P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Part IV

Applications and Operational Systems

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

14

Multitemporal and multisensor image registration

JACQUELINE LE MOIGNE, ARLENE A. COLE-RHODES,
ROGER D. EASTMAN, NATHAN S. NETANYAHU, HAROLD S. STONE,
ILYA ZAVORIN, AND JEFFREY T. MORISETTE

Abstract

Registration of multiple source imagery is one of the most important issues when dealing with Earth science remote sensing data where information from multiple sensors exhibiting various resolutions must be integrated. Issues ranging from different sensor geometries, different spectral responses, to various illumination conditions, various seasons and various amounts of noise, need to be dealt with when designing a new image registration algorithm. This chapter represents a first attempt at characterizing a framework that addresses these issues, in which possible choices for the three components of any registration algorithm are validated and combined to provide different registration algorithms. A few of these algorithms were tested on three different types of datasets: synthetic, multitemporal and multispectral. This chapter contains the results of these experiments and introduces a prototype registration toolbox.

14.1 Introduction

In Chapter 1, we showed how the analysis of Earth science data for applications, such as the study of global environmental changes, involves the comparison, fusion, and integration of multiple types of remotely sensed data at various temporal, spectral, and spatial resolutions. For such applications, the first required step is fast and automatic image registration which can provide precision correction of satellite imagery, band-to-band calibration, and data reduction for ease of transmission. Furthermore, future decision support systems, intelligent sensors and adaptive constellations will rely on real- or near-real-time interpretation of Earth observation data, performed both onboard and at ground-based stations. The more expert the system and far-reaching the application, the more important will it be to obtain timely and accurately registered data.

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

In Chapter 3, we surveyed many different registration techniques developed for different applications, for example, military, medical, as well as remote sensing, from aircraft and from spacecraft. Despite the wide variety of algorithms available for image registration, no commercial software seems to fully respond to the needs of Earth and space data registration. Many of the methods presented in Chapter 3 are or may be applicable to remote sensing problems but, with such a wide choice, it is necessary to develop a framework to evaluate their performance on well-chosen remote sensing data. Our objective is to carry out systematic studies to support image registration users in selecting appropriate techniques for a remote sensing application based on accuracy and suitability for that application. We carry this out by surveying, designing, and developing different components of the registration process to enable the evaluation of their performance on well-chosen multiple source data, to provide quantitative intercomparison, and to eventually build an operational image registration toolbox. Of course, the main evaluation of automatic image registration algorithms is performed with regards to their accuracy, but it is also useful to relate the accuracy to "initial conditions," that is, the distance between the initial navigation geolocation and the correct result. As was described in Chapter 1, depending on the quality of the navigation model and of the ephemeris data, such initial geolocation may be accurate from within one pixel to tens of pixels. Other ways to evaluate image registration algorithms are in terms of their range of application, geometric and radiometric, and their robustness or reliability. In this chapter, we present the first steps of such an evaluation using representative registration components to build a few image registration algorithms. First, a potential evaluation framework is described. Then, choices for the different components of image registration are reviewed, and the algorithms combining these components are described along with their tests on several test datasets. The first two datasets were synthetic datasets incorporating geometric warping, noise, and radiometric distortion. The algorithms were tested utilizing multitemporal data from the Landsat instrument, and multisensor data acquired over several Earth Observing System (EOS) Land Validation Core Sites. These last datasets include data from the IKONOS, Landsat-7, Moderate Resolution Imaging Spectroradiometer (MODIS), and Sea-viewing Wide Field-of-view Sensor (SeaWiFS) sensors featuring multiple spatial and spectral resolutions.

14.2 A framework for the evaluation of image registration of remote sensing data

The NASA Goddard Image Registration group was started in 1999 with the goal of developing and assessing image registration methodologies that will enable accurate multisource integration. In our work, we assume that the data have already

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

been corrected according to a navigation model and that they are at a level equivalent to the EOS Level 1B (see Chapter 1 for a definition of EOS data levels). Assuming that the results of the systematic correction are accurate within a few or a few tens of pixels, our precision-correction algorithms utilize selected image features or control points to refine this geolocation accuracy within one pixel or a subpixel.

Our studies have been following the first two steps (or components) that define registration algorithms as described in Brown (1992). These are summarized in Chapter 3:

(1) Extraction of features to be used in the matching process.
(2) Feature matching strategy and metrics.
(3) Resampling or indexing of the data.

For alignment, we consider transformations that vary, from translation only in x and y, to rotation, scale and translation (RST).

The long-term goal of our research is to build a modular image registration framework based on these first two components. The concept guiding the development of this framework is that various components of the registration process can be combined in several ways in order to reach optimum registration on a given type of data and under given circumstances. Thereby, the purpose of this framework is twofold:

- It represents a testing framework for:
  - assessment of various combinations of components as a function of the applications,
  - assessment of a new registration component compared to other known ones.
- Eventually, it could be the basis of a registration tool where a user could "schedule" a combination of components as a function of the application at hand, the available computational resources, and the required registration accuracy.

Many choices are available for each of the three components defined above. Our experiments deal with components 1 and 2, first focusing on various types of features utilizing only correlation-based methods and then looking at these features with different similarity metrics and strategies.

14.2.1 Feature extraction

14.2.1.1 Correlation-based experiments

Using cross-correlation as a similarity metric, features such as gray levels, edges, and Daubechies wavelet coefficients were compared using monosensor data (Le Moigne *et al.*, 1998). *Gray level* features were matched using either a basic spatial

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

correlation or a phase correlation. When using *edge features*, the registration was performed in an iterative manner, first estimating independently the parameters of a rigid transformation on the center region of the two images, and then iteratively refining these parameters using larger and larger portions of the images (Le Moigne *et al.*, 1997). *Wavelet features* were also extracted and registered after decomposing both images with a discrete orthonormal basis of wavelets (Daubechies' least asymmetric filters; see Daubechies, 1992 in a multiresolution fashion. Low-pass features, which provide a compressed version of the original data and some texture information, and high-pass features, which provide detailed edge-like information, were both considered as potential registration features (see Le Moigne *et al.*, 2002a, and Chapter 11). This work was focused on correlation-based methods combined with an exhaustive search. One of the main drawbacks of this method is the prohibitive amount of computation required when the number of transformation parameters increases (e.g., affine transformation vs. shift-only), or when the size of the data increases (full-size scenes vs. small portions; multiband processing vs. monoband). To answer some of these concerns, we investigated different types of similarity metrics and different types of feature matching strategies (Subsection 14.2.2).

For this first evaluation, we use three datasets: two synthetic datasets for which the true transformation parameters were known, and one dataset for which no ground truth was available but manual registration was computed. Accuracy and computation times were used as evaluation criteria. Results showed that, as expected, edges or edge-like features like wavelets are more robust to noise, local intensity variations or time-of-the day conditions than original gray level values. On the other hand, when only looking for translation on cloud-free data, phase correlation provides a fast and accurate answer. Comparing edges and wavelets, orthogonal wavelet-based registration is usually faster, although not always as accurate as a full-resolution edge-based registration. This lack of consistent accuracy of orthogonal wavelets is mainly due to the lack of translation invariance, and is presented in more detail in the second set of experiments.

14.2.1.2 Wavelet-based experiments

Chapter 11 describes the set of experiments that we performed using wavelets or wavelet-like features. These experiments verified that separable orthogonal wavelet transforms are not translation- and rotation-invariant. By lack of translation (resp. rotation) invariance, we mean that the wavelet transform does not commute with the translation (resp. rotation) operator. The two studies described in Chapter 11 showed that:

(1) Low-pass subbands of orthogonal wavelets are relatively insensitive to translation, provided that the features of interest have an extent at least twice the size of the

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

wavelet filters, while high-pass subbands are more sensitive to translation, although peak correlations are high enough to be useful (Stone *et al.*, 1999).

(2) Simoncelli's steerable filters perform better than Daubechies' filters. Rotation errors obtained with steerable filters are minimum, independent of rotation size or noise amount. Noise studies also reinforced the results that steerable filters show a better robustness to larger amounts of noise than do orthogonal filters (Zavorin and Le Moigne, 2005).

14.2.2 Feature matching

We then considered various similarity metrics and various matching strategies that can be utilized for feature matching of remote sensing data. As an alternative to correlation, mutual information is another similarity metric that was first introduced in Maes *et al*. (1997) and was used very successfully for medical image registration. Mutual information, or relative entropy, is a basic concept from information theory which measures the statistical dependence between two random variables; or, equivalently, it measures the amount of information that one variable contains about another. Experiments described in Cole-Rhodes *et al.* (2003) show that mutual information may be better suited for subpixel registration as it produces consistently sharper optimum peaks than correlation, thereby yielding higher accuracy.

Mutual information is particularly efficient when used in conjunction with an optimization method, for example, a steepest gradient-based type method like the one described in Irani and Peleg (1991) or a Levenberg-Marquardt optimization like the one utilized for medical image registration and described in Thévenaz *et al*. (1998). Different optimization methods are described in Chapter 12 and in Eastman and Le Moigne (2001). Gray levels, edge magnitudes or low-frequency wavelet information could be used as input to these optimization methods. In Cole-Rhodes *et al.* (2003), mutual information was combined with a stochastic gradient search and the results showed that mutual information is generally found to optimize with one-third the number of iterations required by correlation.

We also studied the use of a statistically robust feature matching method based on the use of nearest-neighbor matching and a generalized Hausdorff distance metric (Mount *et al*., 1999; Netanyahu *et al.*, 2004). This method (also described in Chapter 8) is based on the principle of point mapping with feedback. Specifically, given corresponding sets of control points in the reference and the input images within a prespecified transformation (e.g., rigid, affine), this method derives a computationally efficient algorithm to match these point patterns. The algorithms described use the partial Hausdorff distance and derive the matching transformation either by a geometric branch-and-bound search of transformation space or by

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

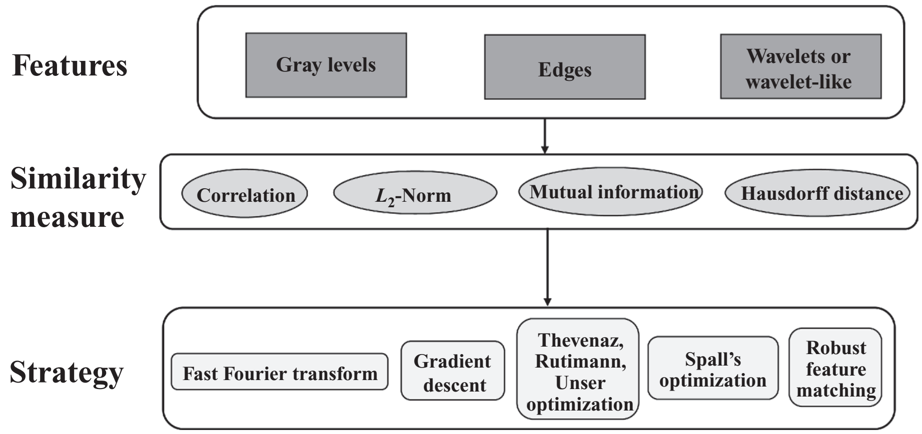

Figure 14.1. Modular approach to image registration combining various choices for feature extraction, similarity metrics, and matching strategy. (Source: Le Moigne *et al*., 2003, © IEEE, reprinted with permission.)

using point alignments. This method has been applied successfully to multitemporal Landsat data using Simoncelli's overcomplete wavelet features; results are described in Subsection 14.3.2.2 (see also Netanyahu *et al*., 2004, for details).

14.2.3 Testing framework

The investigations described in Subsections 14.2.1 and 10.4.2 led to a first version of a testing framework illustrated in Fig. 14.1, in which a registration algorithm is defined as the combination of a set of features, a similarity measure, and a matching strategy. In this framework, *features* can be either gray levels, low-pass features from Simoncelli steerable filters decomposition or from a spline decomposition, or Simoncelli band-pass features; *similarity metrics* can be either cross-correlation, the L_2-norm, mutual information or a Hausdorff distance; *matching strategies* are based either on a fast Fourier correlation, one of the three optimization methods (steepest gradient descent, a Levenberg-Marquardt technique and a stochastic gradient algorithm), or a robust feature matching approach.

By combining these different components, five algorithms were developed and tested. These are compared in Subsection 10.4.3:

- **Method 1**: Gray levels matched by fast Fourier correlation (Stone *et al*., 2001). *We will label it FFC*.
- **Method 2**: Spline or Simoncelli pyramid features matched by optimization and an L_2-norm using the algorithm developed by Thévenaz *et al.* (1998) and Zavorin and Le Moigne (2005). *We will label it TRU*.

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

- **Method 3**: Spline or Simoncelli pyramid features matched by optimization and a mutual information criterion using the algorithm developed by Thévenaz *et al.* (1998) and Zavorin and Le Moigne (2005). *We will label it TRUMI*.
- **Method 4**: Spline or Simoncelli pyramid features matched by optimization of the mutual information criterion using the Spall algorithm (Cole-Rhodes *et al*., 2003). *We will label it SPSA*.
- **Method 5**: Simoncelli wavelet features using a robust feature matching algorithm and a generalized Hausdorff distance (Netanyahu *et al*., 2004). *We will label it RFM*.

For some of the methods (1 and 5), registration is computed on individual subimages and then integrated by computing a global transformation. For the others (2 through 4), registration is computed on the entire images but iteratively, using pyramid decompositions. Another method, called *GGD*, and based on gray-level matching using a gradient descent algorithm with an L_2-norm, was utilized as a reference in early experiments, but since it can be considered as a special case of Method 2, TRU, it will not be systematically evaluated in most of the experiments described in Section 14.3. More details on GGD can be found in Eastman and Le Moigne (2001).

14.3 Comparative studies

In this section, we describe systematic studies that were performed to compare the five algorithms defined in Subsection 14.2.3. Assessing an image registration algorithm for subpixel accuracy and for robustness to noise and to initial conditions, using remote sensing data, presents some difficulty since often ground truth is not available. Interleaving two images for visual assessment can detect gross mismatches and global misalignment but is difficult to extend to quantitative subpixel evaluation. In a few cases, a limited number of control points are known with highly accurate and absolute Ground Positioning System (GPS) information and are used to compute an approximation of the algorithm's accuracy. Manual registration can be used to calculate the unknown transformation but it is uncertain if it is accurate enough to test subpixel accuracy on small regions. Another approach is to generate synthetic image pairs by matching one image against a transformed and resampled version of itself with or without added noise. To avoid some resampling issues this can be done by using high-resolution imagery and downsampling to a lower resolution using an appropriate point spread function to generate both images in a pair. While useful, this approach is limited in realistically modeling noise, temporal scene changes or cross-sensor issues. Yet another approach is to use circular registration results on natural imagery when three or more overlapping images are available. In this case, the transformations should compose to yield the identity – for

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

three images registered pairwise by T_1, T_2, and T_3, the composition T_1 o T_2 o T_3 should be close to the identity transformation (Le Moigne *et al*., 2002b).

For our experiments, we utilize three types of test data: synthetic data, multitemporal data and multisensor data. This section describes the test datasets, followed by the corresponding experiments and the results obtained when testing the five algorithms described in Subsection 14.2.3.

