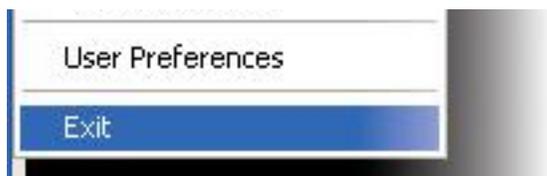
The **Max Allowed RMSDE** value can be changed to suit user requirements. This variable determines how well the cumulative match error must agree with the polynomial transform (1st or 2nd Degree) model for the entire image. When utilizing the <Registration Tools> <Delete Matches> <Delete RMSDE > User Defined> pull-down option, the match with the greatest error will be removed and the transform model will be recomputed. In this way the matches that deviate most from the transform model will be iteratively removed from the matched point list until the overall model contain less cumulative RMS Distance Error than the value identified within this field. This preference is very useful to ensure the overall model error is within desired parameters (i.e. requirement for subpixel registration accuracy). For predictive transformations, where the model error will be increased by the same factor as the scale modulator, it is necessary to keep the registration error to a low sub-pixel level. For example, the transform model of RMSDE = 0.25 becomes RMSE = 1.0 if the predictive transform warps an original image that is four times the scale (i.e. $0.25 * 4 = 1.0$).

The **Interpolation Method** determines the type of sampling that is utilized for the image transformation. Choices for this preference include *Nearest Neighbor* (NN), *Bilinear Interpolation* (BL), and *Bicubic Interpolation* (BC). The NN sampling method retains the image radiometric accuracy but sacrifices edge smoothness. BL interpolation provides a relatively fast sampling technique with good edge characteristics. Whereas, BC sampling delivers the ‘best looking’ edges of the sampling techniques, but at the expense of processing speed (some artifacts may result with 8-bit overflow).

The **I/O Optimization** preference allows the user to process large images even without access to large amounts of virtual memory when using the *Min Virtual Memory* option. This option saves many of the large interim results to temporary files. The optimization for *Processing Speed* preference can be utilized when an adequate amount of virtual memory is available to hold the interim processing results. This option greatly increases the processing speed since writing the temp files to the hard drive can be time consuming.

The **Sort LoG Extrema** preference can be changed to mitigate some of the harmful effects that caused by clouds within an image. Although LoGWaR is very effective in minimizing the effects of moving features from one dataset to the next (i.e. parallax, moving objects), clouds still pose a problem due to the well-defined edges that they produce within an image. This function allows the user to change the sorting procedure of the LoG derived extrema points from Max-to-Min to Min-to-Max. This can minimize the effects of cloud-cover by truncating the most well defined edges, but, keep the edges that have been selected by adaptive thresholding and fall within the *Max # Region Points* limit.

Exit:



The **Exit** function is utilized to gracefully exit from the LoGWaR program. This function overwrites any temporary files that were created when utilizing the *Min Virtual Memory option* within the user preferences.

REGISTRATION TOOLS Pull-down Menu Options:

Deleting ROI Matches:

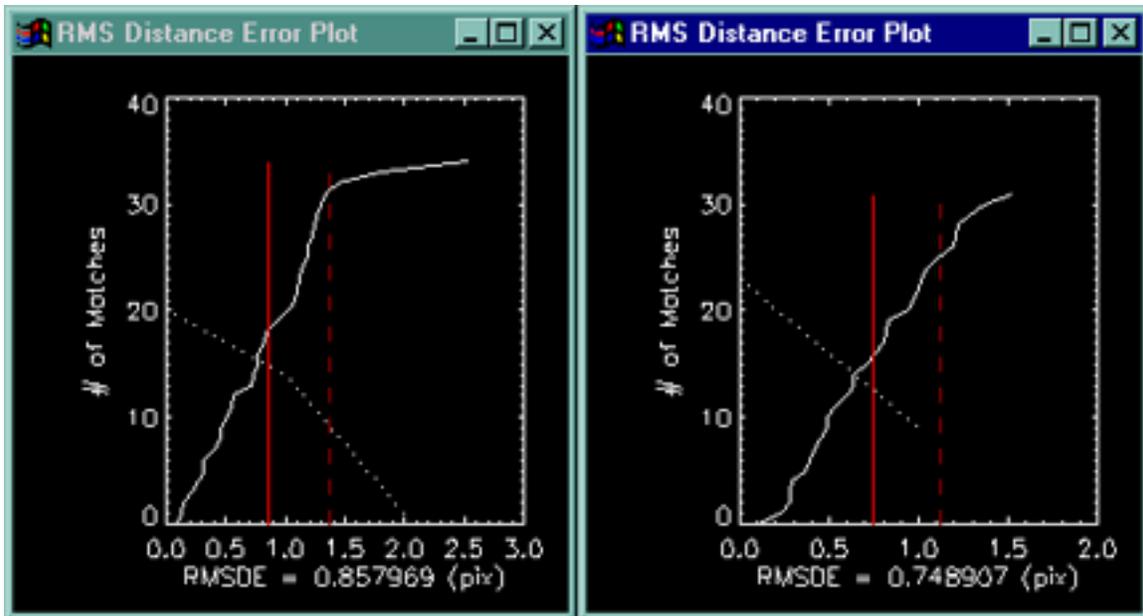


It is possible to create a Region of Interest (ROI) within LoGwaR, by depressing the left mouse button at one corner of the ROI and holding it down while moving the mouse icon to opposite corner of the ROI and releasing the left mouse button. Once the ROI has been defined (in either the Reference or Warp image), it is possible to delete any matches within this area through the use of this pull-down menu option.

