

026097

An application of rotation- and translation-invariant overcomplete wavelets to the registration of remotely sensed imagery

Jacqueline Le Moigne¹
Ilya Zavorine²

1. Applied Information Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771
2. Department of Mathematics, University of Maryland, College Park, MD 20742
and USRA/CESDIS, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

ABSTRACT

A wavelet-based image registration approach has previously been proposed by the authors. In this work, wavelet coefficient maxima obtained from an orthogonal wavelet decomposition using Daubechies filters were utilized to register images in a multi-resolution fashion. Tested on several remote sensing datasets, this method gave very encouraging results. Despite the lack of translation-invariance of these filters, we showed that when using cross-correlation as a feature matching technique, features of size larger than twice the size of the filters are correctly registered by using the low-frequency subbands of the Daubechies wavelet decomposition. Nevertheless, high-frequency subbands are still sensitive to translation effects. In this work, we are considering a rotation- and translation-invariant representation developed by E. Simoncelli and integrate it in our image registration scheme. The two types of filters, Daubechies and Simoncelli filters, are then being compared from a registration point of view, utilizing synthetic data as well as data from the Landsat/ Thematic Mapper (TM) and from the NOAA Advanced Very High Resolution Radiometer (AVHRR).

1. INTRODUCTION

Automatic registration and resampling of remotely sensed data will be an essential element of future Earth satellite observation systems. New remote sensing systems will generate enormous amounts of data representing multiple observations of the same features at different times and/or by different sensors with, most often, these sensors being spread over multiple platforms. Automatic registration and resampling methods are indispensable for such tasks as data fusion, navigation, achieving super-resolution, or optimizing communication rates between spacecraft and ground systems. For all these tasks, accurate image registration is the first step, since a number of distortions prevent two images acquired either by the same sensor at different times or by two sensors at the same or different times from being "perfectly registered" to each other or to a fixed coordinate system. Distortions usually correspond to orbit and attitude anomalies, but some continuous nonlinear distortions are also due to altitude, velocity, yaw, pitch, and roll. It is very difficult to determine exact location within an image using only ancillary data and geo-location is usually computed by combining *navigation* and *registration*. Navigation corresponds to a "systematic correction" based on image acquisition models taking into account satellite orbit and attitude, sensor characteristics, platform/sensor relationship, Earth surface and terrain models and brings the registration accuracy within a few pixels. Image registration, on the other hand, corresponds to a "precision correction" based on landmarks and image features, and refines the geolocation to a subpixel accuracy. Registration is either applied after the navigation process, or both processes are integrated in a closed feedback loop. In this paper we will only consider the issue of feature-based, precision-correction automatic image registration.

In general, image registration can be defined as the process which determines the best match of two or more images acquired at the same or at different times by different or identical sensors. One set of data is taken as the *reference data*, and all other data, called *input data (or sensed data)*, is matched relative to the reference data. The general process of image registration includes three main steps: (1) the extraction of features to be used in the matching process, (2) the feature matching strategy and metrics, and (3) the resampling of the data based on the correspondence computed from matched features. This paper only deals with steps (1) and (2). Currently, the most common approach to satellite image registration is to perform step (1) manually by interactive extraction of a few outstanding characteristics of the data, which are called *control points* (CP's), *tie-points*, or *reference points*. The CP's in both images (or image and map) are matched by pair and used to compute the parameters of a geometric transformation. But such a point selection represents a repetitive, labor- and time-intensive task which becomes prohibitive for large amounts of data, and often leads to large registration errors [1]. A large number of automatic image registration methods have been proposed and surveys can be found in [2-4]. Some of the features which are being utilized for step (1) are: original gray levels, edges, regions, and more recently wavelet features. According to [2], step (2) itself can be separated into:

- the *search space*, i.e. the class of potential transformations that establish the correspondence between input data and reference data. Transformations that are often used are rigid transformations (composed of a scaling, a translation and a rotation), affine transformations (composed of a rigid transformation, a shear and an aspect-ratio change; a shear in the x-axis transforms the x-coordinate into a linear combination of both x and y-coordinates, and the aspect-ratio is defined as the numerical ratio of image width to height), and polynomial transformations.
- a *search strategy*, which is used to choose which transformations have to be computed and evaluated. Local or global search, multi-resolution search or optimization techniques are examples of various search strategies.
- a *similarity metric*, which evaluates the match between input data and transformed reference data for a given transformation chosen in the search space. Correlation measurement has been the most often used but other methods such as a Hausdorff distance [5] can also be utilized.

