

Contents lists available at [ScienceDirect](#)

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

AISIR: Automated inter-sensor/inter-band satellite image registration using robust complex wavelet feature representations

Alexander Wong*, David A. Clausi

Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Image registration
Inter-sensor
Inter-band
Complex wavelet feature representations
Remote sensing

ABSTRACT

An automated registration system named AISIR (Automated inter-sensor/inter-band satellite image registration) has been designed and implemented for the purpose of registering satellite images acquired using different sensors and spectral bands. Sensor and environmental noise, contrast non-uniformities, and inter-sensor and inter-band intensity mapping differences are addressed in the AISIR system. First, a novel modified Geman–McClure M-estimation scheme using a robust phase-adaptive complex wavelet feature representation is introduced for robust control point matching. Second, an iterative refinement scheme is introduced in the AISIR system for improved control point pair localization. Finally, the AISIR system introduces a robust mapping function estimation scheme based on the proposed modified Geman–McClure M-estimation scheme. The AISIR system was tested using various multi-spectral optical, LIDAR, and SAR images and was shown to achieve better registration accuracy than state-of-the-art M-SSD and ARRSI registration algorithms for all of the test sets.

© 2009 Published by Elsevier B.V.

1. Introduction

An open problem in remote sensing is image registration, where remote sensing images taken under different conditions (e.g., time, perspective, sensor technology, spectral band) are aligned. Image registration is important in a number of different remote sensing image analysis applications such as building extraction (Rottensteiner and Jansa, 2002), map rectification, canopy modeling (Chen et al., 2005), and change detection. Image registration of remote sensing images is traditionally performed by human experts by manually selecting ground control point pairs between the images and then estimating the mapping function that aligns the images together. This process of manually selecting ground control point pairs is not only very time consuming, particularly given the large quantity of remote sensing images acquired on a regular basis, but also prone to human bias. As such, automated image registration methods are highly desired to both reduce processing time and improve registration accuracy.

Of particular challenge in automated remote sensing image registration is that of inter-band and inter-sensor image registration, where images acquired from different spectral bands or different sensors (e.g., optical, SAR, LIDAR) are aligned with each other. Different spectral bands and sensing technologies provide insight on

different characteristics of the same scene. As such, the same image content within the scene is represented by different intensity values, thus to perform similarity evaluation using pixel intensities directly is very difficult.

The paper is organized as follows. Given the extensive research literature on inter-band and inter-sensor remote sensing image registration, a summary of background literature in this field is presented in Section 2. The main contributions of the proposed image registration system are outlined in Section 3. The underlying details of the proposed image registration system are presented in Section 4. Experimental results are presented in Section 5 and conclusions are drawn in Section 6.

2. Background

A number of different methods have been proposed to tackle the problem of registering inter-band and inter-sensor remote sensing images, and can be generally divided into one of the following three techniques:

1. *Intensity Remapping-based* (Orchard, 2005, 2007): Intensity remapping-based techniques attempt to estimate an intensity mapping function that transforms the intensity values from an image to that of another image in such a way that the same image content in both images share the same intensity values. These methods assume that the relationship between the intensity values of the images being registered can in fact be

* Corresponding author. Tel.: +1 519 888 4567; fax: +1 519 746 3077.

E-mail addresses: alexanderwong@ieee.org (A. Wong), dclausi@uwaterloo.ca (D.A. Clausi).

represented by a fixed mapping function (e.g., affine and polynomial functions), which is often not true and can lead to poor registration results (Orchard, 2008).

2. *Entropy-based* (Chen et al., 2003; Cole-Rhodes et al., 2003): Entropy-based techniques, in particular those based on mutual information, are popular multimodal image registration methods that attempt to determine the alignment between the images being registered by minimizing the joint intensity entropy between the images. There are several drawbacks to entropy-based methods. First, entropy-based methods are highly under-constrained as they make no assumptions about the intensity relationships between the images. As such, the convergence planes associated with entropy-based methods are typically highly non-monotonic and possess many local optima (Roche et al., 2000). Second, spatial consistency is generally not maintained in entropy-based methods, which is important in distinguishing similar looking characteristics within remote sensing images. Third, entropy-based methods are computationally expensive due to the need to compute marginal and joint entropies.
3. *Feature-based* (Xiaolong and Khorram, 1999; Ali and Clausi, 2002; Eugenio et al., 2002; Wong and Clausi, 2007; Fan et al., 2007; Yasein and Agathoklis, 2008): Feature-based techniques utilize features extracted from the images in a common feature space being registered to estimate the mapping function that align the images together. Such features include invariant moments (Xiaolong and Khorram, 1999), shape characteristics (Ali and Clausi, 2002), contours (Eugenio et al., 2002), local frequency information (Wong and Clausi, 2007), and intensity gradients (Fan et al., 2007). Feature-based techniques share some important benefits when compared to the intensity remapping-based methods and entropy-based methods. Since the features extracted from the images exist in a common feature space, more constrained similarity metrics can be used to compare the images. This results in convergence planes with reduced non-monotonicity and thus fewer local optima. Furthermore, more efficient similarity metrics can be used to compare the images since the extracted features can be compared in a direct manner.