Delete RMSDE > 1 Standard Deviation from Mean:



Before this function can be utilized it is necessary to compute the statistical accuracy of the model compared to each individual matched point (<Analyze> <Compute Matched RMS Error>). This function then allows the user to reject any matches that deviate from the transform model mean by more than one standard deviation (1STD). This technique is often quite useful in identifying anomalous matches and deleting them from the current matched points. Unlike the following technique, it does not recompute the model for each deleted match. It is possible to quickly reject many poor matches with this technique. Unfortunately, you may risk rejecting some good matches if there are bad matches that unduly skew the transform model. This technique is normally a good method to identify the “knee in the curve” for the RMSDE plot. A knee in the plot will normally indicate that there are some anomalous matches that could be removed to increase the transform model accuracy. The figure below shows how the 1STD (dashed line) threshold can be utilized to extract poor matches for rejection. Once these matches have been removed, the model is recomputed (<Analyze> <Compute Matched RMS Error>) for analysis. The RMSDE plot should appear fairly straight with cumulative error increasing linearly for accurate transform models.



The RMSDE plot shows the cumulative error (solid-white), Mean (red-solid), Mean + 1STD (red-dashed), and matched point error histogram (white-dotted).

Delete RMSDE > User Defined:



Before this function can be utilized it is necessary to compute the statistical accuracy of the model compared to each individual matched point (<Analyze> <Compute Matched RMS Error>). This function iteratively strips of the matched point that has the greatest RMSDE from the transform model and recomputes the model error after each deletion. This is repeated until the overall model error falls below the level dictated under the user preferences menu (<File> <User Preferences> [Max Allowed RMSDE]). The default value for this preference is set to 1 pixel and will ensure that the transform model has subpixel accuracy when this function is utilized. The following results are text-based output from the LoGwaR program. It demonstrates how the overall RMSDE Mean decreases monotonically with the iterative rejection of the matching point with greatest error from the transform model:

```

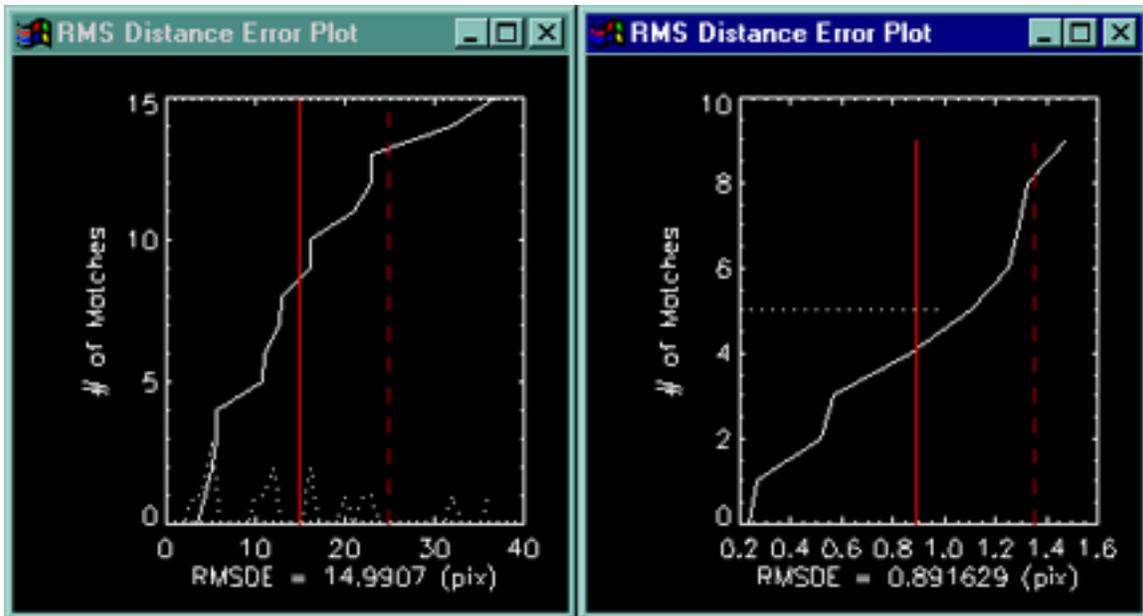
Number of Matches Loaded =      16
RMSDE Mean =      14.9907
Degree of Polynomial utilized for Transform =      1
RMSDE Mean =      12.0535

```

RMSDE Mean = 6.71098
 RMSDE Mean = 3.99343
 RMSDE Mean = 1.14675
 RMSDE Mean = 1.01661
 RMSDE Mean = 0.891629

 RMSDE SDev = 0.464510
 # of Good Matches = 10

The following figure plots the results for this example. Note the relatively linear character of the cumulative error (white-solid) for the final plot.



Pad Reference Image:



As the name suggests, this tool simply pads the Reference image with surrounding zeroes. The dimensions of the padding are dyadic to enable FWT analysis. This tool is useful increase the size of the reference image to preclude harsh cropping of the

Warped image. Since the dimensions of the Warped image are constrained to the dimensions of the Reference image, this tool can be utilized to increase the transformed image dimensions.

Rotate Image:



This function allows rudimentary rotation of both the Reference and Warp images. Special care must be taken when rotating images and using the predictive transform. ***Since any predictive transformations operate on the original images, it is important that the warp image rotation is executed before any FWT or scaling operations.***

This will ensure that the composite matrix multiplications are done correctly. For the same reason, it is important that the Reference image is not rotated when performing composite or predictive operations. Remember that if you do choose to rotate the Reference image, you should save this result as an interim file (<File> <Save Reference Image>) or the transformation results will not relate to the current image. Also, note that any 90 degree rotations do not change the data, they merely represent a swapping of columns for rows in the image array.

Histogram Matching:



Since LoGWaR attempts to register images based on detecting similar regions with high rates-of-variation, it is often useful to make these images appear as similar as possible. One useful technique to accomplish this task is to match the histogram distribution of grayscale values from the Reference image to the Warp image or vice-versa. This function allows for both of those possibilities.

Since the histogram match results are only used to determine matching GCPs and the resulting transformation coefficients, the nonlinear change to grayscale values is only temporary.