A wavelet-based image registration approach has previously been proposed by the authors [4,6,7]. In this work wavelet coefficient maxima, obtained from an orthogonal wavelet decomposition using Daubechies filters [8], were utilized to register images in a multi-resolution fashion. Tested on several remote sensing datasets, this method gave very encouraging results. In the study reported in [9], we showed that when using cross-correlation as a feature matching technique, features of size larger than twice the size of the filters are correctly registered using the low-frequency subbands of the Daubechies wavelet decomposition. Nevertheless, features extracted from the high-frequency subbands are still sensitive to translation effects.

In this work, we are utilizing filters developed by E. Simoncelli [10-12] and we integrate them in our wavelet-based image registration scheme. The two types of filters, Daubechies and Simoncelli filters, are then being compared from a registration point of view, utilizing synthetic data as well as data from the Landsat/ Thematic Mapper (TM) and from the NOAA Advanced Very High Resolution Radiometer (AVHRR). Results are presented in section 3.

2. SOME BACKGROUND ON WAVELET-BASED IMAGE REGISTRATION OF SATELLITE IMAGERY

2.a Previous Wavelet-Based Image Registration Method

Wavelet transforms provide a space-frequency representation of an image. In a wavelet representation, the original signal is filtered by the translations and the dilations of a basic function, called the "mother wavelet". In our wavelet-based registration, only discrete orthonormal bases of wavelets have been considered and are implemented by filtering, separately in rows and in columns, the original image by a high-pass and a low-pass filter, thus in a multi-resolution fashion [13]. At each level of decomposition, four new images are computed; each of these images is a quarter the size of the previous original image and represents the low frequency or high frequency information of the image in the horizontal and/or the vertical directions; images LL (Low/Low), LH (Low/High), HL(High/Low), and HH (High/High). Starting again from the "compressed" image (or image representing the low-frequency information, "LL"), the process can be iterated, thus building a hierarchy of lower and lower resolution images. Figure 1 summarizes the multi-resolution decomposition.

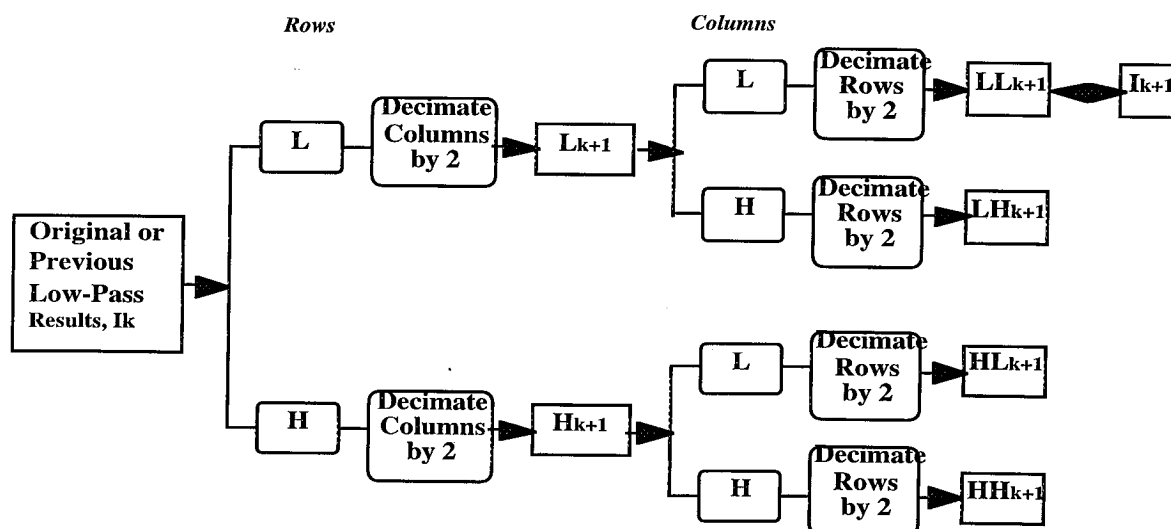


Figure 1
Decomposition by an Orthonormal Basis of Wavelets

Our wavelet-based method represents a three-step approach to automatic registration of remote sensing imagery. The first step involves the wavelet decomposition of the reference and input images to be registered. In the second step, we extract at each level of decomposition domain-independent features from both reference and input images. Finally, we utilize these features to compute the transformation function by following the multiresolution approach provided by the wavelet decomposition. Following the registration framework described in the previous section, our algorithm will utilize the four following components:

- *The feature space*

Features are either chosen as the gray levels provided by the low-frequency LL compressed versions of the original image (for non-noisy images), or are based on the high-frequency information (e.g., maxima points of LH and HL images) extracted from the wavelet decomposition. In this second option, only those points whose intensities belong to the top $x\%$ of the histograms of these images are kept (x being a parameter of the program whose selection can be automatic); we call these points "maxima of the wavelet coefficients." These maxima are computed for all levels of the wavelet decomposition, for reference as well as input images.