Of the aforementioned methods, feature-based image registration methods have been shown to be particularly promising for the purpose of registering inter-band and inter-sensor images and have been the focus of considerable recent research efforts in this field (Wong and Clausi, 2007; Fan et al., 2007; Yasein and Agathoklis, 2008). Unfortunately, current feature-based methods suffer from several drawbacks:

- a. *Noise sensitivity*: Sensor and environmental noise can have a significant effect on the accuracy of the control point detection and matching processes. Despite attempts made to reduce the effect of noise on the control point detection and matching processes (Wong and Clausi, 2007), the effectiveness of current techniques degrade significantly in the presence of high levels of noise.
- b. *Contrast sensitivity*: Remote sensing images are frequently characterized by global and local contrast non-uniformities due to various imaging and environment conditions. This results in the same image content being represented by differing pixel intensities in the images being registered. While techniques have been proposed to handle contrast non-uniformities (Wong and Clausi, 2007), the effectiveness of such techniques drop significantly in situations characterized by high levels of noise.
- c. *Control point location inaccuracies*: Current feature-based methods are based on the assumption that the locations of

matched control points between the images are exactly the same with respect to geographic coordinates. Unfortunately, this is often not true due to the inaccuracies associated with the automated control point detection process. Geographic coordinate differences in the location of the matched control points can lead to poor estimates of the mapping function, thus negatively affecting the registration accuracy of such methods.

3. Contributions

The main contribution of this paper is AISIR (Automated Inter-sensor/Inter-band Satellite Image Registration), an automated registration system designed to address the drawbacks associated with the existing feature-based inter-sensor and inter-band registration methods. The AISIR system can be broken into the following five steps:

- I. A robust phase-adaptive complex wavelet feature representation is computed from the images being registered.
- II. Control point detection and matching is performed on the feature representations in a robust manner using a novel modified Geman–McClure M-estimation scheme (Geman and McClure, 1987).
- III. The locations of control points are adjusted using an iterative refinement scheme.
- IV. Mismatched control points are removed using Maximum Distance Sample Consensus (MDSAC) (Wong and Clausi, 2007).
- V. The mapping function that aligns the images is estimated in a robust manner using the modified Geman–McClure M-estimation scheme.

Details of these five steps are found in Section 4.

The AISIR system extends significantly upon our preliminary work on inter-sensor and inter-band registration (Wong and Clausi, 2008) in three important ways:

- The AISIR system incorporates novel control point matching and mapping function estimation schemes based on a novel modified Geman–McClure M-estimation scheme to further improve robustness to outliers such as noise and poorly matched control point pairs that were not pruned by the MDSAC algorithm.
- The AISIR system incorporates a novel iterative refinement scheme based on the modified Geman–McClure objective function to further improve the localization accuracy of control point pairs.
- The AISIR system takes geometric distortions into account during the control point matching and control point refinement processes.

The registration accuracy improvements of the AISIR system over our preliminary work will be further demonstrated in the experimental results discussed in Section 5.

4. The AISIR image registration system

The underlying details of the AISIR image registration system is described in this section.

4.1. Problem formulation

Given sensed image $f(x)$ and a reference image $g(x)$ acquired under differing spectral bands or sensor technologies, where $x = (x, y)$, the goal of inter-sensor and inter-band image registra-

tion is to determine a mapping function $T(\underline{x})$ that aligns $f(\underline{x})$ to $g(\underline{x})$ such that

$$\hat{T}(\underline{x}) = \arg \min_{T(\underline{x})} [\Psi(f(T(\underline{x})), g(\underline{x}))] \quad (1)$$

where $\Psi(\cdot)$ is the similarity objective function. In the AISIR system, $T(\underline{x})$ may be represented using either a projective transformation model or high-order polynomial transformation model for situations where the images exhibit more complex geometric distortions relative to each other.

4.2. Control point detection using robust complex wavelet feature representations

The first step in the AISIR image registration system is the automatic identification of control points in both the sensed image and the reference image. This process of extracting control points in an automated manner is very challenging when dealing with situations involving inter-sensor and inter-band images for several reasons. Images acquired using differing sensor technologies or under differing spectrum bands represent similar image content within a scene using different intensity values. Hence, to automatically extract corresponding control points from both images is difficult. Furthermore, remote sensing images are often characterized by contrast and illumination non-uniformities, making automatic identification of distinctive control points based solely on intensities difficult.