14.3.1 Test data

14.3.1.1 Synthetic datasets

Our goal is to evaluate the strength of various algorithms when applied to many types of satellite data, so test data sets should include imagery from different platforms, with different spatial and spectral resolutions, taken at various dates. The disadvantage of such data is that in the majority of cases, ground truth, if available at all, is approximate at best. Therefore, in our experiments, we first use synthetic images created by a controlled process, designed to emulate real data (Zavorin and Le Moigne, 2005). Three types of transformations were applied in various combinations to a given "source" image to produce synthetic test data, namely, (1) geometric warping, (2) radiometric variations, and (3) addition of noise:

- *Geometric warping* was introduced by simply applying an RST transformation with predetermined amounts of shift, rotation and/or scale to the source image. The resulting warped image is radiometrically identical to the source. The scale was fixed at a value close to 0.95 while "bundling together" the different shifts and rotations. This was done by varying an auxiliary parameter α and assigning its value to t_x, t_y and θ. The use of this parameter decreases the amount of required computations, thus making the experiments faster and easier to interpret, while still keeping the essence of significantly varying shifts and rotations.
- *Radiometric variations* were introduced to mimic how an instrument would actually process a scene. To do this an image representing the "real" scene is convolved with a point spread function (PSF) (Lyon *et al*., 1997). The PSF may or may not correspond to a specific sensor, but it is very important that it does not introduce any geometric warping to the image. In this chapter, we use a simple PSF that was constructed by convolving with itself a 512-by-512 image that was "black" except for the 5-by-5 "white" center. A similar approach for synthetic image generation was used in Stone *et al.* (2001) and Foroosh *et al.* (2002), where Gaussian point spread functions were applied. This general approach can potentially be used to synthesize various multisensor satellite data.
- *Gaussian noise* was added to emulate imperfections of optical systems and of models used in preprocessing of satellite data. The amount of additional noise is usually specified in terms of signal-to-noise ratio (SNR). The SNR of n dB is defined as:

$$n = 10 \cdot \log_{10} \frac{Var\,(\text{image})}{Var\,(\text{noise})}.$$

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

By using these transformations, two synthetic datasets were created:

(1) *Warping & noise* (or "SameRadNoisy"): The first dataset was created by combining geometric warping with noise. The auxiliary parameter α, defining the shifts and rotation, was varied between 0 and 1 (corresponding to shifts between 0 and 1 pixels, and rotations between 0 and 1 degrees), with a step of 0.025, while noise varies between -15 dB and 20 dB with a step of 1. The parameter α was chosen relatively small, compared to the experiments reported in Zavorin and Le Moigne (2005), to ensure convergence in most cases. A total of 1476 ref-input pairs was generated for this dataset.

(2) *Warping & PSF* (or "DiffRadNoiseless"): The second dataset was created by combining geometric warping with radiometric variations using the PSF described above. Again shifts and rotation were varied, using the auxiliary parameter α, before the warped image was convolved with the PSF. A total of 3321 ref-input pairs was generated for this case.

Figures 14.2(a)–(c) show the original image and examples of the synthetic images created from it.

14.3.1.2 Multitemporal datasets

The multitemporal datasets have been acquired over two different areas: (1) the Washington DC/Baltimore area (Landsat WRS-2 Path 15, Row 33) and (2) Central Virginia (Path 15, Row 34). A multitemporal dataset of Landsat-5/Thematic Mapper (TM) and Landsat-7/Enhanced Thematic Mapper (ETM+) images was assembled for each region (see Table 14.1 and Netanyahu *et al.*, 2004). For each region, one ETM+ scene was picked as the "reference" scene; the systematic navigational information provided with the reference scene was considered to be "the truth." A number of reference chips (eight 256×256 pixel subregions for Washington DC/Baltimore and six 256×256 pixel subregions for Central Virginia) were extracted from these reference scenes. For an operational system, it is reasonable to assume that a database would include between 5 and 10 well-distributed "reference chips" per Landsat scene; they are usually defined as small subimages representing well-contrasted visual landmarks, such as bridges, city grids, islands, or high-curvature points in coastlines, and correspond to cloud-free, different seasons and/or different reflectance conditions of each landmark area. With regards to our multitemporal datasets, we have only one reference chip for each landmark area, so only one season and one radiometry are available for reference; therefore these datasets present the following challenges. The Washington DC/Baltimore data involves multiple sensors (ETM+ and TM), and although the band definitions of these sensors are identical, their spectral responses are different; thus those scenes present a challenge due to spectral differences. On the

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

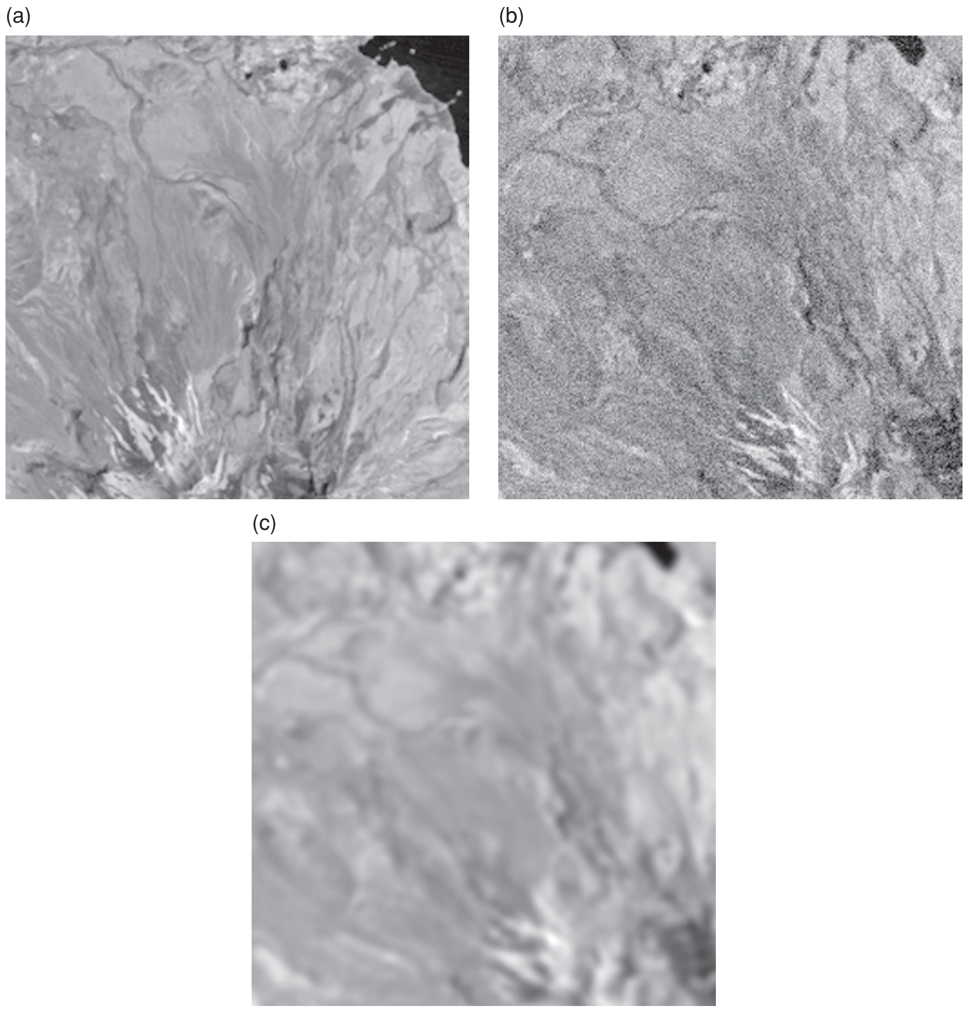

Figure 14.2. Synthetic image samples: (a) Original image; (b) warping & noise; (c) warping & PSF. (Source: Le Moigne *et al*., 2004, © IEEE, reprinted with permission.)

other hand, the Central Virginia dataset spans multiple seasons, and thus presents a challenge in matching features that have very different appearances due to seasonal effects.

All scenes, reference and input, were projected using a WGS-84 model (National Imagery and Mapping Agency, 2000). Using the Universal Transverse Mercator (UTM) coordinates of the four corners of each chip and the UTM coordinates of the four corners of each input scene, a corresponding window was extracted for each chip and each input scene. Figures 14.3 to 14.6 show four of the reference

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.1 *Multitemporal Landsat datasets. (Source: Netanyahu* et al.*, 2004, © IEEE, reprinted with permission)*

Location	Acquisition date	Platform/sensor
Washington, DC (P15 R33)	*July 28, 1999 (990728) (reference)*	*Landsat-7 ETM+*
	August 27, 1984 (840827)	Landsat-5 TM
	May 16, 1987 (870516)	Landsat-5 TM
	August 12, 1990 (900812)	Landsat-5 TM
	July 11, 1996 (960711)	Landsat-5 TM
Central Virginia (P15 R34)	*October 7, 1999 (991007) (reference)*	*Landsat-7 ETM+*
	August 4, 1999 (990804)	Landsat-7 ETM+
	November 8, 1999 (991108)	Landsat-7 ETM+
	February 28, 2000 (000228)	Landsat-7 ETM+
	August 22, 2000 (000822)	Landsat-7 ETM+

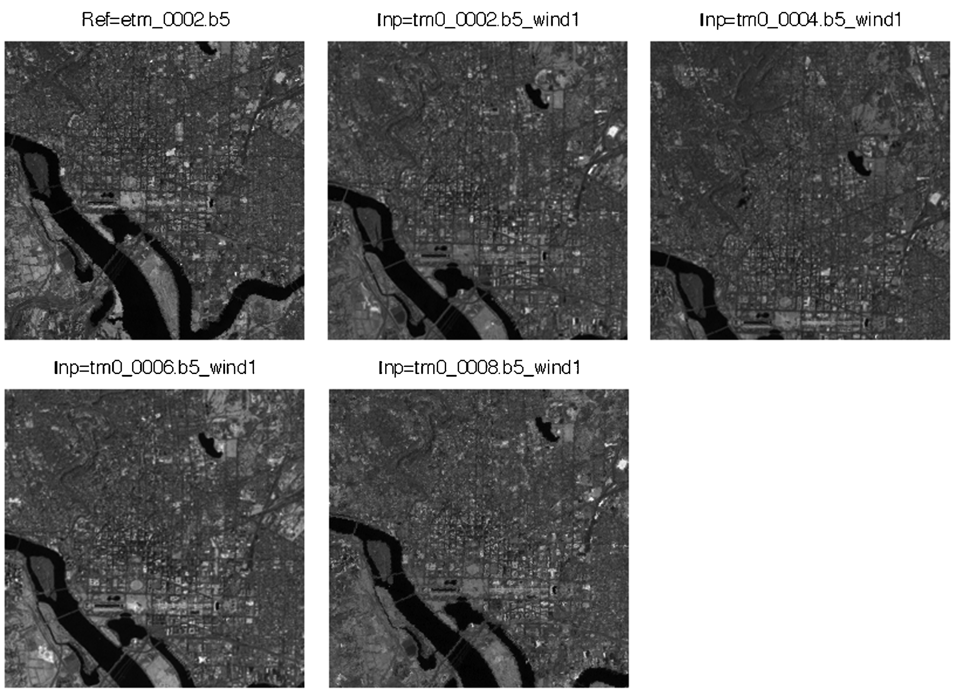

Figure 14.3. Washington DC/Baltimore area: Landsat multitemporal dataset. A reference chip and four input subwindows.

chips and, for each, four corresponding windows from the input scenes, for both the Washington DC/Baltimore and the Central Virginia areas.

For the experiments presented below, each input Landsat-5/ETM or Landsat-7/ETM+ window is registered to its corresponding chip. In our work, we also assume that the transformation between incoming Landsat scenes and reference

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

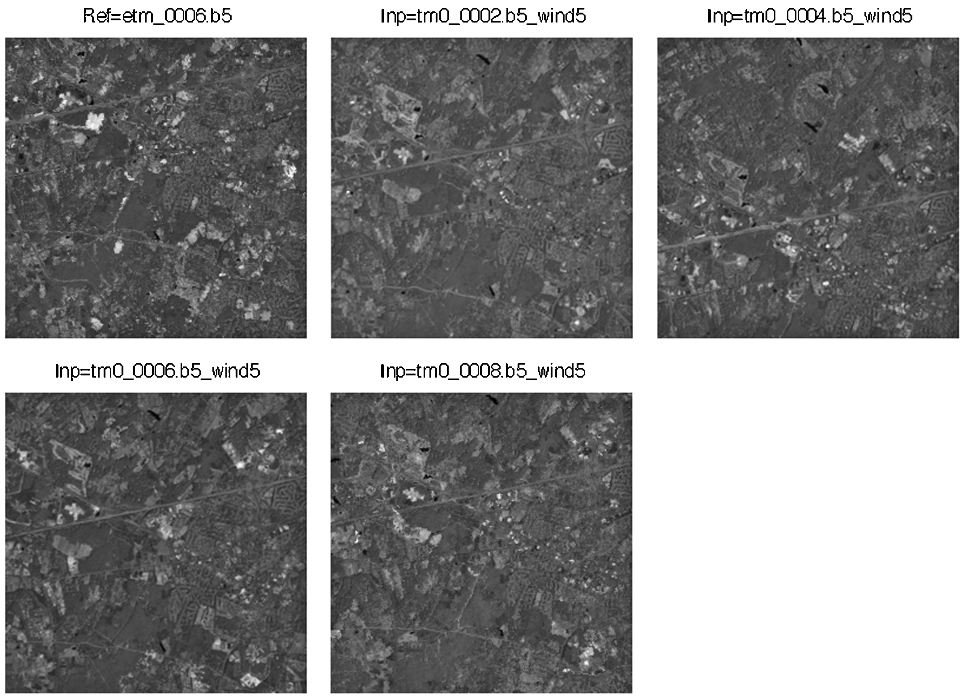

Figure 14.4. Washington DC/Baltimore area: Landsat multitemporal dataset. Another reference chip and four input subwindows.

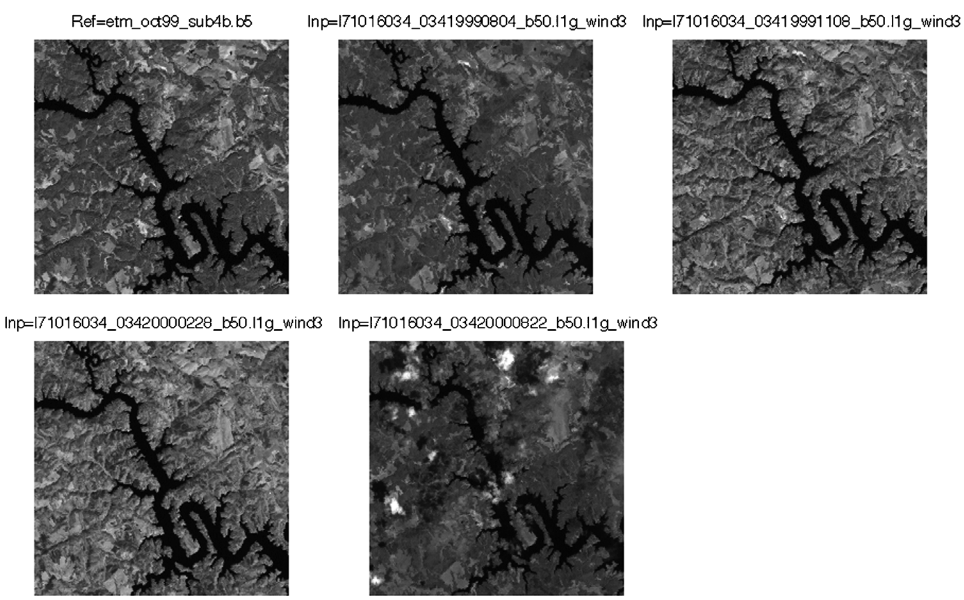

Figure 14.5. Central Virginia area: Landsat multitemporal dataset. A reference chip and four input subwindows.