- *The search space*

In a first step, we assume the transformation to be either a rigid or an affine transformation. Both types of transformations include compositions of translations and rotations; therefore, as a preliminary study, our search space is composed of 2-D rotations and translations, and will be extended later to rigid and affine transformations. We look for rotations with angles included in the interval $[0, 90\text{degrees}]$ and for translations in the interval $[0, \text{half pixel-size of reference image}]$.

- *The search strategy*

Our search strategy follows the multi-resolution wavelet decomposition, iteratively from the deepest level of decomposition (where the image size is the smallest), until the first top level of decomposition, i.e. going from low resolution up to high resolution. For all levels of decomposition, the subband images of the reference image are successively transformed by all the transformations included in the search space. Then maxima of the transformed reference wavelet images and of the input wavelet images are extracted. The accuracy of this search increases when going from low resolution to high resolution. At each level the search focuses in an interval around the "best" transformation found at the previous level with an accuracy D and is refined at the next level up with an accuracy $D/2$. More details on this algorithm can be found in [4,6,7].

- *The similarity metric*

At each level of decomposition and for each of the transformations, a correlation measure is computed between transformed reference maxima and input maxima. Another measure, based on a generalized Hausdorff distance, has also been studied and very encouraging results are reported in [5].

The previous algorithm is summarized in Figure 2. Tested on several datasets, the wavelet-correlation-based method described in this section performs with an accuracy of 1.66 pixels [14]. When using a statistically robust matching method based on a generalized Hausdorff distance, the first tests show that a subpixel accuracy can be obtained. More details on the results can be found in [4,14,15].

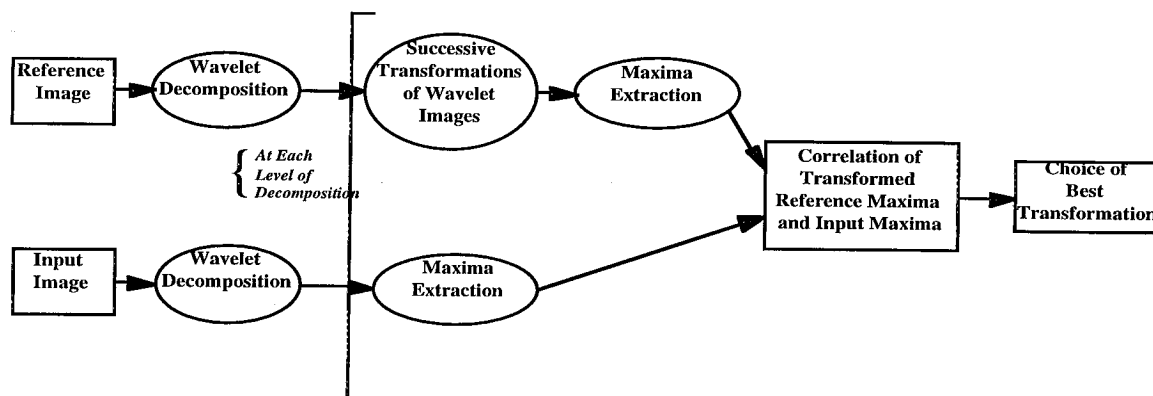


Figure 2
Summary of Our Wavelet-Correlation Image Registration Method

2.b Rotation and Translation Invariance Issues

According to the Nyquist criterion, in order to distinguish between all frequency components and to avoid aliasing, the signal must be sampled at least twice the frequency of the highest frequency component. Therefore, as pointed out in [10],

“translation invariance cannot be expected in a system based on convolution and subsampling.” When using a separable orthogonal wavelet transform (described in Figure 1), information about the signal changes within or across subbands. By lack of translation invariance, we mean that the wavelet transform does not commute with the translation operator, and similar remarks can be made relative to the rotation operator. Following these remarks, we conducted a study where the use of wavelet subbands is quantitatively assessed as a function of features’ sizes. The study reported in [9] shows that when using cross-correlation, the method described in 2.a is still a useful registration scheme in spite of translation effects. The results are summarized here, see [9] for more details:

- the low-pass subband is relatively insensitive to translation, provided that the features of interest have an extent at least twice the size of the wavelet filters.
- the high-pass subband is more sensitive to translation, but the peak correlations are still high enough to be useful.

Following this study, the work presented in this paper only considers the high-pass subbands and look at rotation- and translation-invariant filters [10] in order to create the feature space. Although the scheme described in Figure 2 would be more optimal if a different similarity metrics were used, we keep the same correlation framework for the only purpose of comparing Daubechies and Simoncelli’s filters under the same conditions and for registration purposes. Experiments involving a different search strategy and different similarity metrics are currently being performed as a continuation of this work.