One approach that has proven to be effective for automatically extracting control points in inter-sensor and inter-band images is that proposed by Wong and Clausi (2007), where the locations of controls points are identified based on complex wavelet phase moment representations of the images. This approach is highly robust to contrast and illumination non-uniformities, as well as invariant to intensity mapping differences between the images. Unfortunately, the robustness of the method to contrast non-uniformities drops significantly in the presence of sensor and environmental noise. To achieve invariance to intensity mapping differences between inter-sensor and inter-band images while maintaining high robustness to contrast non-uniformities and noise, the AISIR system introduces an automated control point detection scheme based on robust complex wavelet feature representations of the images constructed using iterative phase-adaptive bilateral estimation (Wong, 2008).

The control point detection scheme in AISIR is described as follows. A Log-Gabor Complex Wavelet Transform (Field, 1987) is used to decompose $f(\underline{x})$ into an n -scale over-complete decomposition. Each point \underline{x} at each scale s and orientation θ consists of a set of complex wavelet responses $\Upsilon_{s,\theta}$, as defined by

$$\Upsilon_{s,\theta} = A_{s,\theta}(\underline{x}) e^{i\phi_{s,\theta}(\underline{x})}. \quad (2)$$

The amplitude $A_{s,\theta}(\underline{x})$ and phase $\phi_{s,\theta}$ for a complex wavelet response $\Upsilon_{s,\theta}$ is defined as

$$A_{s,\theta}(\underline{x}) = \sqrt{(f(\underline{x}) * R_{s,\theta}^e(\underline{x}))^2 + (f(\underline{x}) * R_{s,\theta}^o(\underline{x}))^2}, \quad (3)$$

and,

$$\phi_{s,\theta}(\underline{x}) = \tan^{-1} \left(\frac{f(\underline{x}) * R_{s,\theta}^e(\underline{x})}{f(\underline{x}) * R_{s,\theta}^o(\underline{x})} \right), \quad (4)$$

where $R_{s,\theta}^e(\underline{x})$ and $R_{s,\theta}^o(\underline{x})$ are the even- and odd-symmetric Log-Gabor quadrature pairs, and $*$ denotes a convolution.

In AISIR, the complex wavelet feature representation is based on the concept of phase coherence, which has been demonstrated to be effective at capturing structural characteristics within an image

(Wong, 2008; Morrone and Burr, 1988; Kovesi, 2003). The computation of local phase coherence relies primarily on the phase information of an image, and as such is highly robust to illumination and contrast non-uniformities, as well as intensity mapping differences between images. An estimate of phase coherence ρ at orientation θ can be computed as the normalized weighted summation of complex phase deviations (Kovesi, 2003),

$$\rho_{\theta}(\underline{x}) = \frac{\sum_n W_{\theta}(\underline{x}) |A_{n,\theta}(\underline{x}) \Delta \Phi_{n,\theta}(\underline{x}) - T|}{\sum_n A_{n,\theta}(\underline{x}) + \varepsilon}, \quad (5)$$

where W is a weighting factor that favors phase coherence across a wider frequency spread, T is a noise threshold, ε is a small constant used to avoid division by zero, and $\Delta \Phi_{n,\theta}(\underline{x})$ is the complex phase deviation,

$$\Delta \Phi_{n,\theta}(\underline{x}) = \cos(\phi_{n,\theta}(\underline{x}) - \bar{\phi}_{\theta}(\underline{x})) - |\sin(\phi_{n,\theta}(\underline{x}) - \bar{\phi}_{\theta}(\underline{x}))|, \quad (6)$$

where $\bar{\phi}_n$ is the mean phase. A high value of ρ signifies strong structural characteristics. The details and the parameters used to estimate the phase coherence are described in (Kovesi, 2003). The minimum and maximum complex wavelet phase coherence moments φ and v are determined based on ρ ,

$$\varphi(\underline{x}) = \frac{1}{2} \sum_{\theta} \rho_{\theta}(\underline{x})^2 - A \quad (7)$$

$$v(\underline{x}) = \frac{1}{2} \sum_{\theta} \rho_{\theta}(\underline{x})^2 + A \quad (8)$$

where,

$$A = \frac{1}{2} \left[4 \left(\sum_{\theta} (\rho_{\theta}(\underline{x}) \sin(\theta)) (\rho_{\theta}(\underline{x}) \cos(\theta)) \right)^2 + \left(\sum_{\theta} [(\rho_{\theta}(\underline{x}) \cos(\theta))^2 - (\rho_{\theta}(\underline{x}) \sin(\theta))^2] \right)^2 \right]^{\frac{1}{2}}. \quad (9)$$