Inverting the Images:



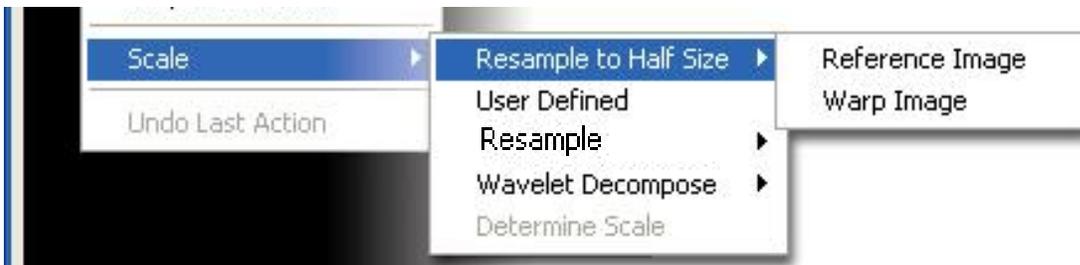
This function is only useful when attempting to work with image datasets that have been previously registered in ENVI with the resulting matched points file saved to a text file. If these images are subsequently brought into LoGWaR and the ENVI matched points file loaded; the GCPs will appear to be inverted due to the way that IDL loads in images bottom-to-top format vs ENVI's top-to-bottom format. For more info read the section above on *Saving Match Points*.

Crop Boxed Area:



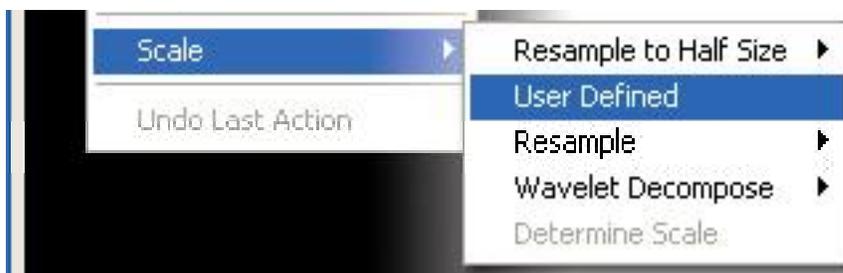
This function is very similar to most cropping function in today's of the shelf image processing software. Once a ROI has been defined, either in the Reference or Warp image window, the *Crop Function* can be utilized to trim the image to the desired dimensions. It is important to remember to save these results (<File> <Save Reference Image>) if the ensuing registration transformation is to be utilized for comparison to the Reference Image.

Scale Comparison:

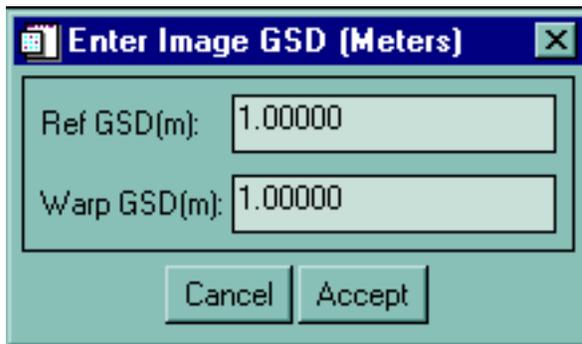


The **Resample to Half Size** option allows the user to down-sample (through pixel averaging) the Reference or Warp image to half of its current size for registration at a lower resolution. This function is very useful when working with large images and allows for the registration at a much lower resolution and eventual predictive transformation of the original image. Since the LoGWaR program automatically keeps track of any scale manipulations, it is possible to include this information into a composite affine transformation that can be utilized to register the original Warp image to the original Reference image based solely on the relationship between the low-resolution counterparts. This function can be repeated several times to compare images at half, quarter, eighth, sixteenth scale etc. It should be noted that this operations is the counterpart to the FWT decimation process. However, where the FWT conserves the high spatial frequency information for later use, this down-sampling procedure does not.

The **User Defined** scale function calls a GUI for the user to input the scale relationship

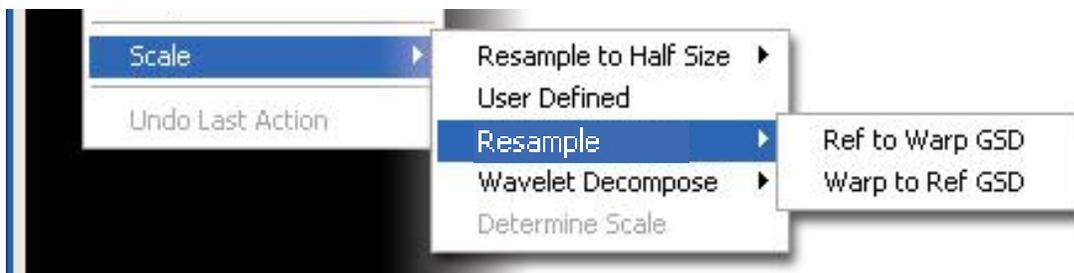


between the Reference and Warp images. It is not necessary to input the absolute scale relationship between the two images; only the



relative scale relationship needed. This scale relationship can then be utilized to resample the images to comparable dimensions or to a unique scale ratio, such as may be required for wavelet analysis (the FWT requires dyadic scale relationships). The following function can be utilized to scale the images to comparable resolutions.

The **Resample** scale function allows the works in conjunction with the previous function and allows the user to scale either the Reference or Warp image to the relative dimensions of the other. The sampling method (NN, BL, BC) is determined within the user preferences section of the *File* pull-down menu.