3. USE OF A ROTATION- AND TRANSLATION-INVARIANT REPRESENTATION FOR IMAGE REGISTRATION

3.a Rotation- and Translation Invariant Representation

The method described by E. Simoncelli in [10-12] enables to build translation- and rotation-invariant filters by relaxing the critical sampling condition of the wavelet transforms. By invariance, it is meant that the information contained in a given subband will be invariant to translation or rotation. The resulting representation is equivalent to an overcomplete wavelet transform, it is not an orthogonal representation but is an approximation of a “tight-frame” [8], i.e. invertible. The Steerable Pyramid described in [10] is summarized in Figure 3, where only the analysis decomposition is shown. $H0$ is a high-pass filter, $L0$ and $L1$ are two low-pass filters, and $\{B0, \dots, Bk\}$ represents a set of oriented band-pass filters which ensure the representation to be rotation-invariant. In order to ensure translation-invariance, the output of the high-pass filter and of the band-pass filters are not subsampled. In addition, the portion of the signal which is iteratively decomposed by the band-pass and the low-pass filters does not contain the larger high frequency components and has been preprocessed by the low-pass filter, $L0$, thus removing most aliased components.

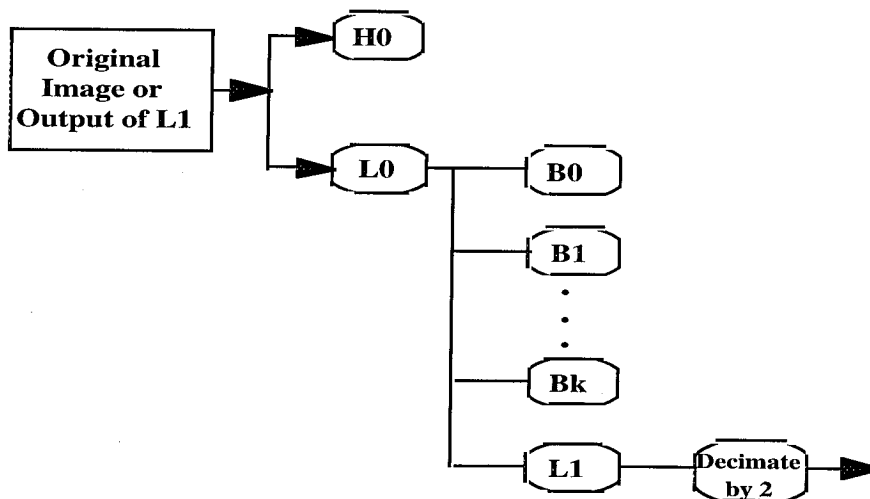


Figure 3
Decomposition by a Steerable Pyramid

As stated in [10], this representation is overcomplete by a factor of $4k/3$, where k is the number of oriented band-pass filters. In the experiments shown below, in order to optimize the computational speed, we chose $k=1$. The decomposition was iterated 3 times and the subbands which were considered for feature selection in the registration algorithm are $\{S0, S1, S2, S3\}$ as shown in Figure 4.

grid that extends from -30.20 S, 15.39 E (upper left) -34.79 S, 24.59 E (lower right). The navigation process uses an orbital model developed at the University of Colorado [16] and assumes a mean attitude behavior (roll, pitch and yaw) derived using Ground Control Points [17]. A map of the coastline derived from the Digital Chart of the World (DCW) is generated for the same geographic grid and is also available for this dataset. In this case, the actual transformation is not known, but results of a manual registration are used to assess the accuracy of the automatic registration.

3.b.3 Results

Registration results obtained with the two different types of filters are summarized in Tables 1 and 2 and graphically represented in Figure 10. These results show that the two types of filters give similar results for ideal or low-noise images but as soon as the noise level increases, the registration accuracy is much more stable with Simoncelli's filters. Overall the translation accuracy obtained with these filters stays around 1 pixel, and even reaches subpixel accuracy for the "GIRL" and "TM" datasets; while the accuracy using Daubechies' filters greatly varies depending on the contents of the images. The results are consistent for all levels of thresholds chosen in the maxima selection, even when the noise level increases.

4. CONCLUSION AND FUTURE WORK

This paper presented an image registration method based on wavelets and overcomplete wavelets. It was shown that, as expected and due to their translation- and rotation- invariance, Simoncelli's filters perform better than Daubechies' filters. Quantitative measurements support this conclusion.

As we recognize that the exhaustive search involving multiple cross-correlations is not optimal, we are looking at other search strategies and similarity metrics, such as optimization and robust matching. Future work will also exploit the full rotation-invariance capability of the steerable filters by varying the number of band-pass filters.