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.2 *Manual registration of all multitemporal datasets. (Source: Netanyahu* et al.*, 2004, © IEEE, reprinted with permission)*

Reference scenes	Input scenes	Manual ground truth		
		θ	t_x	t_y
DC – 990728	840827	0.026	5.15	46.26
	870516	0.034	8.58	45.99
	900812	0.029	15.86	33.51
	960711	0.031	8.11	103.18
VA – 991007	990804	0.002	0.04	3.86
	991108	0.002	1.20	13.53
	000228	0.008	1.26	2.44
	000822	0.011	0.35	9.78

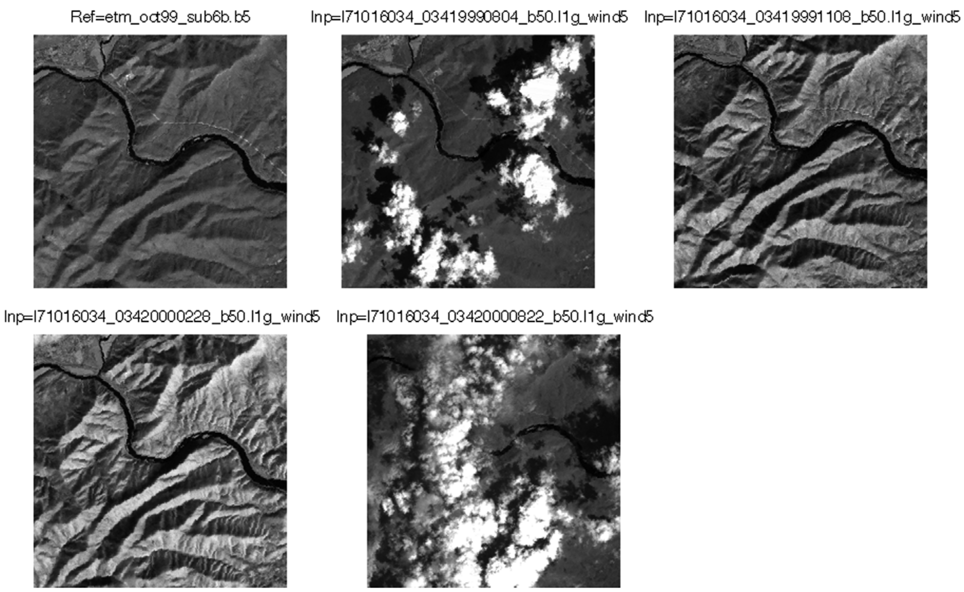

Figure 14.6. Central Virginia area: Landsat multitemporal dataset. Another reference chip and four input subwindows.

chips is limited to the composition of a rotation and a translation. Then, for each pair of scenes, a global registration can be computed with a generalized least-squares method that combines all previous local registrations (see Netanyahu *et al.*, 2004, for more details). Manual registration is available for this dataset to compute algorithm accuracies; see Table 14.2. According to the manual ground truth, the DC datasets present much larger transformations, with rotations of about

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

0.03 radians and shifts varying between 33 and 103 pixels in the vertical direction. On the other hand, the transformations of the VA datasets have rotations ranging from 0.002 radians to 0.011 radians, with the largest translation shift of about 13 pixels.

14.3.1.3 Multisensor datasets

The multisensor datasets used for this study were acquired by four different sensors over four of the MODIS Validation Core Sites (Morisette *et al.*, 2002). The four sensors and their respective bands and spatial resolutions are:

(1) IKONOS bands 3 (red; 632–698 nm) and 4 (near-infrared (NIR); 757–853 nm), at a spatial resolution of 4 meters per pixel,
(2) Landsat-7/ETM+ bands 3 (red; 630–690 nm) and 4 (NIR; 750–900 nm), at a spatial resolution of 30 meters per pixel,
(3) MODIS bands 1 (red; 620–670 nm) and 2 (NIR; 841–876 nm), at a spatial resolution of 500 meters per pixel,
(4) SeaWiFS bands 6 (red; 660–680 nm) and 8 (NIR; 845–885 nm), at a spatial resolution of 1000 meters per pixel.

The four sites represent four different types of terrain in the United States:

(1) *A coastal area* with the Virginia site, data acquired in October 2001;
(2) *An agricultural area* with the Konza Prairie in the state of Kansas, data acquired July to August 2001;
(3) *A mountainous area* with the Cascades site, data acquired in September 2000;
(4) *An urban area* with the USDA, Greenbelt, Maryland site, data acquired in May 2001.

Figures 14.7 to 14.9 show some examples of extracted subimages from the IKONOS, Landsat and MODIS sensors.

14.3.2 Experiments and results

For all experiments, when accurate ground truth is available, a standard way of assessing registration accuracy is by using the root mean square (RMS) error. Details of how to compute RMS are given in Zavorin and Le Moigne (2005), but briefly, if a ground truth transformation is given by $(t_{x_1}, t_{y_1}, \theta_1)$ and a computed transformation is given by $(t_{x_2}, t_{y_2}, \theta_2)$, then the RMS error is given by the following:

$$\begin{aligned} E = {} & \frac{(N_x^2 + N_y^2)}{3} \cdot 2\cos\theta_\varepsilon + \left(t_{x_\varepsilon}^2 + t_{y_\varepsilon}^2\right) \\ & + (N_x t_{x_\varepsilon} + N_y t_{y_\varepsilon})\cos\theta_\varepsilon - (N_x t_{y_\varepsilon} - N_y t_{x_\varepsilon})\cos\theta_\varepsilon, \end{aligned} \tag{14.1}$$

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

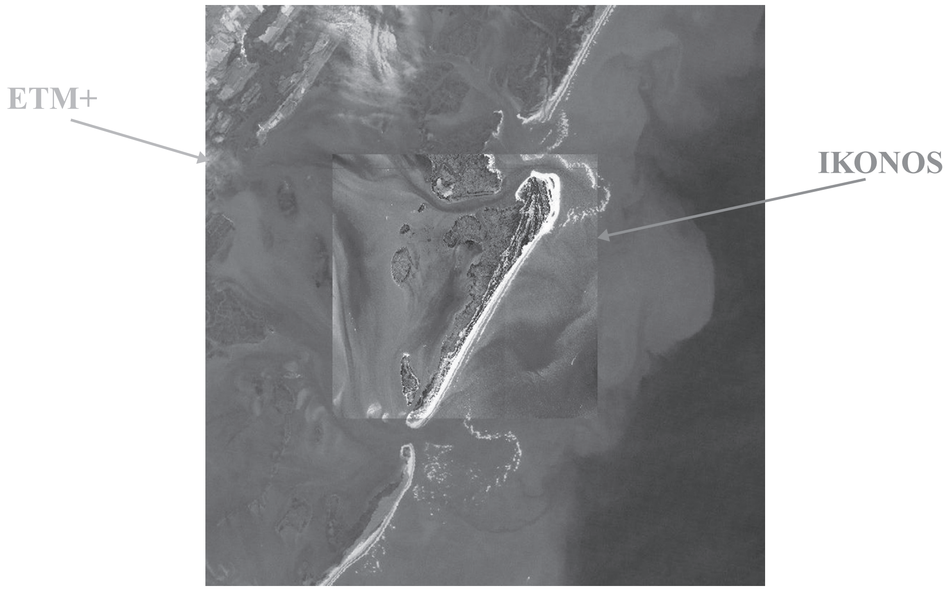

Figure 14.7. ETM+ and IKONOS data of the Virginia coastal area. See Plate 7 in color plates section. (Source: Le Moigne *et al*., 2004, © IEEE, reprinted with permission.)

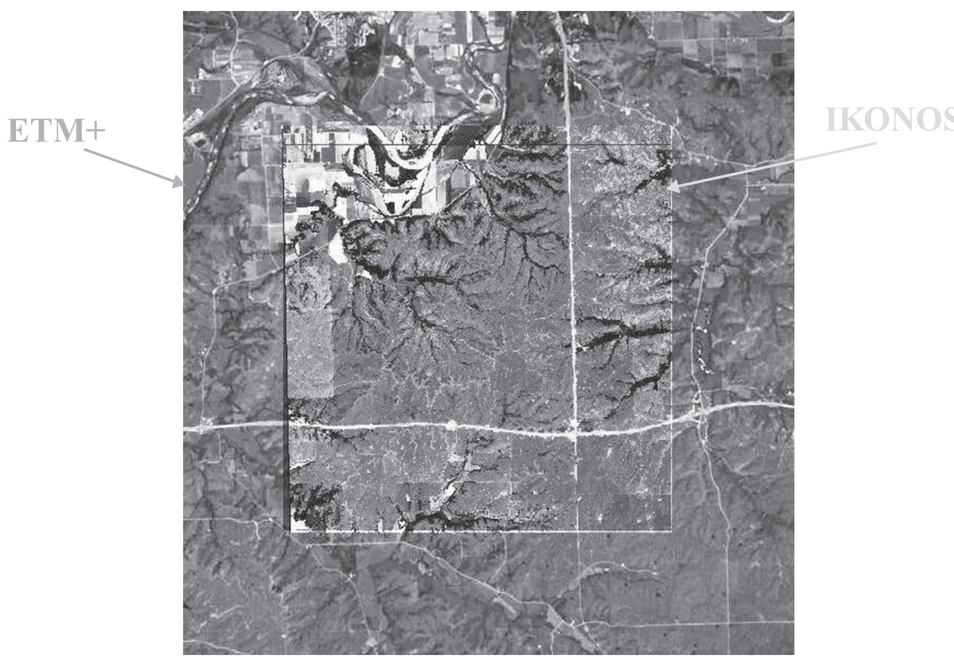

Figure 14.8. ETM+ and IKONOS data of the Cascades mountainous area. See Plate 8 in color plates section. (Source: Le Moigne *et al*., 2004, © IEEE, reprinted with permission.)

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

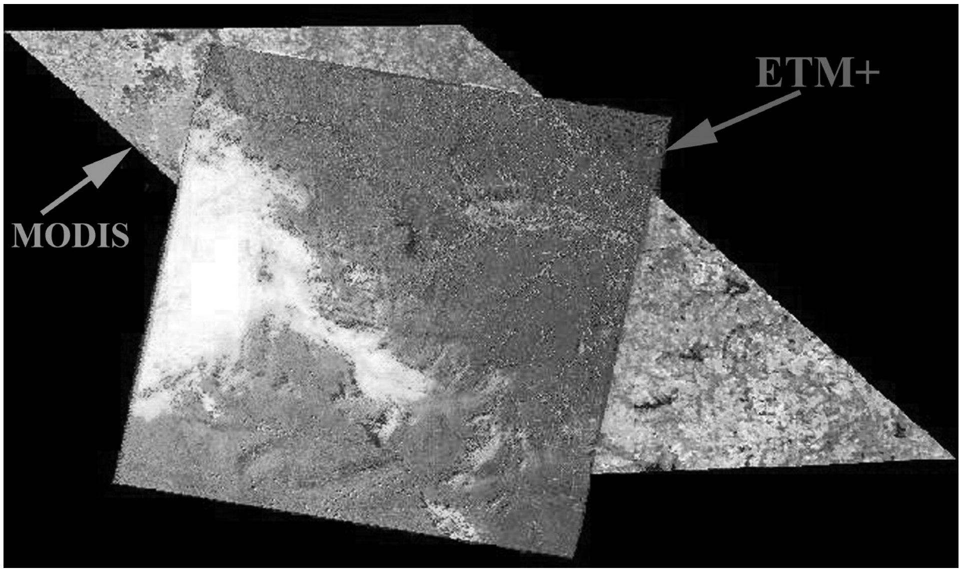

Figure 14.9. ETM+ and MODIS data of the Konza agricultural area. See Plate 9 in color plates section.

where $(t_{x_\varepsilon}, t_{y_\varepsilon}, \theta_\varepsilon)$ represents the "error" transformation between $(t_{x_1}, t_{y_1}, \theta_1)$ and $(t_{x_2}, t_{y_2}, \theta_2)$, N_x represents the size of the images in the x direction, and N_y represents the size of the images in the y direction.

14.3.2.1 Synthetic data experiments

Using the two synthetic datasets described in Subsection 14.3.1.1, we tested four of the algorithms defined in Subsection 14.2.3, namely FFC, TRU, TRUMI and SPSA. For each of the three last algorithms, we utilized three wavelet-like types of features: spline (*SplC*), Simoncelli low-pass (*SimL*) and Simoncelli band-pass (*SimB*). For all three variations of all algorithms, the initial guess was chosen as $(t_x, t_y, \theta) = (0, 0, 0)$, and four pyramid levels were utilized.

Accuracy is measured using Eq. (14.1), so that the RMS error is obtained as a function of either shifts, rotation, and noise (for the first group of datasets), or as a function of shifts, rotation, and radiometric difference (for the second group). Each time an algorithm is run on a pair of images, the resulting RMS error is computed and compared to several thresholds, {0.025, 0.05, 0.075, 0.1, 0.2, 0.25, 0.5, 0.75, 0.75, 1}. Table 14.3 shows the results obtained in this experiment, where for each data type, algorithm, and threshold value, an integer value between 0 and 100 represents the percentage of cases for which the RMS error was below the corresponding threshold value, out of all cases tested. Essentially, the bigger the

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.3 *Summary of experimental results for synthetic data; percentage of cases for which RMS is below threshold*

	TRU			TRUMI			SPSA			FFC
Thresh	SplC	SimB	SimL	SplC	SimB	SimL	SplC	SimB	SimL	Gray
SameRadNoisy										
0.025	*27*	<u>31</u>	**35**	1	0	2	2	3	12	10
0.05	*35*	**47**	<u>41</u>	25	26	29	11	16	28	33
0.075	39	**53**	*43*	33	42	41	18	26	37	<u>44</u>
0.1	42	**59**	44	41	<u>52</u>	45	23	34	41	*51*
0.2	51	**75**	50	61	<u>70</u>	*62*	45	52	54	58
0.25	53	**80**	53	*67*	<u>74</u>	65	50	58	58	60
0.5	63	**91**	63	*81*	<u>85</u>	75	63	72	75	67
0.75	68	**95**	68	*83*	<u>90</u>	80	70	80	83	70
1	74	**97**	72	83	<u>93</u>	83	74	83	*88*	71
DiffRadNoiseless										
0.025	**1**	<u>2</u>	0	**1**	1	1	1	**3**	1	1
0.05	1	**28**	0	5	5	*9*	1	<u>19</u>	6	3
0.075	1	**47**	0	17	14	*19*	5	<u>39</u>	18	9
0.1	1	**63**	0	27	23	29	19	<u>54</u>	*44*	20
0.2	1	<u>86</u>	1	59	51	60	*85*	80	**98**	41
0.25	20	*86*	17	72	62	76	<u>89</u>	86	**100**	43
0.5	83	86	85	86	74	<u>97</u>	90	*92*	**100**	50
0.75	88	86	91	86	74	<u>97</u>	90	*92*	**100**	56
1	90	86	*93*	86	74	<u>97</u>	90	92	**100**	60

Legend: **Best** (bold); <u>Second best</u> (underlined); *Third best* (italic).

number the better is the algorithm. The table also shows, for each threshold, the best algorithm (in bold), the second best (underlined), and the third best one (in italic).

In summary, the results show that for the "SameRadNoisy" dataset, TRU with SimB features performs consistently the best for nearly all thresholds, with an accuracy of 0.25 pixels 80% of the time, with TRUMI using SimB features being second best most of the time. For the "DiffRadNoiseless" dataset, TRU with SimB is best for smaller thresholds, which means that when it converges, it is more accurate, but SPSA with SimL converges more often for higher thresholds and we can say that it reaches accuracies of 0.2 pixels with a 98% probability. Overall, these results show that:

(1) Simoncelli-based methods outperform those with the spline pyramid.
(2) TRUMI (based on the mutual information) does not really perform better than TRU (based on an L_2-norm).