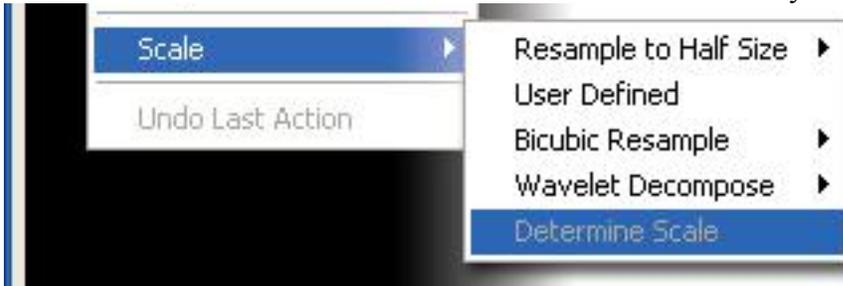


The **Wavelet Decompose** function allows scaling of the Reference or Warp images to the



closest dyadic scale as defined by the image dimensions and the user defined scale relationship. *Currently under construction!!!*

The **Determine Scale** function can be utilized to automatically determine the scale ratio



between the Reference and Warp images. This technique is based on the premise that although the distance may

change between related points, the ratio of distances will remain the same. Since the resolutions (or GSDs) of remotely sensed images are normally known, this function is often unnecessary. This function needs to be optimized and so is *currently under construction*.

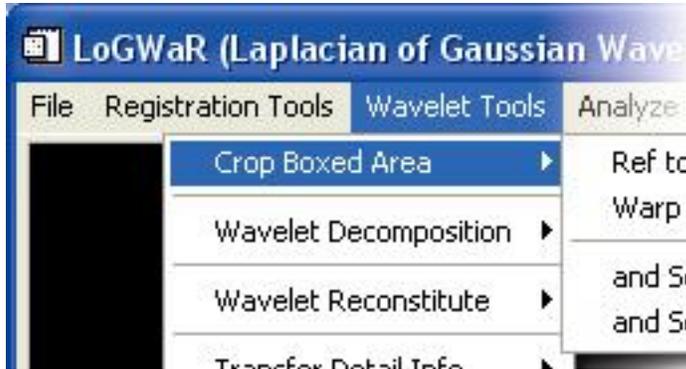
Undo Last Action:



The final function within the *Registration Tools* pull-down menu is the *Undo Last Action* feature. This option allows most actions to be undone, including changes to the Reference and Warp images and to the match points file. This feature will appear bold when available and subdued when unavailable.

Wavelet TOOLS Pull-down Menu Options:

Crop Boxed Area:



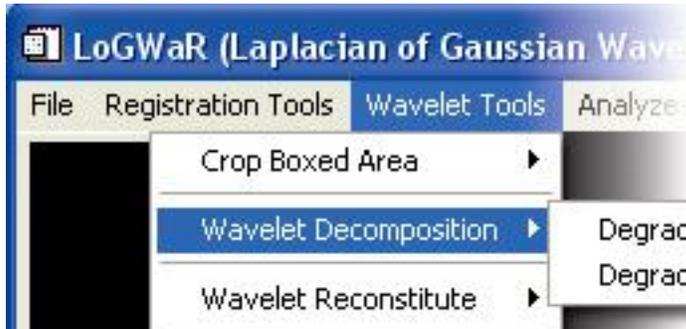
The *Ref/Warp to nearest Dyadic* function crops the reference/warp ROI to the closest Dyadic size

(powers of 2; i.e. 128, 256, 512,...). This is useful when trying to

identify an approximate ROI and still retain proper image dimensions for FWT analysis without changing the scale of the images.

The *Crop and Scale Ref/Warp to Dyadic* functions provide an exact cropping of the ROI area and then scales this region to the nearest dyadic size. Care must be taken with this function since it changes the relative resolution of the images through the scaling procedure.

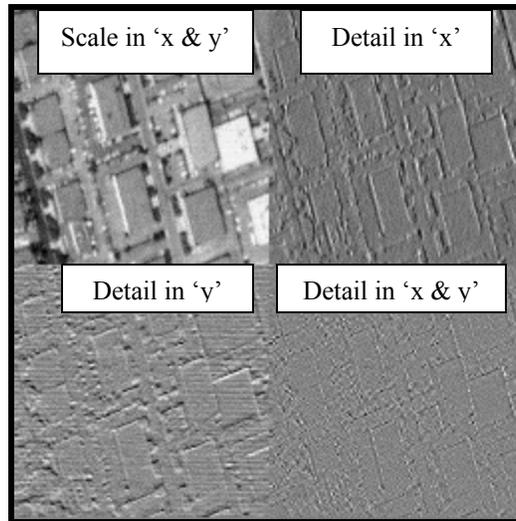
Wavelet Decomposition:



This function provides forward transform for the FWT Decimation of the Reference or Warp images. For all practical purposes, the image is decomposed into four component

'subbands'. The four components are the half-scale image, and the vertical, horizontal, and the diagonal high frequency components (edges) at the current resolution. These results can be portrayed in the Mallat representation of the image resolution pyramid as seen below. The scale component is simply the half-scale resolution image that is treated by LoGwaR as the new Reference or Warp working image. The current FWT level can be identified in the image title by the number in parenthesis. For example, if the reference image was 512x512, after one FWT decimation the title would appear as Reference Image (LL1): 256x256.

Mallat's Image Resolution Pyramid



Wavelet Reconstitute:



This function provides the inverse FWT of the Reference or Warp images. The strength of wavelet analysis with the FWT is the ability to easily move between image resolutions without loss of information. The *Wavelet Reconstitute* function

recombines the frequency subbands to obtain the 'sharpened' Reference or Warp image at twice the resolution. This process can only be utilized if the image was previously decimated or if high frequency information was transferred in from another image.

Transfer Detail Information:

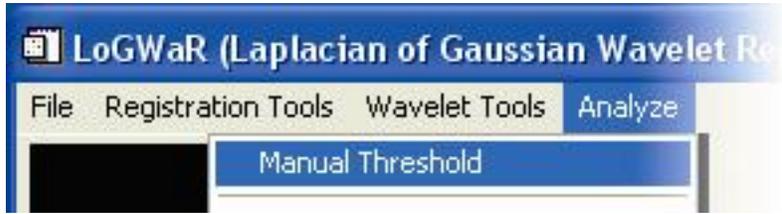


The ability to transfer high frequency information from the Reference image to the Warp image or vice-versa is accomplished with this function. This function can be utilized as long as the images have the same dimensions and spatial frequency information is available for transfer. high-frequency

This process is utilized to transfer information for eventual sharpening applications. This FWT sharpening technique is shown to retain overall radiometric integrity.