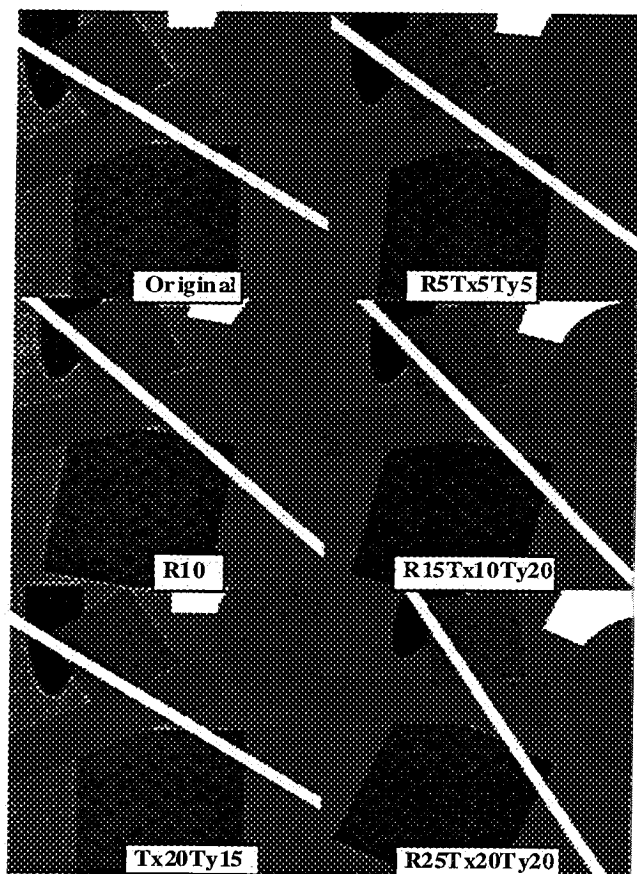


Figure 5
Dataset #1 ("SYNTH")
Reference and Some Transformations

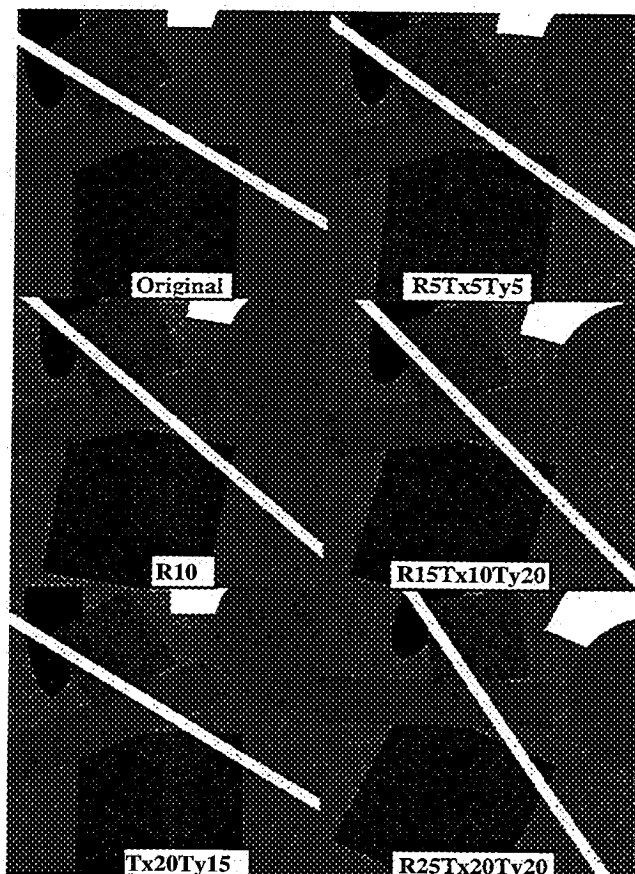


Figure 6
Dataset #5 ("NSYNTH.G20")
Reference and Some Transformations

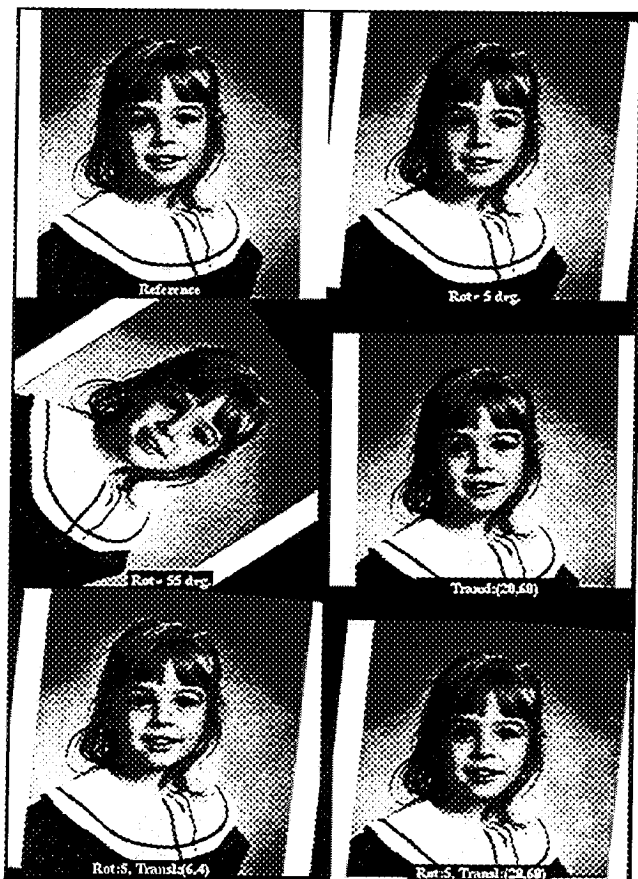


Figure 7 - Dataset #6 ("GIRL")
Reference and Some Transformations

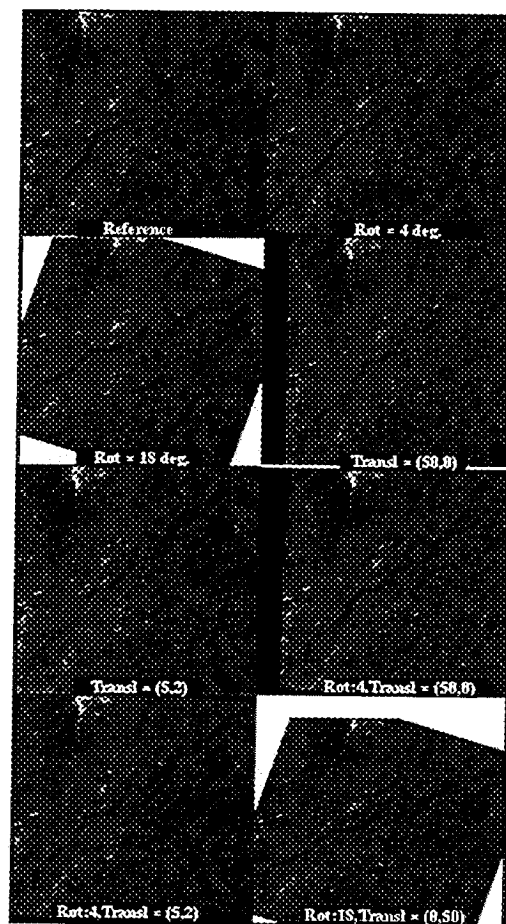


Figure 8 - Dataset #7 ("TM")
Reference and Some Transformations

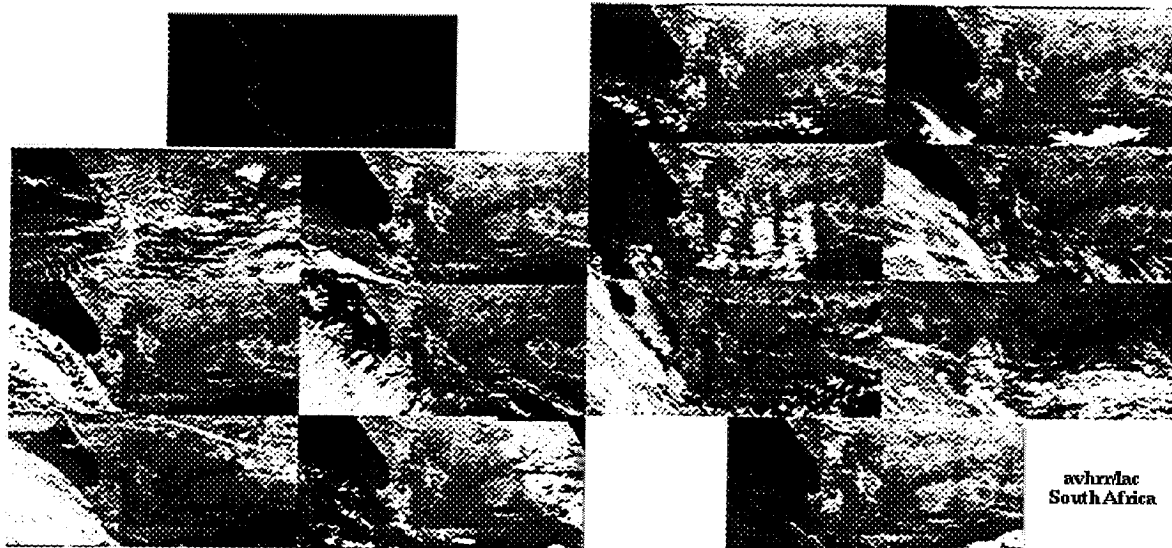


Figure 9 - Dataset #8 ("AVHRR")
Series of Thirteen Multi-Temporal AVHRR Images

ACKNOWLEDGMENTS

The authors wish to acknowledge Dr. Milton Halem for his support to this work, as well as the continuing support of the Applied Information Sciences Branch (NASA/Code 935) and of the Center of Excellence in Space and Data Information Sciences (CESDIS). The authors would also like to thank Pr. Eero Simoncelli for a very useful discussion about his research and his useful comments about our work.

REFERENCES

1. J. R. G. Townshend, and C.O. Justice, "The Impact of Misregistration on Change Detection," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30-5, 1992.
2. L. Brown, "A Survey of Image Registration Techniques," *ACM Computing Surveys*, vol. 24, no.4, 1992.