A bilateral phase-adaptive image estimate f' is then computed based on v (Wong, 2008):

$$f'(\underline{x}) = \frac{\sum_{\zeta} w(\underline{x}, \zeta, v(\underline{x})) f(\zeta)}{\sum_{\zeta} w(\underline{x}, \zeta, v(\underline{x}))}, \quad (10)$$

where the estimation weighting function w is defined as the product of a phase-adaptive spatial weighting function w_{α} and an phase-adaptive amplitudinal weighting function w_{β} over a neighborhood ζ around \underline{x} ,

$$w(\underline{x}, \zeta, v(\underline{x})) = w_{\beta}(\underline{x}, \zeta, v(\underline{x})) w_{\alpha}(\underline{x}, \zeta, v(\underline{x})), \quad (11)$$

$$w_{\alpha}(\underline{x}, \zeta, v(\underline{x})) = \exp \left[-\frac{1}{2} \left(\frac{\|\underline{x} - \zeta\|}{\sigma_{\alpha}(v(\underline{x}))} \right)^2 \right], \quad (12)$$

$$w_{\beta}(\underline{x}, \zeta, v(\underline{x})) = \exp \left[-\frac{1}{2} \left(\frac{\|f(\underline{x}) - f(\zeta)\|}{\sigma_{\beta}(v(\underline{x}))} \right)^2 \right]. \quad (13)$$

The estimated image $f'(\underline{x})$ is then used to estimate the local phase coherence ρ during the next iteration. The robust complex wavelet feature representations used by AISIR is composed of the maximum and minimum complex wavelet phase coherence moment estimates computed after m iterations of the estimation process has been completed to achieve convergence,

$$\zeta = \{\varphi_m, v_m\}. \quad (14)$$

The pseudo-code for the iterative complex wavelet representation estimation is presented in Algorithm 1.

Algorithm 1. Iterative Representation Estimation

```

1: for  $i = 1$  to  $m$  do
2:   Compute amplitude  $A$  (3) and phase  $\phi$  (4) from  $f(\underline{x})$ .
3:   Compute phase coherence  $\rho$  from  $A$  and  $\phi$  (5).
4:   Compute complex wavelet moments  $\varphi$  (7) and  $v$  (8).
5:   Compute image estimate  $f'(\underline{x})$  (10) from  $f(\underline{x})$  and  $v$ .
6:   Set  $f(\underline{x}) = f'(\underline{x})$ .
7: end for
8: return  $\zeta = \{\varphi, v\}$ .

```

To identify a set of control points $p_{f,0}, p_{f,1}, \dots, p_{f,r}$ for $f(\underline{x})$, a non-maximum suppression scheme is applied to the minimum complex wavelet phase coherence moments φ_m of each image and the n strongest local maxima of φ_m are chosen as control points. The minimum complex wavelet phase coherence moment v was used to identify control points as high values of v represent strong structural characteristics across multiple directions, and thus such points are more likely to be structural distinct when compared to other points. A quadratic estimation scheme is then performed based on the neighboring pixels of each control point to refine the position of the control points at a sub-pixel level. The control point extraction process is repeated for $g(\underline{x})$ to determine a set of control points $p_{g,0}, p_{g,1}, \dots, p_{g,q}$.

4.3. Modified Geman–McClure control point matching

After control points have been identified, the second step in the AISIR system is to match the control points from the sensed image to that of the reference image. Similar to control point detection, the process of matching control points in an automated manner is very challenging for inter-sensor and inter-band images due to a number of issues. Since images acquired using differing sensor technologies or under differing spectrum bands represent the same image content within a scene using different intensity values, similarity objective functions that utilize intensity values in a direct manner are ill-suited. Furthermore, remote sensing images are often characterized by sensor and environmental noise and contrast non-uniformities, which results in poor intensity similarity between the same image content captured in the inter-sensor and inter-band images.

The AISIR system attempts to address the issues associated with inter-band and inter-sensor intensity mapping differences, contrast non-uniformities, and noise using two steps. In the first step, rather than performing similarity evaluation based on intensity values, which is ill-suited due to intensity mapping differences, the AISIR system evaluates the similarity between control points based on the robust phase-adaptive complex wavelet feature representations constructed during the control point detection process. Since the feature representation is highly robust to contrast non-uniformities and intensity mapping differences, it allows the control points to be evaluated using well-constrained similarity objective functions based on their corresponding complex wavelet features.