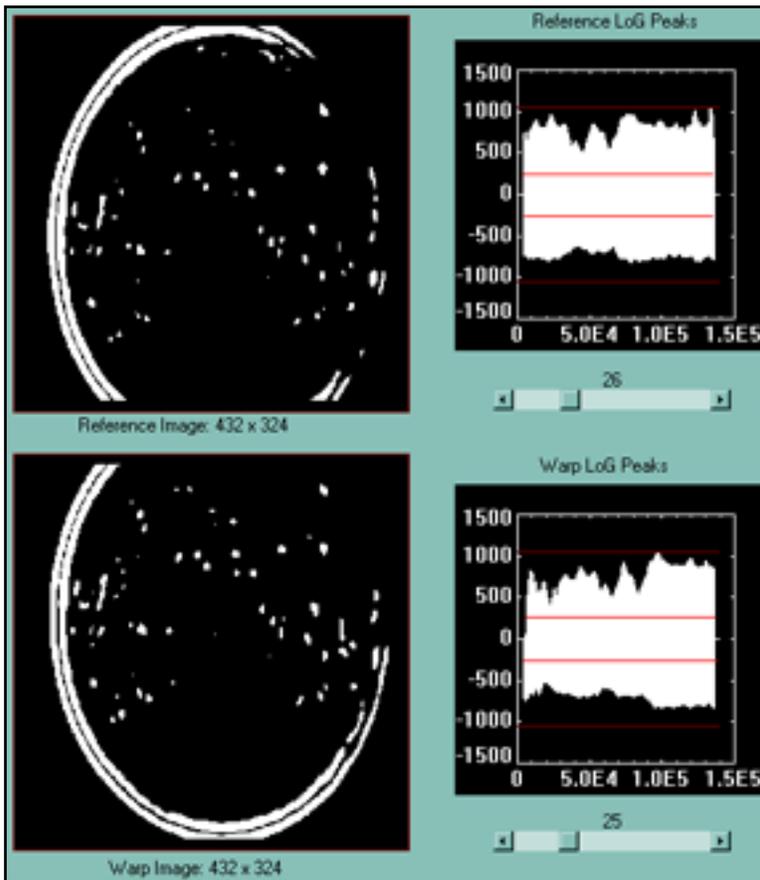
Analyze Pull-down Menu Options:

Manual Threshold:



The *Manual Threshold* procedure is a semi-automated process that requires the user to move the threshold slider-bars until the point sets

(connected components regions) within the Reference and Warp have similar content.



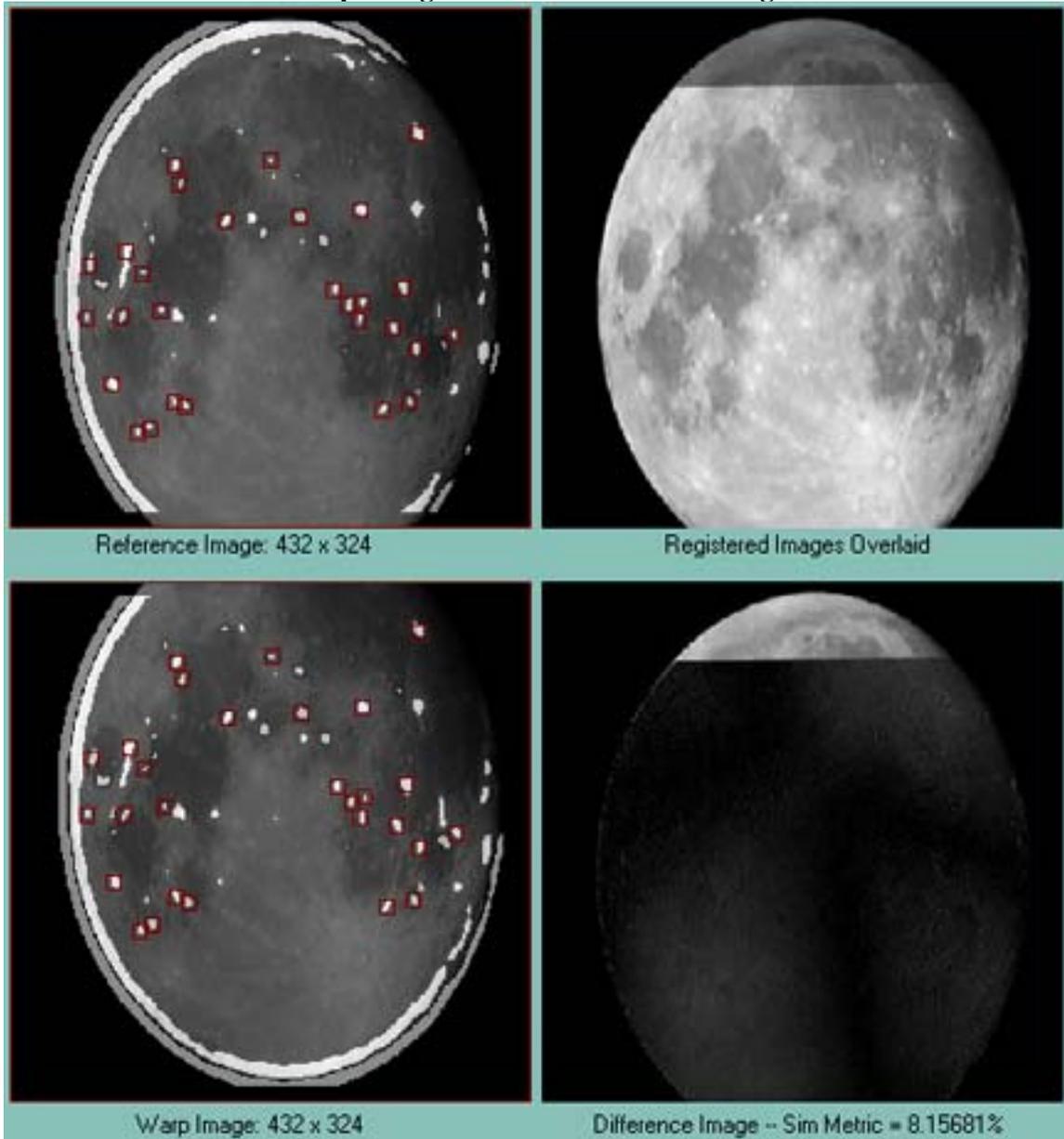
This was achieved in the figure (to right) by sliding the Reference threshold bar to 26% and the Warp threshold bar to 25%. Out of these regions, LoGWaR will isolate the peak pixels in each region and limit the point sets to the 50 most extreme LoG filtered values.