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

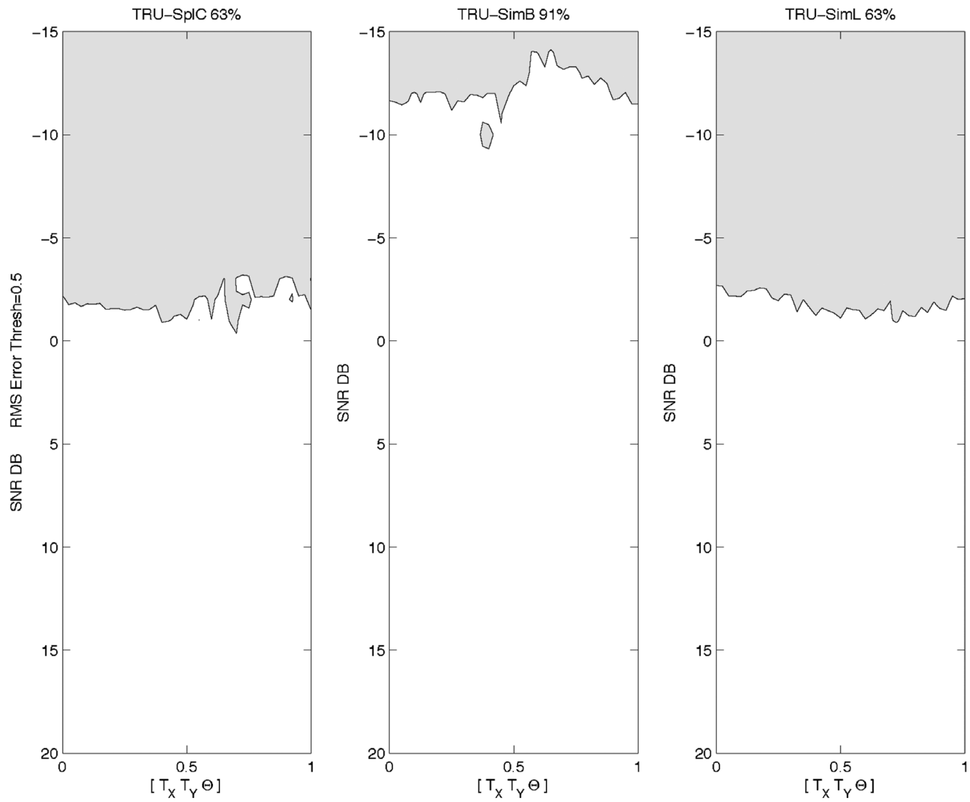

Figure 14.10. Contour plot “SameRadNoisy” dataset; TRU algorithm and a threshold of 0.5.

(3) SimL performs better than SimB for images of different radiometry, but overall all other algorithms seem to perform more poorly for different radiometries than for noisy conditions.

Figures 14.10 to 14.13 show the contour plots corresponding to the results obtained by the four algorithms for the “SameRadNoisy” dataset and a threshold of 0.5, where the white areas depict the regions of convergence of the algorithms with an error less than the threshold. As expected from the results shown in Table 14.3, the plots show that TRU used with SimB (Fig. 14.10) has the largest convergence region, followed by TRUMI with any features (Fig. 14.11). Similarly, Figs. 14.14–14.17 show the contour plots for the “DiffRadNoiseless” dataset and a threshold of 0.2. Again, as expected from Table 14.3, the plots show that for this dataset, SPSA has the largest region of convergence for all types of features, followed by TRU-SimB.

These experiments do not include a study of the sensitivity of the different algorithms to the initial conditions, but previous results reported in Le Moigne

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

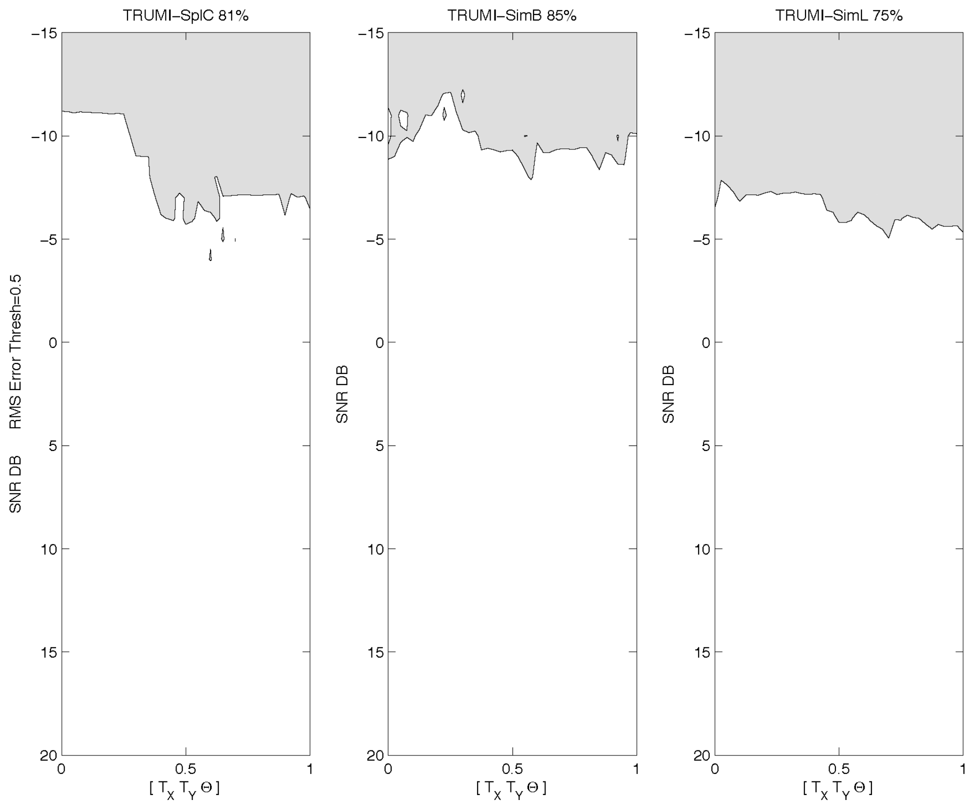

Figure 14.11. Contour plot "SameRadNoisy" dataset; TRUMI algorithm and a threshold of 0.5.

et al. (2004) showed that SimB was more sensitive to the initial guess than SplC or SimL and that SPSA was more robust to initial conditions than TRU or TRUMI.

14.3.2.2 Multitemporal experiments

FFC and optimization-based methods Similarly to the experiments performed on the synthetic datasets, for both DC and VA datasets, we compare the four algorithms, FFC, TRU, TRUMI and SPSA, using the three types of wavelets, SplC, SimB, and SimL for the three latter ones. In this case, not only do we compare the accuracy of the different algorithms but we also assess the sensitivity of the optimization-based methods to the initial conditions, by setting the initial guess of the ground truth to the values given in Table 14.2; that is, if $(t_{x_0}, t_{y_0}, \theta_0)$ is the ground truth between Scenes 1 and 2, the registration is started with the initial guess $(d \cdot t_{x_0}, d \cdot t_{y_0}, d \cdot \theta_0)$ with d taking the successive values {0.0, 0.1, 0.2, 0.3, ..., 0.9, 1.0}.

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

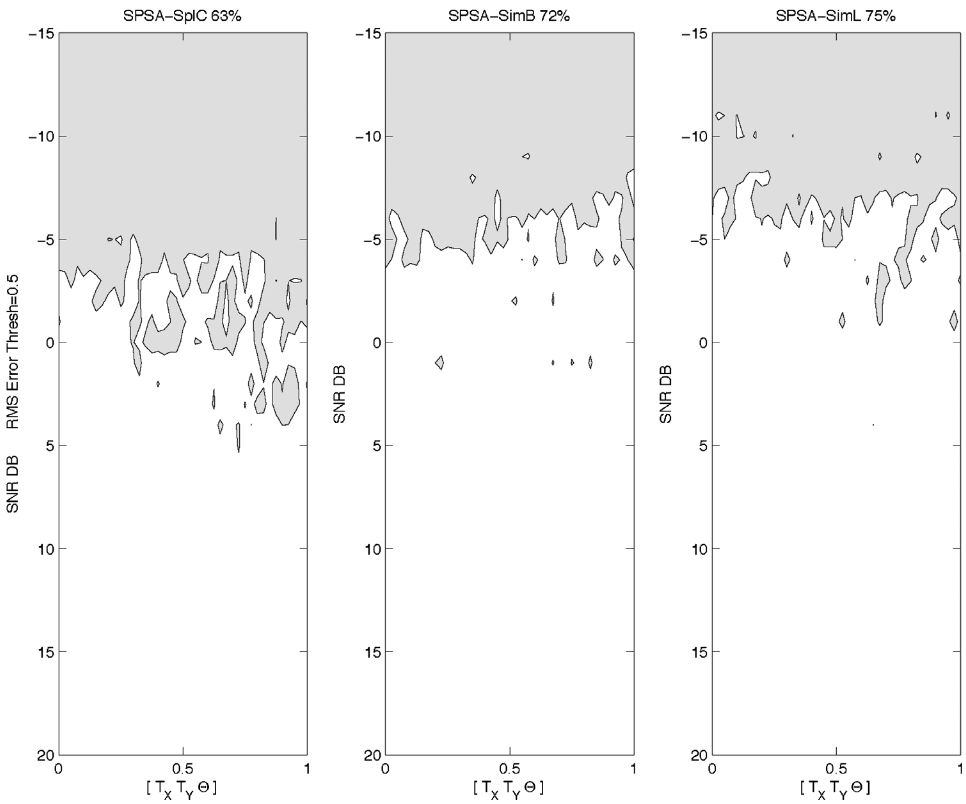

Figure 14.12. Contour plot "SameRadNoisy" dataset; SPSA algorithm and a threshold of 0.5.

Results show that for the VA dataset, all four algorithms perform well. There is very little difference between their accuracy regardless of the initial guess or of the wavelet type used. Tables 14.4a–14.4d show these results for $d = 0.0$ and the four algorithms.

For the DC dataset, unlike the results obtained for VA, we observe that TRU, TRUMI, and SPSA exhibit significant sensitivity to the initial guess, while FFC is essentially insensitive and produces overall the best results. Table 14.5 shows the numbers of correct runs, out of the total of 32 (4 scenes by 8 chip-window pairs) for each algorithm, each pyramid type, and each value of d between 0.0 and 1.0. The results show the following:

- Except for FFC, which is the most insensitive to the initial guess, the algorithms improve performance more or less monotonically as the initial guess gets closer to the ground truth.

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

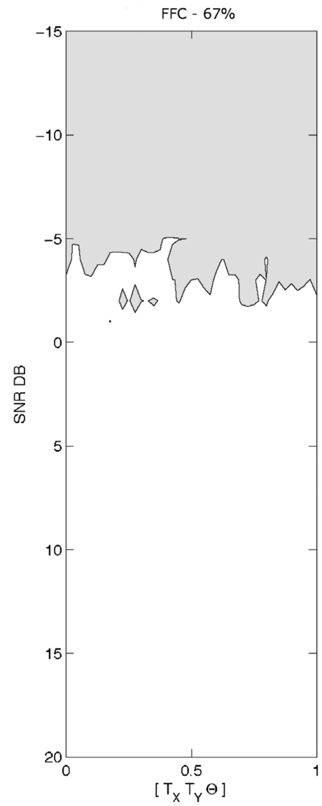

Figure 14.13. Contour plot "SameRadNoisy" dataset; FFC algorithm and a threshold of 0.5.

- TRU, TRUMI, and SPSA are comparable in terms of initial guess sensitivity, with SPSA performing slightly better than the other two.
- Among the three pyramids, SimL seems to perform best in terms of initial guess sensitivity.

These very different results between the DC dataset and the VA dataset might be explained by the characteristics described for the DC area, different sensors and different seasons, as well as by the fact that the DC images tend to have higher frequencies than the VA images. When the images contain high frequencies there is more probability for the optimization algorithms to fall into a local optimum, especially when the initial guess is not very close to the correct solution. Also, FFC is not a local algorithm, it finds the best correlation for each chip, wherever that correlation lies in the image; the algorithm diminishes the effect of false features by correlating a large number of chips, and removing the outliers. The difference

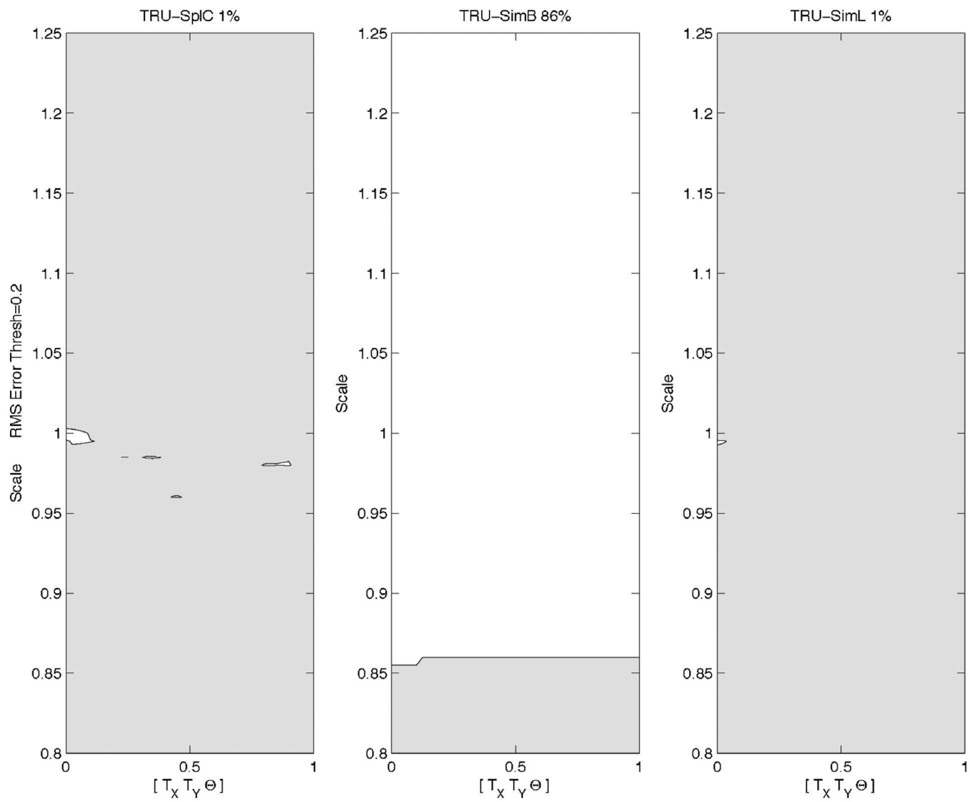

Figure 14.14. Contour plot "DiffRadNoiseless" dataset; TRU algorithm and a threshold of 0.2.

in results between the VA and DC datasets can also be explained by the amount of shifts between reference and input scenes in the DC area being much larger than the ones found in the Virginia area (probably due to the fact that the input DC scenes were acquired by Landsat-5 and the reference by Landsat-7, which has much better navigation capabilities).

Robust feature matching Although the RFM algorithm was not tested simultaneously with the previous algorithms on the multitemporal datasets, results previously obtained in Netanyahu *et al*. (2004), which are summarized in Tables 14.6a and 14.6b, show the following:

- The rotation angle obtained in each case is very small (on the order of a few hundredths of a degree at most). Thus the affine transformation computed can be viewed as essentially "translation only," which is in accordance with Landsat's specifications.