In the second step, the AISIR system introduces a control point similarity evaluation process based on a modified Geman–McClure M-estimation scheme (Geman and McClure, 1987). Remote sensing images are often characterized by sensor and environmental noise, as well as other sensor related-artifacts. Popular quadratic (least-squares) and correlation-based schemes perform poorly under these situations since the error distribution is often heavy-tailed due to noise and artifacts. One effective approach to improving robustness to outlier error is to utilize an M-estimation approach (Huber, 1981). Given a neighborhood ζ centered at \underline{x}_i in $f(\underline{x})$, the generalized M-estimation method for finding the location of the matching control point at $\underline{x}_i + \underline{\delta}_i$ in $g(\underline{x})$ can be defined as,

$$\hat{\underline{\delta}}_i = \arg \min_{(\underline{\delta})} \left[\sum_{\zeta(\underline{x}_i)} \Psi(v_m^f(\Theta \underline{x}), v_m^g(\underline{x} + \underline{\delta})) \right], \quad (15)$$

where Ψ is the similarity objective function and v_m^f and v_m^g are the estimated maximum complex wavelet phase coherence moments of $f(\underline{x})$ and $g(\underline{x})$ respectively, and Θ is the rotation on ζ that minimizes Ψ . Since the primary objective in the design of the control point matching scheme is to improve outlier error robustness, the AISIR system introduces the following modified Geman–McClure similarity objective function,

$$\Psi(v_m^f(\underline{x}), v_m^g(\underline{x} + \underline{\delta})) = \frac{(v_m^f(\underline{x}) - v_m^g(\underline{x} + \underline{\delta}))^4}{1 + (v_m^f(\underline{x}) - v_m^g(\underline{x} + \underline{\delta}))^4}. \quad (16)$$

The advantage of the proposed similarity objective function over commonly used quadratic and correlation based similarity objective functions is that it is an effective soft-re-descending error norm where the influence of outlier error is bounded and tends to zero asymptotically. The influence of outlier error on an objective function can be illustrated by studying its derivative (Black and Rangarajan, 1994). Fig. 1 shows the influence of outlier error on the quadratic and the proposed modified Geman–McClure objective function. The influence of outliers using a quadratic function increases linearly and without bound, while the influence of outlier error using the proposed objective function is bounded and tends to zero asymptotically. As such, the proposed objective function is highly robust to outlier error and well-suited for control point matching for inter-sensor and inter-band registration.

Since evaluating all possible point locations in $g(\underline{x})$ is not computationally tractable, only the set of control points $p_{g,0}, p_{g,1}, \dots, p_{g,q}$ in $g(\underline{x})$ determined by the control point detection scheme are evaluated to find the best match for each of control point $p_{f,0}, p_{f,1}, \dots, p_{f,r}$ in $f(\underline{x})$,

$$\underline{p}_{g,i} = \arg \min_{(\underline{p}_g)} \left[\sum_{\zeta(\underline{p}_{f,i})} \Psi(v_m^f(\Theta \underline{x}), v_m^g(\underline{x} + (\underline{p}_g - \underline{p}_{f,i}))) \right]. \quad (17)$$

4.4. Iterative control point pair refinement and mapping function estimation

After identifying matching control point pairs, the mapping function that aligns the sensed image and the reference image together must be estimated. A typical approach to estimating the mapping function is through the use of least squares solvers such as the normalized direct linear transformation (DLT) algorithm (Hartley and Zisserman, 2001). Two main factors contribute heavily to the effectiveness of the mapping function estimation process: (i) control point matching accuracy, and (ii) control point pair localization accuracy. While much effort has been focused on addressing the problem of control point matching accuracy through the use of outlier rejection schemes such as the Random Sample Consensus (RANSAC) algorithm (Fischler and Bolles, 1981) and the Maximum Distance Sample Consensus (MDSAC) algorithm (Wong and Clausi, 2007), little attention has been paid to the problem of control point pair localization accuracy. To address the issues associated with control point matching accuracy as well as control point matching pair localization accuracy, the AISIR system introduces a three-stage approach consisting of: (i) a control point rejection stage, (ii) an iterative control point pair control point refinement stage, and (iii) a robust mapping function estimation stage.