This example registration was accomplished on two separate moon-shots taken very close in time. The edge of the moon against the black background provided an excellent edge that required the entire threshold to come down to a very low level. Note that only one peak is identified for each region...including

the edge of the moon.

It is possible to see the matching points that were derived from the threshold regions (by the LoGWaR program) from the following figure. The resulting overlaid/stacked images, difference image and similarity metric are also displayed for reference. The manual technique works quite well due to the ability of the human visual system (HVS) to identify similarity in point sets. For difficult image registration problems, it is often useful to manipulate the manual threshold sliders to get an idea of how the LoG Filtering and thresholding process is isolating similar regions within the scenes. The manual thresholding procedure can be accomplished on ROIs as well as image wide.

Example Registration of two moon images.



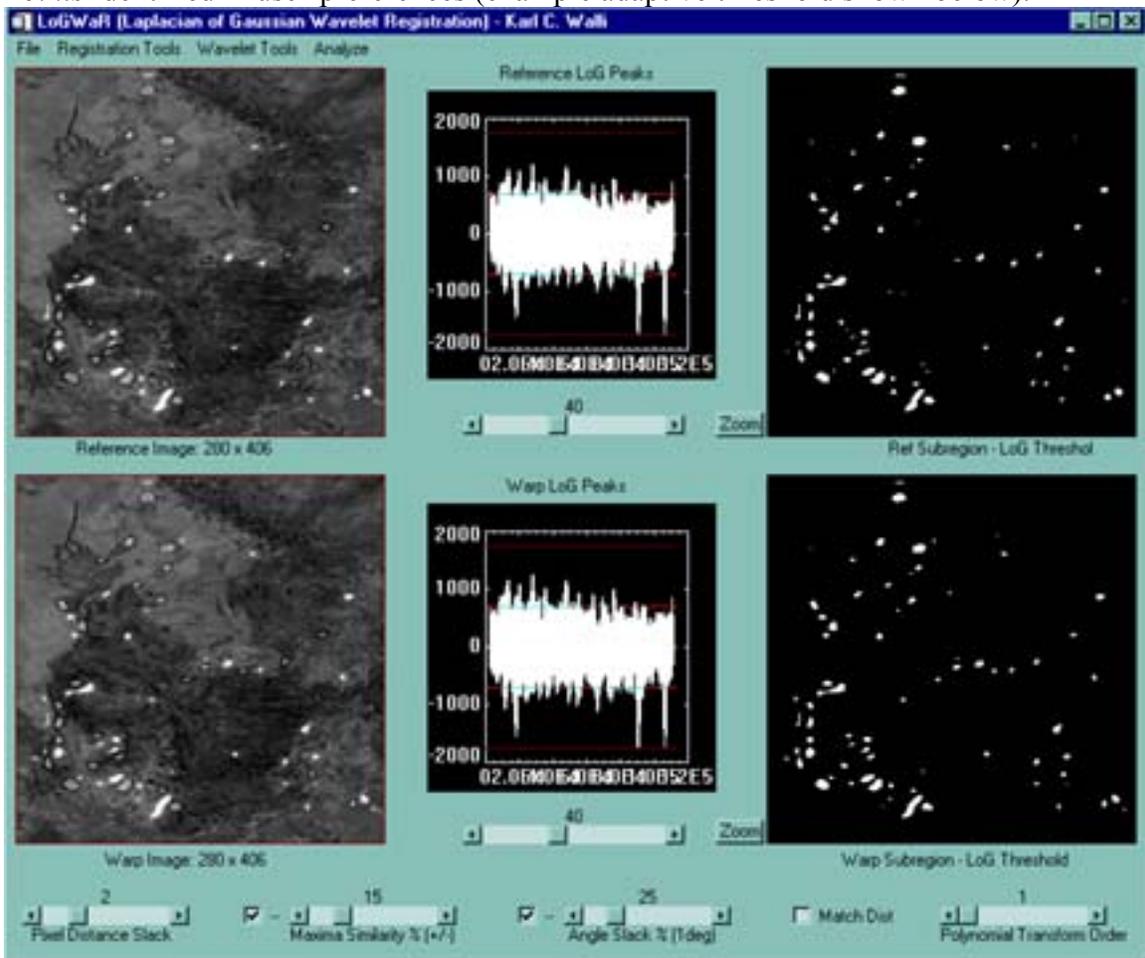
Auto Threshold Processes:



The automatic threshold processes attempt to replicate the manual process identified above with three separate functions.

The **Preset – Imagewide** technique utilizes the user-preference’s *Threshold Level (%)* value to automatically threshold both images at that level. This technique is most useful when dealing with datasets like FITS images that often need very low threshold values (1%) to bring out the maximum number of threshold regions.