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

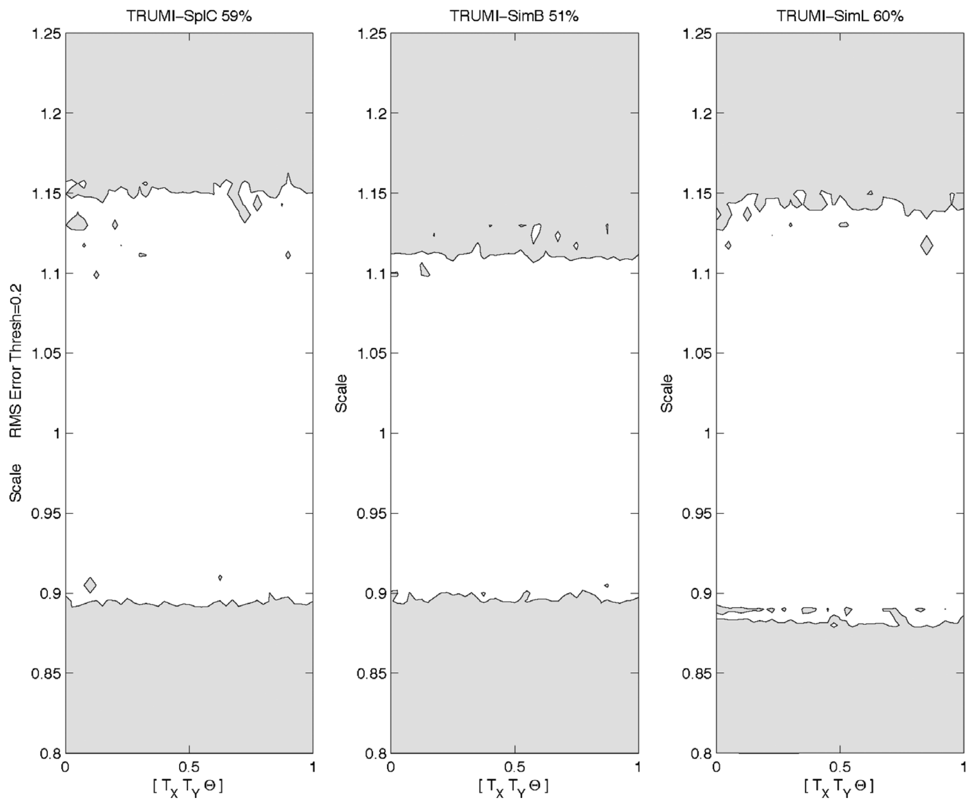

Figure 14.15. Contour plot "DiffRadNoiseless" dataset; TRUMI algorithm and a threshold of 0.2.

- There is a very good agreement between the transformation parameters obtained by RFM and the ground truth shown in Table 14.2. The average errors for the shifts in *x* and *y* were 0.21 and 0.59, respectively, for the DC scenes, and 0.26 and 0.49, respectively, for the Virginia scenes.

RFM was not studied for its sensitivity to initial conditions and this will need to be investigated and compared to the other methods in the future. More details on this method can be found in Netanyahu *et al.* (2004) and in Chapter 8.

14.3.2.3 Multisensor experiments

Algorithm comparison For these experiments, multisensor registrations were performed in "cascade": IKONOS to ETM+, ETM+ to MODIS, and MODIS to SeaWiFS. Wavelet decomposition was utilized, not only to compute registration features, but also to bring various spatial resolution data to similar resolutions, by

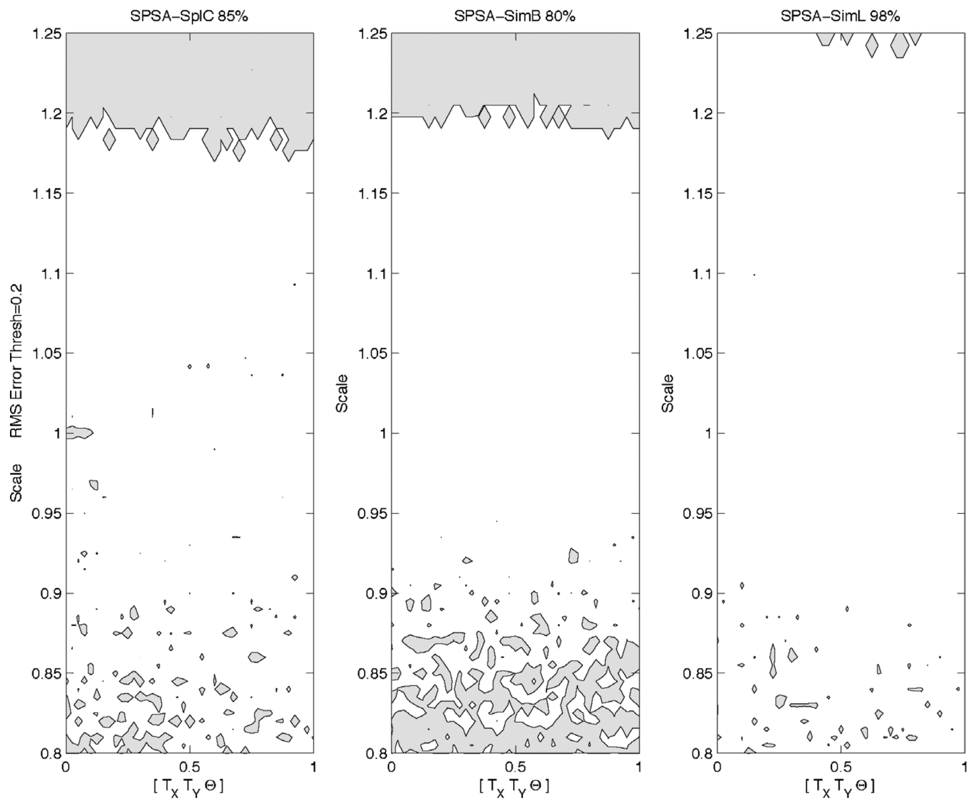

Figure 14.16. Contour plot "DiffRadNoiseless" dataset; SPSA algorithm and a threshold of 0.2.

performing recursive decimation by 2. For example, after three levels of wavelet decomposition, the IKONOS spatial resolution was brought to 32 meters that, compared to the Landsat spatial resolution, corresponds to a scaling of about 1.07. This was the scaling expected when registering IKONOS to Landsat data.

For all scenes, we extracted subimages from the original images so that their dimensions in x and y were multiples of 2^L, where L is the maximum number of wavelet decomposition levels used in the registration process.

We first performed a comparison of the algorithms with the Konza agricultural dataset, using exhaustive search: in order to simplify this comparison, we resampled the IKONOS and ETM+ data to the respective spatial resolutions of 3.91 and 31.25 meters, using the commercial software, PCI®. This slight alteration in the resolution of the data enables us to obtain compatible spatial resolutions by performing recursive decimation by 2 of the wavelet transform, and therefore to only search for translations and rotations. Overall, we considered eight different

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

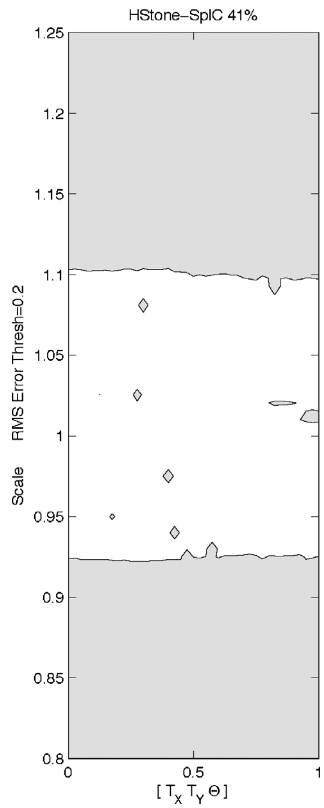

Figure 14.17. Contour plot "DiffRadNoiseless" dataset; FFC algorithm and a threshold of 0.2.

subimages, two for each band of the four sensors. We performed manual registration for two pairs of data (the two bands of IKONOS and ETM data) and found the transformation $t_x = 2$, $t_y = 0$, and $\theta = 0°$. Then, five methods were applied: (a) FFC, (b) GGD, defined in Section 14.2.3, (c) exhaustive search using Simoncelli band-pass and correlation, (d) exhaustive search using Simoncelli band-pass and mutual information, and (e) RFM, defined in Subsection 14.2.3. Results are shown in Table 14.7. All methods confirm the coregistration of the two ETM bands, red and NIR; methods (b), (c), and (d) all found the correct transformation for the registration of IKONOS to ETM, while methods (a) and (e) had 1 or 2 pixel misregistration; all algorithms discovered a misregistration between the MODIS and the SeaWiFS datasets, which was then confirmed manually to be of -8 pixels in the x direction. Overall, all results obtained by the five algorithms were similar within 0.5° in rotation and 1 pixel in translation. More details can be found in Le Moigne *et al.* (2001).

Table 14.4a *Results of multitemporal experiments for the Central Virginia area using TRU and an initial guess of (0, 0, 0)*

TRU algorithm		SplC t_x	SplC t_y	SplC θ	SplC RMS error	SimB t_x	SimB t_y	SimB θ	SimB RMS error	SimL t_x	SimL t_y	SimL θ	SimL RMS error
990804													
Chip1	Wind1	0.312	−3.650	0.026	0.280	0.365	−3.711	0.008	0.328	0.257	−3.565	0.072	0.282
Chip2	Wind2	0.089	−3.897	0.029	0.229	0.110	−3.924	0.007	0.255	0.078	−3.917	0.030	0.247
Chip3	Wind3	2.249	−1.214	−1.248	3.941	0.313	−3.870	−0.024	0.334	2.582	−1.898	−1.437	3.962
Chip4	Wind4	0.035	−3.928	0.021	0.251	0.031	−3.902	0.016	0.224	0.067	−3.958	0.019	0.282
Chip5	Wind5	−0.082	−4.399	−0.431	1.060	0.270	−3.791	−0.014	0.255	−0.354	−6.767	−1.961	4.691
Chip6	Wind6	0.309	−3.875	−0.231	0.519	0.111	−3.919	−0.055	0.266	0.408	−4.192	−0.067	0.639
	MEDIAN	**0.199**	**−3.886**	**−0.105**	**0.400**	**0.190**	**−3.886**	**−0.004**	***0.260***	**0.167**	**−3.937**	**−0.024**	**0.461**
991108													
Chip1	Wind1	1.137	−13.535	0.012	0.064	1.130	−13.526	0.030	0.081	1.097	−13.526	0.012	0.102
Chip2	Wind2	1.416	−13.669	0.007	0.258	1.407	−13.678	0.003	0.255	1.429	−13.696	−0.001	0.282
Chip3	Wind3	1.381	−13.599	0.015	0.198	1.363	−13.600	0.026	0.187	1.373	−13.587	0.017	0.187
Chip4	Wind4	1.322	−13.404	−0.005	0.175	1.298	−13.405	0.001	0.158	1.307	−13.367	−0.012	0.194
Chip5	Wind5	1.270	−13.704	−0.002	0.187	1.254	−13.702	0.000	0.180	1.222	−13.727	−0.007	0.199
Chip6	Wind6	1.229	−13.481	−0.018	0.064	1.208	−13.451	−0.009	0.081	1.234	−13.458	−0.026	0.091
	MEDIAN	**1.296**	**−13.567**	**0.003**	**0.181**	**1.276**	**−13.563**	**0.002**	***0.169***	**1.270**	**−13.557**	**−0.004**	**0.191**

000228													
Chip1	Wind1	−1.177	−2.528	−0.019	0.129	−1.145	−2.517	0.005	0.138	−1.204	−2.544	−0.030	0.134
Chip2	Wind2	−1.136	−2.692	−0.002	0.281	−1.132	−2.682	−0.016	0.276	−1.106	−2.684	−0.007	0.289
Chip3	Wind3	−1.096	−2.535	0.027	0.193	−1.117	−2.546	0.041	0.189	−1.031	−2.480	0.010	0.233
Chip4	Wind4	−1.051	−2.287	0.019	0.260	−1.076	−2.277	0.014	0.246	−1.028	−2.260	0.014	0.294
Chip5	Wind5	−1.112	−2.753	0.019	0.347	−1.120	−2.739	0.021	0.331	−1.105	−2.797	0.019	0.391
Chip6	Wind6	−1.149	−2.365	0.016	0.135	−1.145	−2.337	0.012	0.155	−1.085	−2.503	0.048	0.202
	MEDIAN	**−1.124**	**−2.531**	**0.017**	**0.227**	**−1.126**	**−2.531**	**0.013**	***0.218***	**−1.095**	**−2.524**	**0.012**	**0.261**
000822													
Chip1	Wind1	−3.152	−11.838	1.588	4.749	0.433	−9.565	0.047	0.242	−19.539	−13.491	8.162	23.808
Chip2	Wind2	0.164	−9.774	0.005	0.188	0.166	−9.798	0.009	0.185	0.111	−9.779	−0.007	0.244
Chip3	Wind3	0.356	−9.642	−0.024	0.151	0.385	−9.674	−0.010	0.117	0.309	−9.648	−0.027	0.156
Chip4	Wind4	0.432	−9.329	0.026	0.460	0.432	−9.362	0.024	0.427	0.441	−9.286	0.013	0.503
Chip5	Wind5	0.358	−9.667	0.022	0.115	0.327	−9.705	0.024	0.081	0.334	−9.710	0.023	0.075
Chip6	Wind6	1.512	−21.277	−1.139	11.727	−2.589	−17.576	−2.707	9.809	1.026	−21.184	−2.552	12.310
	MEDIAN	**0.357**	**−9.720**	**0.013**	**0.324**	**0.356**	**−9.689**	**0.016**	***0.213***	**0.322**	**−9.744**	**0.003**	**0.373**

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.4b *Results of multitemporal experiments for the Central Virginia area using TRUMI and an initial guess of (0, 0, 0)*

TRUMI algorithm		SplC				SimB				SimL			
		t_x	t_y	θ	RMS error	t_x	t_y	θ	RMS error	t_x	t_y	θ	RMS error
990804													
Chip1	Wind1	0.407	−3.670	−0.014	0.367	0.306	−3.721	−0.010	0.269	0.471	−3.582	−0.008	0.441
Chip2	Wind2	0.176	−3.956	0.020	0.311	0.121	−3.959	0.016	0.293	0.132	−3.940	0.010	0.276
Chip3	Wind3	0.262	−3.826	0.023	0.271	21.653	−232.311	0.249	229.652	0.222	−3.834	0.016	0.241
Chip4	Wind4	0.063	−3.945	0.017	0.268	0.025	−3.909	0.017	0.232	0.043	−3.981	0.010	0.302
Chip5	Wind5	0.246	−3.886	−0.066	0.310	0.336	−3.820	−0.013	0.327	0.292	−3.691	−0.106	0.307
Chip6	Wind6	0.213	−4.004	0.016	0.369	5.273	−12.356	−5.197	13.609	0.235	−4.255	0.077	0.624
	MEDIAN	**0.230**	**−3.915**	**0.016**	**0.311**	**0.321**	**−3.934**	**0.003**	**0.310**	**0.229**	**−3.887**	**0.010**	***0.304***
991108													
Chip1	Wind1	1.249	−13.508	0.015	0.061	1.151	−13.509	0.017	0.057	1.133	−13.483	0.000	0.083
Chip2	Wind2	1.403	−13.633	0.007	0.228	1.397	−13.694	0.012	0.258	1.419	−13.724	0.014	0.295
Chip3	Wind3	1.393	−13.576	0.008	0.200	1.377	−13.592	0.019	0.194	1.399	−13.615	0.004	0.216
Chip4	Wind4	1.355	−13.373	−0.005	0.219	1.300	−13.394	−0.002	0.168	1.328	−13.358	−0.011	0.214
Chip5	Wind5	1.302	−13.655	0.007	0.162	1.248	−13.706	0.005	0.182	1.234	−13.735	−0.009	0.209
Chip6	Wind6	1.083	1.016	−0.049	14.545	4.528	−1.524	−0.587	12.454	1.191	−13.468	−0.024	0.078
	MEDIAN	**1.328**	**−13.542**	**0.007**	**0.210**	**1.338**	**−13.551**	**0.008**	***0.188***	**1.281**	**−13.549**	**−0.004**	**0.211**