3. L.M.G. Fonseca and B.S. Manjunath, "Registration Techniques for Multisensor Sensed Imagery," *Photogrammetric Engineering and Remote Sensing Journal*, 62 (9), pp. 1049-1056, Sept. 1996.
4. J. Le Moigne, W.J. Campbell, and R.F. Crompt, "An Automated Parallel Image Registration Technique of Multiple Source Remote Sensing Data," *CESDIS TR-96-182* - submitted to *IEEE Transactions in Geoscience and Remote Sensing*.
5. D.M. Mount, N.S. Netanyahu, and J. Le Moigne, "Efficient Algorithms for Robust Feature Matching," to appear in *Special Issue of Pattern Recognition on Image Registration*, 1999.
6. J. Le Moigne, "Parallel Registration of Multi-Sensor Remotely Sensed Imagery Using Wavelet Coefficients," *SPIE's OE/Aerospace Sensing, Wavelet Applications Conference*, Orlando, April 5-8, 432-443.
7. J. Le Moigne, "Towards a Parallel Registration of Multiple Resolution Remote Sensing Data," *IGARSS'95, International Geoscience and Remote Sensing Symposium*, Firenze, Italy, July 10-14, pp.1011-1013.
8. I. Daubechies, *Ten Lessons on Wavelets*, CMBS-NSF Series Applications in Mathematics, SIAM, 1991.
9. H Stone, J. Le Moigne, and M. McGuire, 1997, "Registration Using Wavelet Techniques," *26-th Applied Imagery Pattern Recognition Workshop (AIPR'97)*, Washington DC, Oct. 15-17, 1997, submitted to IEEE-PAMI.