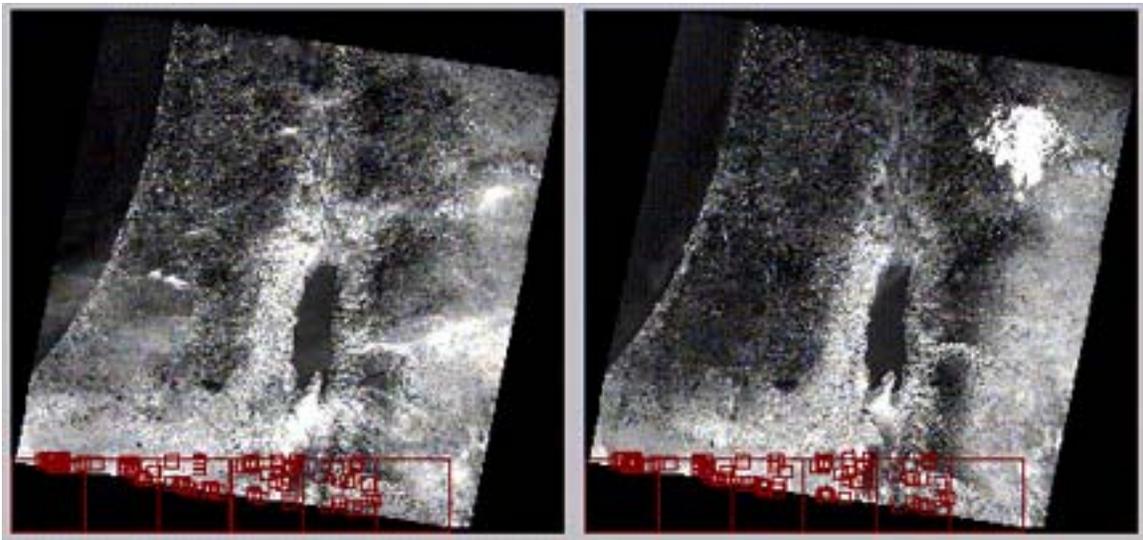
The **Adaptive –Imagewide** technique is an adaptive threshold procedure that requires LoG filtering of the entire image. For this reason, the image should be no larger than what can be efficiently filtered with resident virtual memory. The LoGWaR program will not provide imagewide filtering on datasets that are larger than what is prescribed in the user-preference’s *Max LoG Filtering Size* value. In order to accomplish image-wide filtering on large datasets it is often necessary to either reduce the size through FWT decimation or traditional downsampling. This technique is referred to as *Subband* processing due to the cutoff of high frequency spatial information from the image. Adaptive thresholding is utilized to slowly lower the threshold level in order to isolate more potential GCPs. Once the number of isolated regions grows above the user-preference’s *Max # Region Points*, independently in both the Reference and Warp images, the adaptive thresholding stops. The isolated regions are then converted to maxima pixels, sorted based on LoG value, and then truncated to the *Max # Region Points* identified in user-preferences (example adaptive threshold shown below).



The **Adaptive - Subregion** technique relies on the same procedure as the *Adaptive – Imagewide* process except that it does not filter the whole image at one time. Instead, it divides the image into manageable pieces and independently derives matched GCPs for each region. The ROI size is obtained from the user-preference’s *Subregion Size* value. The LoGwaR program will then isolate no more than the *Max # Region Points* from each ROI. After ‘walking’ through the entire images, the matches are combined into one master GCP match set.

This technique is referred to as *Subregion* processing due to the independent analysis of *related* ROI within the datasets. **For this reason it is necessary that the two images have good rotational agreement before the *Adaptive – Subregion* technique can be utilized effectively.** If they do not, it is first necessary to first rotate the images or determine a rough estimate of the transform through Subband analysis, apply the predicted transform and then re-register with the Subregion technique. LoGwaR allows the user to perform independent affine transformations (such as rotations, scaling, and affine transforms) and automatically logs these manipulations for future inclusion into a composite transformation. The interim rotations and scales are computed automatically at each stage, as affine transformations, and later combined to produce the composite transform for the original image. This final composite transform only needs to be applied once to the original image and mitigates unnecessary sampling degradation.

LoGwaRs Adaptive-Subregion Technique LANDSAT MS (8k x 8k), Jericho, Israel



Register from Matched Points:



Once matched GCPs have been derived from LoGWaR, it is possible to transform the Warp image with either an affine or polynomial model. The affine models are a subset of the polynomial transforms that include shift, rotation, scale, and skew. They also

have a special commutative property that allows combination through multiplication.

The **Affine Warp (4pts)** function determines the four GCP matches with the greatest spread in the horizontal and vertical directions. This allows for a quick estimate of the affine without utilizing the Psuedo-Inverse solution, which is used for most over determined problems within LoGWaR. This procedure is primarily utilized to check results of alternative solutions.

The **Affine Warp (all pts)** function is the ‘workhorse’ for the LoGWaR program. Since most registrations produce several GCPs (possibly hundreds), the Psuedo Inverse solution to the Linear Least Squares problem is utilized. This is where the GCPs that have been generated are transformed into a global equation that can be utilized to relate the two images. The following equations show the relationship between the GCPs and the affine solution via the Psuedo Inverse technique:

$$\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \\ x_4 & y_4 & 1 \end{bmatrix} = \begin{bmatrix} x'_1 & y'_1 & 1 \\ x'_2 & y'_2 & 1 \\ x'_3 & y'_3 & 1 \\ x'_4 & y'_4 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\begin{aligned} X &= UA \\ U^{-1}X &= A \\ (U^T U)^{-1}U^T X &= A \end{aligned}$$

It should be noted that an affine solution is required for any composite/predictive transformation. If an affine solution is inadequate to relate the two datasets it may be necessary to produce an affine estimate with a polynomial model to reduce any sampling degradations to two operations. The image transformation is done with the POLY_WARP function, which utilizes the derived affine coefficients (stored in a 3 x 3 matrix), but reformats them into a polynomial expression to transform the Warp image.