000228													
Chip1	Wind1	−1.155	−2.512	−0.009	0.130	−1.116	−2.510	−0.001	0.160	−1.162	−2.498	−0.030	0.131
Chip2	Wind2	−1.179	−2.749	−0.002	0.319	−1.149	−2.725	−0.006	0.306	−1.165	−2.783	0.005	0.356
Chip3	Wind3	−1.119	−2.660	0.013	0.262	−1.120	−2.598	0.039	0.219	−1.122	−2.665	0.008	0.263
Chip4	Wind4	−1.031	−2.332	0.015	0.254	−1.064	−2.274	0.015	0.257	−1.002	−2.278	0.000	0.305
Chip5	Wind5	−1.130	−2.792	0.014	0.376	−1.131	−2.766	0.024	0.352	−1.188	−2.875	0.009	0.441
Chip6	Wind6	−1.142	−2.417	0.008	0.120	−1.168	−2.312	0.019	0.159	−1.181	−2.399	0.005	0.089
	MEDIAN	**−1.136**	**−2.586**	**0.010**	**0.258**	**−1.125**	**−2.554**	**0.017**	***0.238***	**−1.164**	**−2.581**	**0.005**	**0.284**
000822													
Chip1	Wind1	0.517	−9.525	0.039	0.312	0.436	−9.530	0.042	0.272	6.938	−48.263	10.038	43.298
Chip2	Wind2	0.208	−9.815	0.005	0.148	0.169	−9.801	0.002	0.184	0.207	−9.830	0.004	0.153
Chip3	Wind3	0.409	−9.692	−0.006	0.108	0.408	−9.695	−0.006	0.105	0.387	−9.684	−0.009	0.107
Chip4	Wind4	0.467	−9.430	0.028	0.371	0.431	−9.407	0.024	0.383	0.469	−9.505	0.003	0.299
Chip5	Wind5	0.350	−9.684	0.011	0.096	0.322	−9.712	0.023	0.076	0.330	−9.734	0.012	0.050
Chip6	Wind6	−0.153	1.364	0.367	11.173	−48.823	−1.877	0.142	49.783	−10.486	7.762	−3.828	22.020
	MEDIAN	**0.379**	**−9.604**	**0.020**	**0.230**	**0.365**	**−9.612**	**0.023**	**0.228**	**0.358**	**−9.709**	**0.003**	***0.226***

Table 14.4c *Results of multitemporal experiments for the Central Virginia area using TRUMI and an initial guess of (0, 0, 0)*

SPSA algorithm		SplC				SimB				SimL			
		t_x	t_y	θ	RMS error	t_x	t_y	θ	RMS error	t_x	t_y	θ	RMS error
990804													
Chip1	Wind1	0.390	−3.630	−0.002	0.354	0.284	−3.705	−0.017	0.246	0.440	−3.642	−0.010	0.401
Chip2	Wind2	0.178	−3.863	0.023	0.234	0.117	−3.940	0.015	0.273	0.106	−3.962	0.010	0.291
Chip3	Wind3	0.234	−3.883	0.022	0.285	0.218	−3.880	0.003	0.268	0.235	−3.868	0.019	0.275
Chip4	Wind4	0.032	−3.941	0.000	0.261	0.027	−3.898	0.019	0.221	0.018	−3.980	0.009	0.302
Chip5	Wind5	0.328	−3.805	−0.008	0.313	0.302	−3.781	−0.003	0.281	0.318	−3.819	−0.004	0.311
Chip6	Wind6	0.147	−4.026	0.012	0.364	0.132	−3.906	−0.007	0.245	0.244	−4.040	0.029	0.419
	MEDIAN	**0.206**	**−3.873**	**0.006**	**0.299**	**0.175**	**−3.889**	**0.000**	***0.257***	**0.240**	**−3.915**	**0.010**	**0.306**
991108													
Chip1	Wind1	1.218	−13.519	0.034	0.064	1.148	−13.517	0.018	0.057	1.139	−13.479	0.016	0.081
Chip2	Wind2	1.388	−13.666	0.033	0.244	1.398	−13.687	−0.001	0.252	1.412	−13.722	0.013	0.289
Chip3	Wind3	1.380	−13.619	0.025	0.210	1.361	−13.613	0.019	0.187	1.390	−13.613	0.020	0.214
Chip4	Wind4	1.335	−13.424	−0.008	0.170	1.294	−13.402	−0.009	0.158	1.315	−13.373	−0.017	0.195
Chip5	Wind5	1.245	−13.735	0.004	0.210	1.244	−13.726	0.007	0.201	1.235	−13.765	−0.004	0.238
Chip6	Wind6	1.193	−13.519	−0.009	0.024	1.193	−13.472	−0.004	0.060	1.219	−13.504	−0.010	0.037
	MEDIAN	**1.290**	**−13.569**	**0.015**	**0.190**	**1.269**	**−13.565**	**0.003**	***0.173***	**1.275**	**−13.559**	**0.004**	**0.204**

000228													
Chip1	Wind1	−1.156	−2.503	−0.020	0.130	−1.141	−2.506	0.003	0.136	−1.139	−2.507	−0.023	0.147
Chip2	Wind2	−1.203	−2.756	−0.006	0.322	−1.150	−2.733	−0.012	0.314	−1.141	−2.776	−0.006	0.357
Chip3	Wind3	−1.189	−2.616	0.013	0.190	−1.124	−2.612	0.044	0.230	−1.122	−2.675	0.016	0.273
Chip4	Wind4	−1.071	−2.303	0.012	0.234	−1.056	−2.292	0.020	0.253	−1.028	−2.290	0.005	0.277
Chip5	Wind5	−1.188	−2.800	0.040	0.372	−1.147	−2.773	0.027	0.354	−1.170	−2.855	0.021	0.426
Chip6	Wind6	−1.175	−2.396	0.015	0.097	−1.190	−2.321	0.019	0.140	−1.187	−2.377	0.008	0.096
	MEDIAN	**−1.181**	**−2.560**	**0.013**	***0.212***	**−1.144**	**−2.559**	**0.020**	**0.241**	**−1.140**	**−2.591**	**0.007**	**0.275**
000822													
Chip1	Wind1	0.378	−9.499	0.003	0.283	0.357	−9.456	−0.015	0.328	0.482	−9.447	0.022	0.360
Chip2	Wind2	0.201	−9.842	0.010	0.161	0.185	−9.825	−0.005	0.176	0.185	−9.843	−0.001	0.180
Chip3	Wind3	0.418	−9.645	−0.005	0.153	0.373	−9.705	−0.001	0.080	0.392	−9.673	−0.002	0.117
Chip4	Wind4	0.408	−9.459	0.021	0.327	0.414	−9.369	0.022	0.417	0.428	−9.456	0.015	0.334
Chip5	Wind5	0.347	−9.741	0.016	0.041	0.318	−9.709	0.025	0.080	0.323	−9.729	0.019	0.059
Chip6	Wind6	−4.301	5.894	−1.409	16.603	−18.685	9.732	0.642	27.209	−9.858	11.557	0.500	23.636
	MEDIAN	**0.362**	**−9.572**	**0.007**	***0.222***	**0.337**	**−9.581**	**0.010**	**0.252**	**0.358**	**−9.565**	**0.017**	**0.257**

Table 14.4d *Results of multitemporal experiments for the Central Virginia area using FFC and an initial guess of (0, 0, 0)*

FFC Algorithm		t_x	t_y	θ	RMS error
990804					
Chip1	Wind1	0.218	−3.724	−0.053	0.204
Chip2	Wind2	0.038	−4.001	−0.008	0.322
Chip3	Wind3	0.056	−4.030	0.025	0.358
Chip4	Wind4	−0.001	−3.975	0.007	0.298
Chip5	Wind5	0.145	−4.022	−0.009	0.382
Chip6	Wind6	−0.022	−4.115	0.006	0.455
	MEDIAN	**0.047**	**−4.011**	**−0.001**	***0.340***
991108					
Chip1	Wind1	1.152	−13.537	0.007	0.114
Chip2	Wind2	1.297	−13.782	0.035	0.294
Chip3	Wind3	1.286	−13.675	0.070	0.247
Chip4	Wind4	1.138	−13.267	0.005	0.285
Chip5	Wind5	1.140	−13.913	0.027	0.398
Chip6	Wind6	0.975	−13.296	0.014	0.324
	MEDIAN	**1.146**	**−13.606**	**0.020**	***0.289***
000228					
Chip1	Wind1	−1.234	−2.348	−0.084	0.191
Chip2	Wind2	−1.078	−2.899	−0.052	0.504
Chip3	Wind3	−1.057	−2.600	0.030	0.272
Chip4	Wind4	−1.017	−2.083	0.025	0.434
Chip5	Wind5	−0.986	−2.891	−0.013	0.529
Chip6	Wind6	−1.268	−2.187	−0.001	0.263
	MEDIAN	**−1.068**	**−2.474**	**−0.007**	***0.353***
000822					
Chip1	Wind1	0.524	−9.513	−0.021	0.406
Chip2	Wind2	0.053	−9.677	−0.104	0.398
Chip3	Wind3	0.186	−9.848	−0.059	0.227
Chip4	Wind4	0.614	−9.594	−0.106	0.370
Chip5	Wind5	0.127	−9.887	0.021	0.262
Chip6	Wind6	−126.590	−45.147	76.912	273.851
	MEDIAN	**0.157**	**−9.763**	**−0.040**	***0.384***

We performed registrations of all ETM and IKONOS multisensor data using five of the algorithms defined in Subsection 14.2.3: (a) FFC, (b) TRU with SplC, (c) TRU with SimB, (d) TRU with SimL, and (e) SPSA with SimB (Le Moigne *et al.*, 2003). Overall, for each site, six different registrations are performed, corresponding to inter- and intra-sensor registrations, including cross-spectral (multimodal) matching. Results are shown in Tables 14.8 and 14.9 for two of the sites, Cascades-Mountainous and Virginia-Coast. For this study, no exact ground truth is

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.5 *Number of cases that converged (out of 32) for the DC dataset, running four algorithms with the initial guess varying between the origin ($d = 0.0$) and ground truth ($d = 1.0$)*

	TRU			TRUMI			SPSA			
d	SplC	SimB	SimL	SplC	SimB	SimL	SplC	SimB	SimL	FFC
0.0	7	5	12	10	2	14	5	3	7	30
0.1	8	4	14	8	4	12	5	4	11	30
0.2	8	6	16	8	7	15	7	5	15	30
0.3	8	8	16	11	6	19	11	12	17	30
0.4	10	14	21	10	9	17	16	16	20	30
0.5	15	19	25	15	12	21	17	17	24	30
0.6	16	23	27	15	16	25	22	26	27	30
0.7	22	26	28	20	26	29	24	27	28	30
0.8	24	31	31	27	28	30	31	29	32	30
0.9	30	32	31	29	32	31	32	32	32	30
1.0	31	32	31	32	32	31	32	32	32	30

Table 14.6a *Global transformation versus ground truth parameters for the four scenes in the DC/Baltimore area. The rotation angle is in degrees. (Source: Netanyahu* et al.*, 2004, © IEEE, reprinted with permission)*

	RFM registration			Manual ground truth			Absolute error		
Scene	θ	t_x	t_y	θ	t_x	t_y	$\lvert\Delta\theta\rvert$	$\lvert\Delta t_x\rvert$	$\lvert\Delta t_y\rvert$
840827	0.031	4.72	−46.88	0.026	5.15	−46.26	0.005	0.43	0.62
870516	0.051	8.49	−45.62	0.034	8.58	−45.99	0.017	0.09	0.37
900812	0.019	17.97	−33.36	0.029	15.86	−33.51	0.010	0.11	0.15
960711	0.049	8.34	−101.97	0.031	8.11	−103.18	0.018	0.23	1.21

Table 14.6b *Global transformation versus ground truth parameters for the four scenes in the Virginia area. The rotation angle is in degrees. (Source: Netanyahu* et al.*, 2004, © IEEE, reprinted with permission)*

	RFM registration			Manual ground truth			Absolute error		
Scene	θ	t_x	t_y	θ	t_x	t_y	$\lvert\Delta\theta\rvert$	$\lvert\Delta t_x\rvert$	$\lvert\Delta t_y\rvert$
990804	0.009	0.36	3.13	0.002	0.04	3.86	0.011	0.40	0.73
991108	0.000	1.00	13.00	0.002	1.20	13.53	0.002	0.20	0.53
000228	0.005	0.88	−2.32	0.008	1.26	2.44	0.003	0.38	0.12
000822	0.002	0.41	9.22	0.011	0.35	9.78	0.013	0.06	0.56

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.7 *Results of multisensor registration for the Konza agricultural area using the four different algorithms*

	FFC		GGD		SimB-correl		SimB-MI		RFM	
Pair to register	θ	(t_x, t_y)	θ	(t_x, t_y)	θ	(t_x, t_y)	θ	(t_x, t_y)	θ	(t_x, t_y)
ETM-NIR/ETM-red	Rotation = 0, translation = (0, 0) computed by all methods, using seven subwindow pairs									
IKO-NIR/ETM-NIR	–	(2, 1)	0.0001	(1.99, −0.06)	0	(2, 0)	0	(2, 0)	0.00	(0.0, 0.0)
IKO-red/ETM-red	–	(2, 1)	−0.0015	(1.72, 0.28)	0	(2, 0)	0	(2. 0)	0.00	(0.0, 0.0)
ETM-NIR/MODIS-NIR	–	(−2, −4)	0.0033	(−1.78, −3.92)	0	(−2, −4)	0	(−2, −4)	0.00	(−3.0, 3.5)
ETM-red/MODIS-red	–	(−2, −4)	0.0016	(1.97, −3.90)	0	(−2, −4)	0	(−2, −4)	0.00	(−2.0, −3.5)
MODIS-NIR/SeaWiFS-NIR	–	(−9, 0)	0.0032	(−8.17, 0.27)	0	(−8, 0)	0	(−9, 0)	0.50	(−6.0, 2.0)
MODIS-red/SeaWiFS-red	–	(−9, 0)	0.0104	(−7.61, 0.57)	0	(−8, 0)	0	(−8, 0)	0.25	(−7.0, 1.0)

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.8 *Results of five algorithms on the Cascades-Mountainous area (initial guess = (1.0, 0.0, 0.0, 0.0))*

Cascades	FFC	TRU/SplC	TRU/SimB	TRU/SimL	SPSA/SimB	Median
1. IKO-red/IKO-NIR						
Scale	1.000	1.000	1.000	1.000	1.000	*1.000*
θ	0.000	0.001	0.001	0.001	0.018	*0.001*
t_x	0.014	−0.024	−0.036	−0.046	0.020	*−0.024*
t_y	0.014	−0.160	−0.183	−0.209	0.054	*−0.160*
2. IKO-red/ETM-red						
Scale	1.064	**1.138**	1.064	**1.179**	1.065	*1.065*
θ	0.092	**1.567**	0.074	**2.542**	0.130	*0.130*
t_x	8.674	**10.918**	8.652	**8.993**	8.777	*8.777*
t_y	10.162	**15.750**	10.044	**11.330**	10.039	*10.162*
3. IKO-red/ETM-NIR						
Scale	1.065	**1.064**	1.065	**1.000**	1.064	*1.064*
θ	0.088	**0.091**	0.084	**0.000**	0.114	*0.088*
t_x	8.694	**8.542**	8.641	**0.000**	8.898	*8.641*
t_y	10.217	**10.153**	10.129	**0.000**	10.224	*10.153*
4. IKO-NIR/ETM-red						
Scale	1.064	**1.097**	**1.150**	**no convrg**	1.066	*1.081*
θ	0.039	**−1.153**	**2.108**	**no convrg**	0.128	*0.083*
t_x	8.562	**13.130**	**3.150**	**no convrg**	8.732	*8.647*
t_y	10.164	**12.494**	**9.572**	**no convrg**	9.924	*10.044*
5. IKO-NIR/ETM-NIR						
Scale	1.065	1.065	1.065	1.065	1.065	*1.065*
θ	0.109	0.068	0.070	0.066	0.110	*0.070*
t_x	8.668	8.687	8.704	8.663	8.663	*8.668*
t_y	10.167	10.148	10.140	10.153	10.156	*10.153*
6. ETM-red/ETM-NIR						
Scale	1.000	1.000	1.000	1.000	1.000	*1.000*
θ	−0.001	0.000	0.000	0.000	0.093	*0.000*
t_x	0.079	0.000	0.000	0.000	0.734	*0.000*
t_y	−0.029	0.000	0.000	0.000	0.942	*0.000*
7. IKO-red to IKO-NIR to ETM-NIR to ETM-red						
t_x	8.761	8.663	8.668	8.617	9.417	
t_y	10.151	9.988	9.957	9.944	11.152	
Round-robin error |7 – 2|						
t_x	0.087	2.255	0.016	0.377	0.641	
t_y	0.010	5.762	0.087	1.387	1.112	
8. IKO-red to ETM-red to ETM-NIR						
t_x	8.754	10.918	8.652	8.993	9.511	
t_y	10.133	15.750	10.044	11.330	10.981	

(*cont.*)

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.8 (*cont.*)

Cascades	FFC	TRU/SplC	TRU/SimB	TRU/SimL	SPSA/SimB	Median
Round-robin error |8 – 3|						
t_x	0.059	2.377	0.011	8.993	0.613	
t_y	0.085	5.597	0.085	11.330	0.757	
9. IKO-NIR to ETM-NIR to ETM-red						
t_x	8.747	8.687	8.704	8.663	9.397	
t_y	10.138	10.148	10.140	10.153	11.098	
Round-robin error |8 – 4|						
t_x	0.186	4.443	5.554	no convrg	0.665	
t_y	0.026	2.346	0.567	no convrg	1.174	

available, but we expect the multimodal intra-sensor registrations to be scale $s = 1$, $t_x = 0$, $t_y = 0$, and $\theta = 0°$, with $s = 1.07$ for the IKONOS to Landsat registrations. All results for which these results are not obtained or which seem to fall far from the median result are highlighted in red. The results of Tables 14.8 and 14.9 show that, as expected, the registrations based on gray levels are less reliable on cross-spectral data than those based on edge-like (band-pass) features, but, when reliable, these results are more accurate. Also, most results are within 1/4 to 1/3 pixels of each other (by looking at the median values).