- *Stage 1:* The Maximum Distance Sample Consensus (MDSAC) algorithm (Wong and Clausi, 2007) is applied to the set of con-

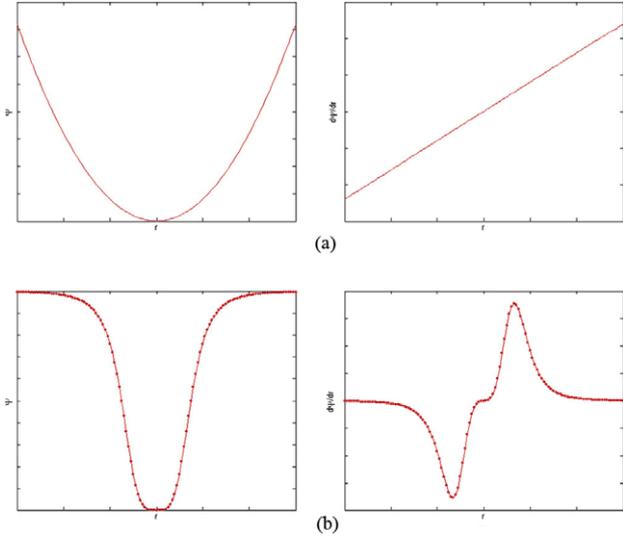


Fig. 1. Objective functions (left) and the corresponding influence of outliers on the objective function (right) for (a) quadratic objective function, and (b) proposed modified Geman–McClure objective function.

control point pairs identified during the control point matching stage in an attempt to reject erroneous control point pairs. The objective of this stage is to reduce the influence of the control point matching accuracy problem on the estimated mapping function.

- *Stage 2:* The locations of the remaining γ control points in the sensed image are refined in an iterative manner to minimize the modified Geman–McClure objective function between the maximum complex wavelet phase coherence moments around the control point in the sensed image \underline{p}_f and the control point in the reference image \underline{p}_g . The control point refinement problem can be formulated as an optimization problem, where the refined location of the control point in the sensed image $\hat{\underline{p}}_f$ is estimated as

$$\hat{\underline{p}}_{f,i} = \underline{p}_{g,i} - \hat{\delta}, \quad (18)$$

where,

$$\hat{\delta} = \arg \min_{(\hat{\delta})} \left[\sum_{\zeta(\underline{p}_{g,i})} \Psi(v_m^g(\Omega \underline{x}), v_m^f(\underline{x} + \hat{\delta})) \right]. \quad (19)$$

where Ω is the transformation on ζ that minimizes Ψ . In AISIR, Ω is defined by rotation and scale. The current location of the control point in the sensed image is used as the initial estimate for $\hat{\underline{p}}_f$ and is re-estimated iteratively on a sub-pixel basis using an iterative sequential quadratic programming (SQP) solver (Boggs and Tolle, 1995) until convergence is reached. The objective of this stage is to reduce the influence of the control point pair localization problem on the estimated mapping function.

- *Stage 3:* The mapping function that aligns the sensed image with the reference image is estimated based on the refined set of γ control point pairs. To improve the robustness of the mapping function estimation process to the control point mismatches and localization errors that were not dealt with by the previous two stages, AISIR introduces a mapping function estimation scheme based on the proposed modified Geman–McClure M-estimation scheme,

$$\hat{T}(\underline{x}) = \arg \min_{T(\underline{x})} \left[\sum_{i=1}^{\gamma} \frac{(T(\underline{p}_{f,i}) - \underline{p}_{g,i})^4}{1 + (T(\underline{p}_{f,i}) - \underline{p}_{g,i})^4} \right]. \quad (20)$$

where T is the mapping function as defined in Eq. (1). The AISIR system supports a projective transformation model or a high-order polynomial transformation model for the mapping function for situations where the images exhibit more complex geometric distortions relative to each other.

5. Experiments

The effectiveness of the AISIR system was evaluated using six sets of satellite images provided by US Geological Survey (USGS), and one set of satellite image from the Spaceborne Imaging Radar-C/X-Band Synthetic Aperture Radar (SIR-C/X-SAR) project. All test images are 8-bit grayscale images and each set of test images consists of a reference image and a sensed image. For all test cases, the number of initial control points extracted by the control point detection process was set to a maximum of 300. To demonstrate the effectiveness of the AISIR system when no additional information is available, the images in each test set are registered without the aid of georeferencing data. The registration accuracy of the AISIR system is evaluated by computing the root mean squared error (RMSE) between 30 manually selected ground-truth control point pairs chosen by a trained expert between the reference image and the sensed image on a pixel basis for each set of test images. RMSE is computed as the average of five test trials given the randomness in the registration process. For comparison purposes, the state-of-the-art M-SSD method proposed by Orchard (2007), ARRSI method proposed by Wong and Clausi (2007), and our preliminary work with robust complex wavelet feature representations (Wong and Clausi, 2008) (RCWF) were also evaluated. The M-SSD method formulates the registration problem as an affine least squares problem and solves the problem in an efficient manner in the frequency domain, while the ARRSI method provides efficient least squares registration based on a phase congruency model. The M-SSD and ARRSI methods are both state-of-the-art methods and have already been shown to provide improved registration accuracy for registering remote sensing images when compared to other methods such as cross-correlation (Capel and Zisserman, 1998), mutual information (Viola and Wells, 1997), and correlation ratio (Roche et al., 1998). Finally, a perspective transformation model was used for all of the tested methods.