The **Polynomial Warp** function allows for a much higher degree relationship to be defined between the Reference and Warp images. The degree of the polynomial is defined by the *Polynomial Transform Order* slider-bar, which is located in the lower right-hand corner of the LoGWaR GUI. Although the LoGWaR program can only automatically register images with an affine relationship (minus skew), it is still possible to relate regions at the affine level and solve for higher order relationships image-wide. This is because related ROIs may have affine relationships even though the entire dataset may require a higher order relationship. Most remotely sensed image, that have been corrected for sensor artifacts, can be related with no more than a 2nd degree polynomial expression. The polynomial expressions, within LoGWaR, are derived primarily from IDL's *POLY_2D* function. This function also utilizes the Psuedo Inverse solution to the least squares problem. The image transformation is done with the *POLY_WARP* function, which uses the derived polynomial expression to transform the Warp image.

Register from Coefficients:



The LoGWaR program allows user to save and load matched GCP files as well as transform coefficients to warp images. This function allows user to transform the Warp image utilizing either an affine or polynomial model that has been loaded from a saved file. This can be very useful when utilizing the predictive

and/or composite affine transforms that contain information that is not easily converted into matched GCP form. Polynomial coefficients that have been saved from ENVI can also be loaded into LoGWaR to transform the Warp image.

Composite Affine Transform:



The **Update Composite** function allows the latest affine transform to be included into the composite affine structure with any rotation and scale information. The 3 x 3 affines containing this information are then combined through multiplication to produce a composite affine transform. This formulation can be visualized in the example below:

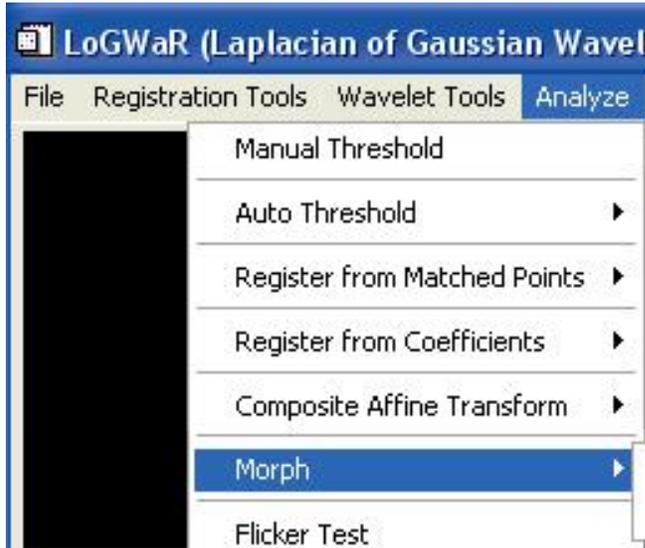
$$\begin{array}{c}
 \textit{Rotation Affine} \qquad \qquad \qquad \textit{Latest Affine} \qquad \qquad \qquad \textit{Scale Affine} \\
 \\
 \textit{Comp_Trans} = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 \\ -1.0000 & 0.0000 & 0.0000 \\ 2047.00 & 0.0000 & 1.0000 \end{bmatrix} \begin{bmatrix} 1.01672 & 0.00112421 & 0.0000 \\ -0.00862428 & 0.998749 & 0.0000 \\ -0.224098 & 0.0169114 & 1.0000 \end{bmatrix} \begin{bmatrix} 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 8.0000 \end{bmatrix} \\
 \\
 \textit{Comp_Trans} = \begin{bmatrix} -0.00112421 & 1.01672 & 0.0000 \\ -0.998749 & -0.00862428 & 0.0000 \\ 2046.86 & -1.79278 & 1.0000 \end{bmatrix}
 \end{array}$$

This example demonstrates the rationale for initiating any rotations before scaling operations. Due to the order in which these matrices are combined, it is essential that any rotations are accomplished first, in order to properly account for the dimensions of the original image.

The LoGWaR program also keeps tabs on whether the Warped image has been loaded as the current Warp image. If it has, the composite transform will retain the old composite affine for possible incorporation into the latest composite structure. This technique is useful when utilizing a supervised selection of GCPs to compensate for higher order deformations that cannot be automatically detected by the LoGWaR point matching scheme. For example, to correct for skew and perspective, it is only necessary to pick four matching GCPs to correct for the 1st order polynomial required to solve for these unknowns. This 1st order equation can also be incorporated into a 3 x 3 formulation for incorporation into the composite structure. Although 1st order polynomials do not have the commutative property of affines, the combination of only one 1st order with an affine has shown good composite transformation results.

Once the *Update Composite* function has incorporated the affine operations into a composite transform, the **Apply Composite** function applies this transform to the original Warp Image.

Morph:



The *Ref to Warp (4 corners)* function provides a visual inspection of how the Reference and Warp images relate.

The *Ref to Warped (match pts)* technique first requires the Reference and Warp image to be registered with matched GCPs. These matches are then utilized, to sequentially morph the Warped image to the Reference image coordinate system.

Flicker Test:



The *Flicker Test* often provides the best visual inspection the registration results. Subtle differences can often be detected

when quickly flickering between the Reference image and the recently Warped image. This is why compression testing often utilizes the *Flicker Test* to detect artifacts and other anomalies.

Compute Matched RMS Error:



This function provides a quantitative analysis of how well the LoGwaR derived matched

GCPs agree with a 1st or 2nd Degree Polynomial model. Once the image-wide model is determined from all the matches, each Warp image match point is compared against the model prediction. Any variation between the LoGwaR derived Warp location and the polynomial model prediction is captured in the RMS Distance Error (RMSDE) calculation. These individual RMSDE values (in the 'x' and 'y') are accumulated and combined into an average RMSDE value for the entire image. This useful metric can be utilized to judge the overall agreement of the matches to a mathematical model and is often a very good judge of registration accuracy. If subpixel accuracy is required for a registration, it is simply a matter of obtaining an average RMSDE value less than one pixel through iterative rejection of matches with greatest error. This function plots the RMS error (cumulative and histogram) and the LoGwaR derived vs. model predicted matched point comparison.