In the absence of ground truth, to assess the accuracy of the registrations in Tables 14.8 and 14.9, one can make use of a technique called *round-robin registration*. The idea is to use three or more images of the same scene, and to form pairwise registrations of those images. For example, the pairwise registrations can be A to C, C to B, and A to B. The registrations of A to C and C to B give one calculation for the relative registration of A with B. That relative registration ideally should be identical to what is obtained when registering A directly to B. In reality, there is always some registration error in the round-robin registrations of A to B, B to C, and C to A. The value of round-robin registrations is that when the error estimate is low, e.g., a fraction of a pixel, there is great confidence that each of the pairwise registrations has low error. Conversely, if the error estimate is high, e.g., several pixels, then at least one and possibly more than one pairwise registration is off by several pixels. However, the analysis gives no indication as to which of the pairwise registrations has high error in the latter case.

Round robin computations performed on the results in Tables 14.8 and 14.9 show that FFC and SPSA/SimB generally result in a smaller round-robin error than the other algorithms. Because round-robin error measures the cumulative error from several pairwise registrations, if only a single pairwise registration has significant error the round-robin error will be significant. For round-robin error to be small, each pairwise registration in the sequence of registrations should be

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.9 *Results of five algorithms on the Virginia-Coast area (initial guess = (1.0, 0.0, 0.0, 0.0))*

VA-COAST	FFC	TRU/SplC	TRU/SimB	TRU/SimL	SPSA/SimB	Median
1. IKO-red/IKO-NIR						
Scale	1.000	1.000	0.999	1.000	1.001	*1.000*
θ	−0.001	0.000	0.002	0.000	0.081	*0.000*
t_x	0.007	−0.148	0.052	−0.243	0.922	*0.007*
t_y	−0.054	−0.484	−0.560	−0.532	0.751	*−0.484*
2. IKO-red/ETM-red						
Scale	1.066	1.064	1.066	1.066	1.066	*1.066*
θ	0.001	0.030	0.019	0.045	0.104	*0.030*
t_x	12.858	13.357	12.944	13.100	13.024	*13.024*
t_y	13.172	12.957	13.200	13.222	14.138	*13.200*
3. IKO-red/ETM-NIR						
Scale	1.619	**1.048**	1.075	1.049	1.066	*1.066*
θ	−0.121	**−1.096**	−1.546	−1.041	0.010	*−1.041*
t_x	12.395	**11.099**	8.465	11.099	12.216	*11.099*
t_y	12.218	**9.276**	12.714	9.529	13.156	*12.218*
4. IKO-NIR/ETM-red						
Scale	1.061	**1.055**	**0.997**	**1.097**	1.067	*1.061*
θ	−0.903	**−1.095**	**−0.665**	**−1.342**	0.972	*−0.903*
t_x	10.329	**27.921**	**−2.465**	**23.063**	16.090	*16.090*
t_y	11.549	**6.665**	**−3.043**	**12.034**	16.097	*11.549*
5. IKO-NIR/ETM-NIR						
Scale	1.065	**1.000**	1.066	1.064	1.066	*1.065*
θ	−0.109	**−0.001**	0.011	0.024	0.006	*0.006*
t_x	12.591	**−5.760**	12.861	13.123	12.856	*12.856*
t_y	12.898	**9.914**	13.169	13.048	13.246	*13.048*
6. ETM-red/ETM-NIR						
Scale	1.000	**0.098**	0.999	0.995	1.000	*0.999*
θ	0.002	**−1.266**	0.003	−0.111	−0.002	*−0.002*
t_x	−0.067	**−1.918**	−0.272	−0.374	0.851	*−0.272*
t_y	−0.014	**−3.849**	0.358	−0.457	0.665	*−0.014*
7. IKO-red to IKO-NIR to ETM-NIR to ETM-red						
t_x	12.531	−7.826	12.641	12.506	14.629	
t_y	12.831	5.581	12.967	12.059	14.662	
Round-robin error |7 – 2|						
t_x	0.326	21.183	0.302	0.594	1.604	
t_y	0.342	7.376	0.233	1.164	0.524	
8. IKO-red to ETM-red to ETM-NIR						
t_x	12.791	11.439	12.672	12.727	13.875	
t_y	13.159	9.108	13.558	12.766	14.803	

(*cont.*)

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.9 (*cont.*)

VA-COAST	FFC	TRU/SplC	TRU/SimB	TRU/SimL	SPSA/SimB	Median
Round-robin error |8 – 3|						
t_x	0.395	0.340	4.206	1.628	1.659	
t_y	0.941	0.168	0.844	3.237	1.647	
9. IKO-NIR to ETM-NIR to ETM-red						
t_x	12.524	−7.678	12.589	12.749	13.706	
t_y	12.885	6.065	13.527	12.591	13.912	
Round-robin error |8 – 4|						
t_x	2.1949	35.5989	15.0546	10.3137	2.3837	
t_y	1.3357	0.6000	16.5696	0.5567	2.1856	

small. For the Cascades mountainous region study summarized in Table 14.8, FFC and SPSA/SimB yielded very consistent results, with FFC producing results that were consistent within 0.18 pixels, and SPSA/SimB results being consistent within 1.11 pixels for all but one of the offsets. The results were much less robust for the Virginia coast region summarized in Table 14.9, where FFC produced results that were within 2.2 pixels and SPSA/SimB results were within 2.4 pixels. The other algorithms generally produced results that revealed significant inconsistencies in the round-robin sense. Therefore, one or more of the pairwise registrations produced by the other algorithms was inaccurate, but no pairwise registration in the sequences for FFC and for SPSA/SimB were inaccurate by more than that indicated by the round robin results.

The Virginia coast region and the Cascades mountainous region produced very different results for the round-robin data. This is possibly due to the differences in the registration features available in the image sets. The mountainous region has many edges visible in each of the images in the set, and the edges provide excellent registration characteristics. Edges are less prevalent in the Virginia coast dataset.

Since the TRU algorithm produced results less reliable than FFC and SPSA for all three types of features, further experiments were performed with TRU where the initial guess was given closer to the expected transformations. Tables 14.10 and 14.11 show the results of these experiments, and for both areas, Cascades and Virginia, it can be seen that the results improved significantly for TRU used in combination with SimB, although the round-robin results are still better for Cascades than for Virginia. For the low-pass features, spline (SplC) and SimL features, the results improved much less significantly. These results are in agreement with the conclusions drawn from the synthetic data experiments.

Overall, these experiments also show that using several algorithms in combination might be a solution to obtain accurate and robust multimodal registration,

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.10 *Results of TRU algorithm with three different features on the Cascades-Mountainous area (initial guess = (1.07, 0.0, 8.0, 10.0))*

CASCADES	TRU/SplC	TRU/SimB	TRU/SimL
1. IKO-red/IKO-NIR			
Scale	1.000	1.000	1.000
θ	0.001	0.001	0.001
t_x	−0.024	−0.036	−0.046
t_y	−0.160	−0.183	−0.209
2. IKO-red/ETM-red			
Scale	1.067	1.065	1.070
θ	0.015	0.065	0.074
t_x	8.384	8.626	9.225
t_y	10.225	10.083	10.423
3. IKO-red/ETM-NIR			
Scale	1.065	1.065	1.066
θ	0.054	0.078	0.044
t_x	8.292	8.470	8.207
t_y	10.315	10.133	10.235
4. IKO-NIR/ETM-red			
Scale	**1.070**	**1.065**	**no convrg**
θ	**0.000**	**0.084**	**no convrg**
t_x	**8.000**	**8.641**	**no convrg**
t_y	**10.000**	**10.130**	**no convrg**
5. IKO-NIR/ETM-NIR			
Scale	1.065	1.065	1.065
θ	0.068	0.070	0.066
t_x	8.687	8.704	8.662
t_y	10.148	10.140	10.153
6. ETM-red/ETM-NIR			
Scale	1.000	1.000	1.000
θ	0.000	0.000	0.000
t_x	0.000	0.000	0.000
t_y	0.000	0.000	0.000
7. IKO-red to IKO-NIR to ETM-NIR to ETM-red			
t_x	8.663	8.668	8.616
t_y	9.988	9.957	9.944
Round-robin error |7 – 2|			
t_x	0.279	0.042	0.609
t_y	0.237	0.126	0.479

(*cont.*)

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.10 (*cont.*)

CASCADES	TRU/SplC	TRU/SimB	TRU/SimL
8. IKO-red to ETM-red to ETM-NIR			
t_x	8.384	8.626	9.225
t_y	10.225	10.083	10.423
Round-robin error $\|8 - 3\|$			
t_x	0.092	0.156	1.018
t_y	0.090	0.050	0.188
9. IKO-NIR to ETM-NIR to ETM-red			
t_x	8.687	8.704	8.662
t_y	10.148	10.140	10.153
Round-robin error $\|8 - 4\|$			
t_x	0.687	0.063	no convrg
t_y	0.148	0.010	no convrg

for example, by using as a final result the median values of all transformation parameters.

14.3.2.4 Subpixel accuracy assessment

This section discusses a technique for estimating registration accuracy in the absence of ground truth. The registration experiment uses a set of images from different spectra and different resolutions of the same Earth region. When registering images of different resolutions, the registration algorithm matches the coarse image to the fine image and to the nearest fine image pixel. Because the resolution of the fine image is a multiple of the resolution of the coarse image, the nearest pixel of the fine image corresponds to a fractional pixel (a *phase*) of the coarse image. To assess the accuracy of the registration, the registrations were compiled for a collection of image pairs such that there are two or sequences of pairwise registrations from which one can find the relative registration on an image A with respect to an image B, which permits the use of a round-robin analysis as discussed in the previous section. The results discussed in this section show a few instances where the registration error is on the order of a tenth of a pixel, others where it is on the order of 1 or 2 pixels, and still others where the error is substantially higher.

In the experiment described in Le Moigne *et al.* (2002b), our objective was to register a coarse image to a fine image at the resolution of the fine image, and therefore to assess the subpixel registration capabilities of our algorithms. For this purpose, we utilized a multiphase filtering technique, in which all possible phases of the fine image are registered with respect to the coarse image. Each different phase was filtered and downsampled to the coarse resolution. The phase that gives the best registration metric gives the registration to the resolution of the fine image.

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.11 *Results of TRU algorithm with 3 different features on the Virginia-Coast area (initial guess = (1.07, 0.0, 12.0, 12.0))*

VA-COAST	TRU/SplC	TRU/SimB	TRU/SimL
1. IKO-red/IKO-NIR			
Scale	1.000	0.999	1.000
θ	0.000	0.002	0.000
t_x	−0.148	0.052	−0.243
t_y	−0.484	−0.560	−0.532
2. IKO-red/ETM-red			
Scale	1.064	1.066	1.066
θ	0.049	0.019	0.039
t_x	13.179	12.944	13.126
t_y	13.050	13.200	13.176
3. IKO-red/ETM-NIR			
Scale	1.048	1.075	1.049
θ	−1.097	−1.546	−1.041
t_x	11.097	8.465	11.099
t_y	9.279	12.174	9.259
4. IKO-NIR/ETM-red			
Scale	**1.100**	1.075	**1.117**
θ	**0.232**	1.591	**0.395**
t_x	**20.835**	15.209	**24.201**
t_y	**17.181**	16.597	**21.848**
5. IKO-NIR/ETM-NIR			
Scale	**1.215**	1.066	1.064
θ	**−0.396**	0.011	0.015
t_x	**17.695**	12.861	13.127
t_y	**24.171**	13.169	13.126
6. ETM-red/ETM-NIR			
Scale	**0.098**	0.999	0.995
θ	**−1.266**	0.003	−0.111
t_x	**−1.918**	−0.272	−0.374
t_y	**−3.849**	0.358	−0.457
7. IKO-red to IKO-NIR to ETM-NIR to ETM-red			
t_x	15.629	12.641	12.510
t_y	19.838	12.967	12.137
Round-robin error |7 – 2|			
t_x	2.450	0.303	0.616
t_y	6.788	0.233	1.039
8. IKO-red to ETM-red to ETM-NIR			
t_x	11.261	12.672	12.752
t_y	9.201	13.558	12.719

(*cont.*)

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.11 (*cont.*)

VA-COAST	TRU/SplC	TRU/SimB	TRU/SimL
Round-robin error |8 – 3|			
t_x	0.164	4.207	1.653
t_y	0.078	1.384	3.460
9. IKO-NIR to ETM-NIR to ETM-red			
t_x	15.777	12.589	12.753
t_y	20.322	13.527	12.669
Round-robin error |8 – 4|			
t_x	5.0580	2.6200	11.4475
t_y	3.1410	3.0700	9.1788

We registered this data using two different criteria, normalized correlation and mutual information.

In practice, we utilized two of the images prepared in the first part of the experiment described in Subsection 14.3.2.3, where the IKONOS and ETM+ data have been resampled to the respective spatial resolutions of 3.91 and 31.25 meters. IKONOS red and near-infrared (NIR) bands (of size 2048 × 2048) were shifted in the x and y directions by the amounts $\{0, \ldots, 7\}$, thus creating 64 images for each band, for a total of 128 images. We used the centered spline, SplC, filters (Unser *et al.*, 1993) to downsample with no offset bias. The 128 phase images were downsampled by 8 to a spatial resolution of 31.25 meters and dimensions of 256 × 256. At the coarse resolution, the integer pixel shifts now correspond to subpixel shifts of {0, 1/8, ..., 7/8}. We constructed reference chips of size 128 × 128 from the ETM-red and ETM-NIR images by extraction at position (64, 64) of the initial images. We know from the results in Subsection 14.3.2.3 that the offset between the original downsampled IKONOS image and the ETM reference image is (2, 0), we expected to find the (x, y) offset of the IKONOS image to the ETM image at about (66, 64).