5.1. Test sets

A description of each of the inter-band (IB) and inter-sensor (IS) test sets is presented in Table 1. There are two main challenges associated with the inter-band test sets. First, the reference and sensed images in each test set are acquired within different spectral bands (0.63–0.69 μm for Band 3 and 1.55–1.75 μm for Band 5). As such, corresponding intensities between the images are very different and cannot be compared directly. Second, the reference and sensed images in each test set are acquired at different times. Therefore, the environmental conditions under which the images are captured can be very different, with varying illumination and contrast non-uniformities. In addition to these two challenges, an additional challenge associated with the inter-sensor test sets is that the image characteristics captured in the reference and sensed images can differ significantly due to the characteristics of the underlying sensing technologies.

5.2. Experiment 1

The first set of experiments investigates the effectiveness of the AISIR system in registering images acquired under different spectral bands and different sensor technologies under typical conditions where the resolution is known. The registration accuracy

Table 1
Summary of inter-band (IB) and inter-sensor (IS) test image pairs.

Test			Reference			Sensed		
Name	Source	Lat/Long	Sensor	Res	Date	Sensor	Res	Date
IB1	USGS	46.0/–83.8	Landsat 7 ETM+, Band: 3	240 m	2003/4/12	Landsat 4–5 TM, Band: 5	240 m	2006/6/15
IB2	USGS	69.6/–92.7	Landsat 7 ETM+, Band: 5	240 m	2000/7/24	Landsat 4–5 TM, Band: 3	240 m	1999/7/6
IB3	USGS	46.0/–113.1	Landsat 7 ETM+, Band: 5	240 m	2001/8/17	Landsat 4–5 TM, Band: 3	240 m	2006/2/12
IB4	USGS	48.9/–68.8	Landsat 7 ETM+, Band: 5	240 m	1999/12/14	Landsat 4–5 TM, Band: 3	240 m	2005/10/13
IS1	JPL	48.9/–68.8	Landsat 7 ETM+, Band: 3	240 m	1999/12/14	SIR-C/X-SAR, Band: C	200 m	1994/04/15
IS2	USGS	29.0°59' / –90°01'	Orthorectified air-photo	~	~	LIDAR	~	~
IS3	USGS	29.0°57' / –90°02'	Orthorectified air-photo	~	~	LIDAR	~	~

results are shown in Table 2. The AISIR system outperforms the other registration methods under evaluation in terms of RMSE for all sets of test images. An example of the inter-band registration achieved using the AISIR system is shown in Fig. 2. By visual inspection, the registration results produced by the AISIR system are accurate for the set of inter-band images. Successful registration was achieved between the images despite the presence of cloud cover in the Band 3 image, as well as significant intensity differences between the two images. These results demonstrate the effectiveness of the AISIR system in registering inter-band images. Similar to the inter-band registration results, the AISIR system noticeably outperforms the other registration methods under evaluation in terms of RMSE for all sets of test images. An example of the inter-sensor registration achieved using the AISIR system is shown in Fig. 3. By visual inspection, the registration results produced by the AISIR system are accurate for the set of inter-sensor images. Successful registration was achieved between the images despite the image detail and intensity differences between the two images. These results demonstrate the effectiveness of the AISIR system in registering inter-sensor images.

5.3. Experiment 2

The second set of experiments investigates the robustness of the AISIR system under conditions where the resolution and the

Table 2
Registration accuracy for inter-band and inter-sensor images. “–” indicates the failure to register images within RMSE < 100.

Test	RMSE (pixels)			
	M-SSD	ARRSI	RCWF	AISIR
IB1	55.6633	3.7815	2.6837	2.2653
IB2	2.3664	1.6426	1.2880	1.0169
IB3	6.5345	1.2184	1.2013	0.9158
IB4	3.7815	2.7906	1.9380	1.6980
IS1	–	8.5854	3.7326	3.2164
IS2	21.168	4.4519	4.0136	3.8791
IS3	8.962	–	4.6428	4.2677

relative orientation between the images are unknown and relatively large compared to typical conditions. This was accomplished by systematically distorting the test image sets to relatively large scales and orientations. These distortions serve to provide a good test of robustness under uncertainty for the AISIR system.