10. E. Simoncelli, W. Freeman, E. Adelson, and D. Heeger, "Shiftable Multiscale Transforms," *IEEE Transactions on Information Theory*, Vol. 38, No. 2, March 1992.
11. E. Simoncelli, and W. Freeman, "The Steerable Pyramid: A Flexible Architecture for Multi-Scale Derivative Computation," *2-nd Annual IEEE International Conference on Image Processing*, Washington, DC, October 1995.
12. A. Karasiris, and E. Simoncelli, "A Filter Design Technique for Steerable Pyramid Image Transforms," *Proceedings ICASSP-96*, May 7-10, Atlanta, GA.
13. S.Mallat, "A Theory for Multiresolution Signal Decomposition," *IEEE Pattern Analysis and Machine Intelligence*, vol. PAMI-11, no. 7, 1989.
14. J. Le Moigne, W. Xia, J.C. Tilton, T. El-Ghazawi, M. Mareboyana, N. Netanyahu, W.J. Campbell, and R.F. Crompt, "First Evaluation of Automatic Image Registration Methods," *IGARSS'98*, July 6-10, 1998.
15. T. El-Gazhawi, P. Chalermwat, and J. Le Moigne, "Wavelet-Based Image Registration on Parallel Computers," *SuperComputing'97*, San Jose, November 1997.
16. G.W. Rosborough, D.G. Baldwin, W.J. Emery, "Precise AVHRR Image Navigation," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, No. 3, May 1994.
17. D. Baldwin, W. Emery, "Spacecraft Attitude Variations of NOAA-11 Inferred from AVHRR Imagery," *International Journal of Remote Sensing*, Vol. 16, No. 3, 531-548, 1995.