The complete experiment involved the registration of the 128 × 128 extracted red-band (resp. NIR-band) ETM chips to the 64 phased and downsampled 256 × 256 red-band (resp. NIR-band) IKONOS images. For each 128 × 128 ETM chip and for each of the 64 phase IKONOS images, we computed the maximum correlation (resp. mutual information) and the associated location at which this maximum occurs. Then, we found the phase that gave the maximum correlation (resp. mutual information) out of all phases, and we recorded the corresponding shift and the offset computed in this registration. These are shown in Tables 14.12 and 14.13 under "Coarse resolution (X, Y) phase" and under "(Peak X, Peak Y)." Notice that to all peaks, locations have been approximated to the nearest integers.

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.12 *Correlation results. For each image pair, (X, Y) is the subpixel shift that gives the highest correlation, while for each of those offsets, (Peak X, Peak Y) represents the position of the correlation peak*

Pattern	Reference	Coarse resolution (X, Y) phase giving max correlation		Normalized correlation	Peak X giving max correlation	Peak Y giving max correlation	Relative offset X	Relative offset Y
IKO-red	ETM-red	6/8	1/8	0.8521	67	65	66.2500	64.8750
IKO-red	ETM-NIR	0	3.5/8	0.2722	66	65	66.0000	64.5625
IKO-NIR	ETM-red	4/8	7/8	0.2191	67	66	66.5000	65.1250
IKO-NIR	ETM-NIR	6/8	1/8	0.8436	67	65	66.2500	64.8750

Table 14.13 *Mutual information results. For each image pair, (X, Y) is the subpixel shift that gives the highest MI, while for each of those offsets, (Peak X, Peak Y) represents the position of the MI Peak*

Pattern	Reference	Coarse resolution (X, Y) phase giving max mutual information		Mutual information	Peak X giving max mutual information	Peak Y giving max mutual information	Relative offset X	Relative offset Y
IKO-red	ETM-red	0	1/8	1.3017	66	65	66.0000	64.8750
IKO-red	ETM-NIR	6/8	3/8	0.3411	67	65	66.2500	64.6250
IKO-NIR	ETM-red	2/8	1/8	0.3826	66	65	65.7500	64.8750
IKO-NIR	ETM-NIR	7/8	2/8	1.1653	67	65	66.1250	64.7500

Table 14.14 *Self-consistency study of the normalized correlation results*

Image name	Computed x	Computed y	Comes from registered pair
IKONOS-red	0	0	(Starting point)
IKONOS-NIR	−0.2500	−0.2500	IKO-red to ETM-red and ETM-red to IKO-NIR
IKONOS-NIR	−0.2500	−0.3125	IKO-red to ETM-NIR and ETM-NIR to IKO-NIR

To find the relative offset of an IKONOS pattern to an ETM reference (in the last two columns of Tables 14.12 and 14.13), we subtracted the coarse resolution phase offset (X, Y) (in columns 3 and 4) from the corresponding peak offset (Peak X, Peak Y) (in columns 6 and 7). The two tables agree to within 0.25 pixels for all relative offsets except for the X offset from IKONOS-NIR to ETM-red. They disagree by about 0.75 pixels in that case. There are some small inconsistencies in the tables. If the two IKONOS images were registered to the nearest pixel before downsampling, the offsets that produce maximum correlation peaks should be identical when the downsampled patterns are registered to the same image. But when both IKONOS patterns are registered to "ETM-red," the offsets that produce the highest correlations are different. The same phenomenon also occurs in the mutual information-based registrations. One explanation would be that IKONOS-red and IKONOS-NIR are misregistered by 1 or 2 pixels. Another possible explanation is that cross-spectral registration between IKONOS (downsampled by 8) with respect to ETM has an extra offset of 0.25 pixels when compared to within-spectrum registration. It is uncertain where this offset comes from. It is probably an artifact of the cross-spectral data wherein some edges in the image appear to be shifted because of the spectral responses, and these cause registration peaks to shift. More data are required to study this phenomenon. Overall, we can see that the average absolute difference between computed relative offsets and the expected (64, 66) is about 0.5 pixels for both correlation and mutual information metrics.

Another way to look at the data is to analyze the self-consistency of all four measurements. For this analysis, we computed the (x, y) offset of one of the images from the other three in two different ways. If the data are self-consistent, the answers should be the same. To do this, we established an x base point for "IKONOS-red", and let this be $x = 0$. Then, we use the previous relative offsets shown in Tables 14.12 and 14.13 to determine (x, y) offsets for each of the other three images. Tables 14.14 and 14.15 show these results.

Note that both measures show a displacement of "IKONOS-red" from "IKONOS-NIR" of either 0.25 or 0.125 pixels, and the signs of the relative

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Table 14.15 *Self-consistency study of the mutual information results*

Image name	Computed x	Computed y	Comes from registered pair
IKONOS-red	0	0	(Starting point)
IKONOS-NIR	0.2500	0.0000	IKO-red to ETM-red and ETM-red to IKO-NIR
IKONOS-NIR	0.1250	−0.1250	IKO-red to ETM-NIR and ETM-NIR to IKO-NIR

displacements differ for mutual information and normalized correlation registrations. For these two images, the two measures are self-consistent in their estimates of a relative displacement to within 1/8 of a coarse pixel.

14.4 Conclusions

The studies presented in this chapter investigated the use of various feature extraction and feature matching components for the purpose of remote sensing image data registration. Results were provided on a variety of test datasets, synthetic (including noise and radiometric variations), multitemporal, and multisensor. The performances of six different algorithms utilizing gray levels and wavelet-like features combined with correlation, mutual information, and partial Hausdorff distance as similarity metrics, and Fourier transform, optimization, and robust feature matching as search strategies were evaluated. Two of the metrics, correlation and mutual information, were further studied for subpixel registration.

Using synthetic data, we demonstrated that the algorithm based on a Levenberg-Marquardt optimization using the L_2-norm and band-pass wavelet-like features was the most accurate and the most robust to noise. Nevertheless, using Simoncelli's low-pass features with the same type of algorithm was less sensitive to the initial guess. Overall, an approach based on a stochastic gradient technique with a mutual information metric was more robust to initial conditions. If the transformations are very large and if the images contain many high-frequency features, the approach based on a global fast Fourier correlation of multiple chips seemed to work the best.

More generally, we can estimate two regions of interest in the space of registration parameters based on how the collection of registration algorithms we studied behaves for various parameter sets. We say that an algorithm converges for a set of registration parameters if it converges to a global optimum, that is, to the right answer when two images differ by that set of registration parameters. The first region of interest is the *region of convergence*, within which *all* the algorithms that we studied are likely to converge:

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

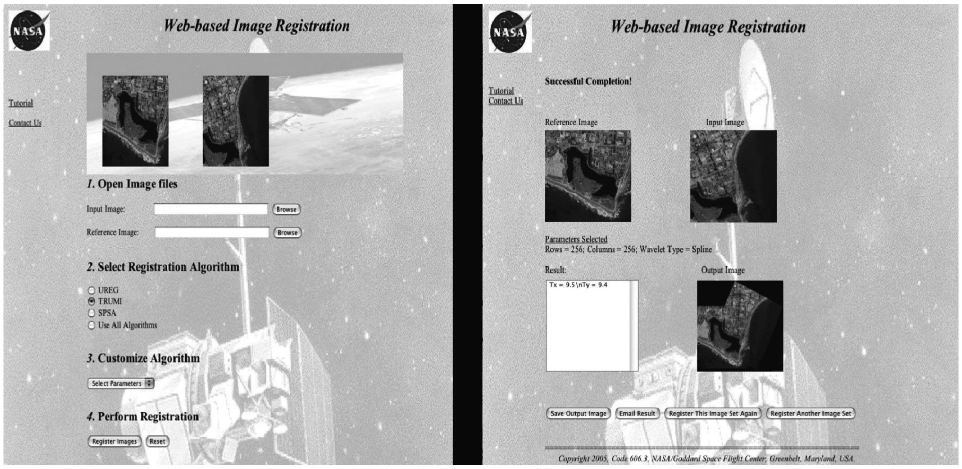

Figure 14.18. User interface of the TARA web-based image registration toolbox. See Plate 9 in color plates section.

- If only shift, it is the region that ranges from −20 to 20 pixels.
- If only rotations, between −10° and 10°.
- If only scale, between 0.9 and 1.1.
- If rotation and shift, then it is when the shift is between −15 and 15 pixels and the rotation between −5° and 5°.
- If rotation, shift and scale, the region of convergence is defined by a shift between −10 and 10 pixels, a rotation between −5° and 5°, and a scale between 0.9 and 1.1.

The second region is the *region of divergence* within which *all* the algorithms will most likely diverge:

- If the shift is more than 30 pixels.
- If the rotation is more than 15°.
- If the scale is less than 0.8 or more than 1.2.
- If together, the rotation and shift are more than 20° and 10 pixels, respectively.
- If rotation, shift and scale, when the shift is larger than 15 pixels, the rotation more than 10°, and the scale less than 0.85 or more than 1.15.

The region between the two is the one where some algorithms converge and some do not. These regions were estimated fairly conservatively, based on a limited number of sample images. To get a more precise estimate, we would need to run more thorough testing with more images of various types.

Based on these first studies, we developed the first prototype of a web-based image registration toolbox (called TARA for "Toolbox for Automated Registration and Analysis") that is depicted in Figure 14.18. At present, this first prototype includes TRU, TRUMI and SPSA with the choice of spline, Simoncelli band-pass

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

or Simoncelli low-pass features. The toolbox's interface is implemented in Java, and the algorithms are implemented in C or C++, but integrated in the toolbox as Java Native Interface (JNI) wrapped functions. The synthetic experiments enable us to define for each method an *applicability range* that will be provided as guidance to the users of the toolbox. At the same time, these methods continue to be tested on other sensor data, for example, the ALI multispectral sensor and the Hyperion hyperspectral sensor, both carried on the EO-1 platform. Eventually, we hope that TARA will be used to assess other registration components and will be extended to include other methods, to compute more general transformations and to process other types of imagery, such as aerial images or other planetary data, for example, from the Moon or Mars.

Acknowledgements

The authors acknowledge the support of NASA under NRA-NAS2–37143, "Research in Intelligent Systems", and of the Goddard Internal Research and Development (IRAD) program, under which many of the experiments have been conducted. The authors also wish to thank Jeffrey Masek for the multitemporal data and the MODIS Validation Core Sites program for the multisensor data.

References

Brown, L. G. (1992). A survey of image registration techniques. *ACM Computing Surveys*, **24**(4), 325–376.

Cole-Rhodes, A. A., Johnson, K. L., Le Moigne, J., and Zavorin, I. (2003). Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *IEEE Transactions on Image Processing*, **12**(2), 1495–1511.

Daubechies, I. (1992). *Ten Lectures on Wavelets*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

Eastman, R. and Le Moigne, J. (2001). Gradient-descent techniques for multi-temporal and multi-sensor image registration of remotely sensed imagery. In *Proceedings of the 4th International Conference on Image Fusion*, Montreal, Canada.

Foroosh, H., Zerubia, J. B., and Berthod, M. (2002). Extension of phase correlation to subpixel registration. *IEEE Transactions on Image Processing*, **11**(3), 188–200.

Irani, M. and Peleg, S. (1991). Improving resolution by image registration, *Computer Vision, Graphics and Image Processing*, **53**(3), 231–239.

Le Moigne, J., El-Saleous, N., and Vermote, E. (1997). Iterative edge- and wavelet-based registration of AVHRR and GOES satellite imagery. In *Proceedings of the Image Registration Workshop*, NASA Goddard Space Flight Center, Greenbelt, MD, pp. 137–146.

Le Moigne, J., Xia, W., El-Ghazawi, T., Mareboyana, M., Netanyahu, N., Tilton, J. C., Campbell, W. J., and Cromp, R. F. (1998). First evaluation of automatic image registration methods. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, Seattle, WA, Vol. 1, pp. 315–317.

P1: SFK Trim: 174mm × 247mm Top: 0.553in Gutter: 0.747in
CUUK1136-14 cuuk1136/Le-Moigne ISBN: 978 0 521 51611 2 September 15, 2010 12:43

Le Moigne, J., Cole-Rhodes, A., Eastman, R., Johnson, K., Morisette, J. T., Netanyahu, N. S., Stone, H. S., and Zavorin, I. (2001). Multisensor registration of Earth remotely sensed imagery. In *Proceedings of the SPIE Conference on Image and Signal Processing for Remote Sensing VII*, Toulouse, France, Vol. 4541, pp. 1–10.

Le Moigne, J., Campbell, W. J., and Cromp, R. F. (2002a). An automated parallel image registration technique based on the correlation of wavelet features. *IEEE Transactions on Geoscience and Remote Sensing*, **40**(8), 1849–1864.

Le Moigne, J., Cole-Rhodes, A., Eastman, R., Johnson, K., Morisette, J., Netanyahu, N., Stone, H., and Zavorin, I. (2002b). Multisensor image registration for on-the-ground or on-board science data processing. In *Proceedings of the Science Data Workshop*, Greenbelt, MD.

Le Moigne, J., Cole-Rhodes, A., Eastman, R., Johnson, K., Morisette, J., Netanyahu, N. S., Stone, H., and Zavorin, I. (2003). Earth science imagery registration. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, Toulouse, France, Vol. I, pp. 161–163.

Le Moigne, J., Cole-Rhodes, A., Eastman, R., Johnson, K., Morisette, J., Netanyahu, N. S., Stone, H., Zavorin, I., and Jain, P. (2004). A study of the sensitivity of automatic image registration algorithms to initial conditions. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, Anchorage, Ak, Vol. 2, pp. 1390–1393.

Lyon, R. G., Dorband, J. E., and Hollis, J. M. (1997). Hubble space telescope faint object camera calculated point spread functions. *Applied Optics*, **36**(8), 1752–1765.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P. (1997). Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, **16**(2), 187–198.

Morisette, J. T., Privette, J. L., and Justice, C. O. (2002). A framework for the validation of MODIS land products. *Remote Sensing of Environment*, **83**, 77–96.

Mount, D. M., Netanyahu, N. S., and Le Moigne, J. (1999). Efficient algorithms for robust feature matching. *Special Issue of Pattern Recognition on Image Registration*, **32**(1), 17–38.

National Imagery and Maping Agency (2000). *Department of Defense World Geodetic System 1984: Its Definition and Relationships with Local Geodetic Systems*. Technical Report 8350.2, 3rd edn.

Netanyahu, N. S., Le Moigne, J., and Masek, J. G. (2004). Georegistration of Landsat data via robust matching of multiresolution features. *IEEE Transactions on Geoscience and Remote Sensing*, **42**(7), 1586–1600.

Stone, H. S., Le Moigne, J., and McGuire, M. (1999). The translation sensitivity of wavelet-based registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **21**(10), 1074–1081.

Stone, H. S., Orchard, M. T., Cheng, E.-C., and Martucci, S. A. (2001). A fast direct Fourier-based algorithm for subpixel registration of images. *IEEE Transactions on Geoscience and Remote Sensing*, **39**(10), 2235–2243.

Thévenaz, P., Ruttimann, U. E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. *IEEE Transactions on Image Processing*, **7**(1), 27–41.

Unser, M., Aldroubi, A., and Murry, E. (1993). The L_2 polynomial spline pyramid. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **15**(4), 364–379.

Zavorin, I. and Le Moigne, J. (2005). Use of multiresolution wavelets feature pyramids. *IEEE Transactions on Image Processing*, **14**(6), 770–782.