SYNTHETIC	Threshold 15		Threshold 20		Threshold 25	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
Orthogonal/Daubechies						
Accuracy	0.0000	1.0990	0.0000	1.0990	0.0000	1.0990
Standard Deviation	0.0000	0.7513	0.0000	0.7513	0.0000	0.7513
Steerable/Simoncelli						
Accuracy	0.0000	1.1089	0.0000	1.1287	0.0000	1.1287
Standard Deviation	0.0000	0.9109	0.0000	0.9299	0.0000	0.9299
SYNTHETIC-NOISE 2	Threshold 15		Threshold 20		Threshold 25	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
Orthogonal/Daubechies						
Accuracy	0.0000	1.5842	0.0000	1.4752	0.0000	1.3366
Standard Deviation	0.0000	1.2127	0.0000	1.1739	0.0000	0.9978
Steerable/Simoncelli						
Accuracy	0.0000	1.2277	0.0000	1.1089	0.0000	1.1287
Standard Deviation	0.0000	1.2733	0.0000	0.9109	0.0000	0.9299
SYNTHETIC-NOISE 5	Threshold 15		Threshold 20		Threshold 25	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
Orthogonal/Daubechies						
Accuracy	0.0000	1.6535	0.0000	1.4752	0.0000	1.3564
Standard Deviation	0.0000	1.4032	0.0000	1.2396	0.0000	1.0301
Steerable/Simoncelli						
Accuracy	0.0000	1.3663	0.0000	1.3267	0.0000	1.1683
Standard Deviation	0.0000	1.5397	0.0000	1.3867	0.0000	1.0056
SYNTHETIC-NOISE 10	Threshold 15		Threshold 20		Threshold 25	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
Orthogonal/Daubechies						
Accuracy	0.1188	3.5248	0.0792	2.9802	0.0792	2.1188
Standard Deviation	0.4288	2.4600	0.3045	2.4655	0.3045	1.6488
Steerable/Simoncelli						
Accuracy	0.0000	1.4257	0.0000	1.3663	0.0000	1.3267
Standard Deviation	0.0000	1.5113	0.0000	1.4331	0.0000	1.3867
SYNTHETIC-NOISE 20	Threshold 15		Threshold 20		Threshold 25	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
Orthogonal/Daubechies						
Accuracy	3.2970	13.3663	3.0693	12.5446	2.9208	12.0297
Standard Deviation	2.9610	4.5917	2.0646	5.0942	2.0904	6.0419
Steerable/Simoncelli						
Accuracy	0.0000	2.1386	0.0000	1.6436	0.0000	1.5842
Standard Deviation	0.0000	2.1937	0.0000	1.7439	0.0000	1.6245

Table 1
Registration Results for Both Types of Filters on Synthetic Datasets #1-#5

GIRL	Threshold 15		Threshold 20		Threshold 25	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
Orthogonal/Daubechies						
Accuracy	0.5556	3.8889	0.2963	2.3333	0.1481	1.7037
Standard Deviation	1.6178	2.8458	0.5316	1.8459	0.3552	1.3282
Steerable/Simoncelli						
Accuracy	0.0000	0.4815	0.1852	1.0741	0.1852	0.9259
Standard Deviation	0.0000	0.4997	0.4743	1.7832	0.4743	1.3032
TM	Threshold 15		Threshold 20		Threshold 25	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
Orthogonal/Daubechies						
Accuracy	1.9091	7.0606	0.2727	2.5455	0.2727	2.2121
Standard Deviation	2.8000	7.2360	0.9301	3.7424	0.9621	3.1019
Steerable/Simoncelli						
Accuracy	0.0000	0.5455	0.0000	0.5455	0.0000	0.5455
Standard Deviation	0.0000	0.4979	0.0000	0.4979	0.0000	0.4979
AVHRR	Threshold 15		Threshold 20		Threshold 25	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
Orthogonal/Daubechies						
Accuracy	3.2500	12.0000	3.0833	9.9167	3.8333	9.0000
Standard Deviation	0.3727	2.4267	0.2764	5.8878	0.0000	2.4944
Steerable/Simoncelli						
Accuracy	0.1667	2.3333	0.0833	4.0000	0.0000	2.6667
Standard Deviation	1.7381	5.4620	2.5644	6.2244	3.3624	6.7330
ALL SETS	Threshold 15		Threshold 20		Threshold 25	
	Rotation	Translation	Rotation	Translation	Rotation	Translation
Orthogonal/Daubechies						
Accuracy	0.0588	1.5650	0.1173	1.6212	0.1098	1.2661
Standard Deviation	1.7035	5.4647	1.5500	5.0611	1.5770	5.0474
Steerable/Simoncelli						
Accuracy	0.0035	1.3744	0.0104	1.3154	0.0087	1.2392
Standard Deviation	0.0588	1.5650	0.1173	1.6212	0.1098	1.2661

Table 2
Registration Results for Both Types of Filters on Datasets #6-#8 and Overall Results

Figure 10
Summary of Results for Translation Accuracy