The AISIR system and the other tested methods were used to register the distorted test image sets without information pertaining to scale or orientation, and the registration accuracy was evaluated in terms of RMSE. Tests were also performed in situations where the images were distorted by translations of {20%, 40%, 80%} of the reference image, but since the tested methods are largely translation-invariant, the registration accuracy results for all methods are within 1% to those shown in Table 2. Similarly, tests were performed in situations where the images were distorted by a set of orientations {5°, 10°, 15°, 20°}, but the registration accuracy results for all methods are within 3% to those shown in Table 2. Of greater interest are the registration accuracy results for large

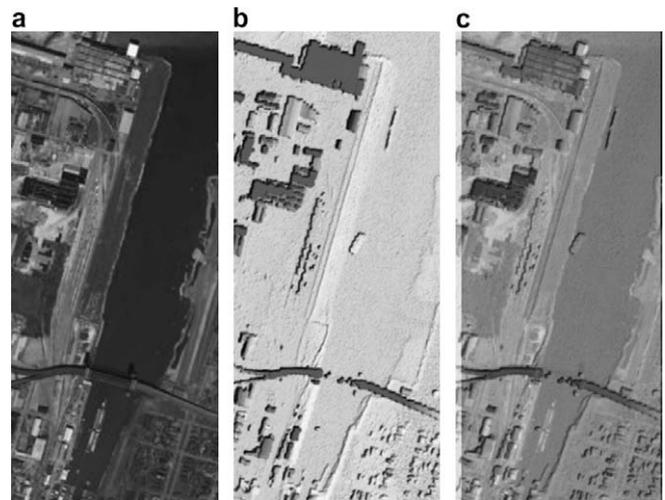


Fig. 3. Image registration from IS2: (a) air-photo; (b) LIDAR; (c) registration result.

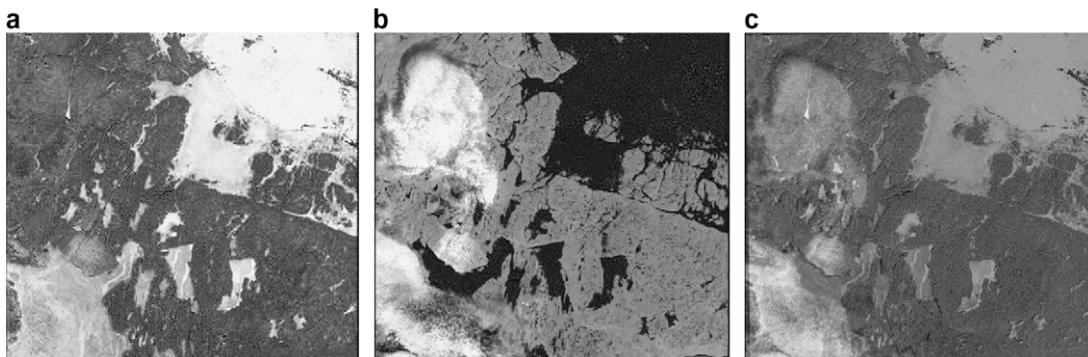


Fig. 2. Image registration from IB2: (a) Landsat 7 ETM+, Band: 5; (b) Landsat 4–5 TM, Band: 3; (c) registration result.

Table 3

Registration accuracy for inter-band and inter-sensor images under orientation distortions of 45°. “–” indicates the failure to register images within RMSE < 100.

Test	RMSE (pixels)			
	M-SSD	ARRSI	RCWF	AISIR
IB1	56.0928	–	–	2.3841
IB2	2.5717	–	–	1.1966
IB3	7.0157	–	–	1.0966
IB4	3.6923	–	–	1.7712
IS1	–	–	–	3.3872
IS2	21.6405	–	–	3.9867
IS3	9.1520	–	–	4.4079

Table 4

Registration accuracy for inter-band and inter-sensor images under relative scale distortions of 1.3. “–” indicates the failure to register images within RMSE < 100.

Test	RMSE (pixels)			
	M-SSD	ARRSI	RCWF	AISIR
IB1	–	4.5693	3.2926	2.3724
IB2	–	3.0296	2.1472	1.1804
IB3	–	2.8631	2.1598	1.0574
IB4	–	4.3173	3.172	1.8228
IS1	–	9.2167	4.1143	3.4338
IS2	–	6.3286	5.2391	3.9469
IS3	–	–	5.8894	4.3816

orientation distortions of 45° and scale distortions of 1.3, which are shown in Tables 3 and 4, respectively.

The AISIR system outperforms the other registration methods in terms of RMSE for all sets of test images. The registration accuracy of the AISIR and M-SSD methods remain comparable to the registration accuracy shown in Table 2 under large orientation distortions, while the other methods fail to produce reasonable registration results as shown in Table 3. An image example of inter-sensor registration achieved using the AISIR system under large orientation distortion is shown in Fig. 4. Successful registration was achieved between the images despite large orientation distortions in the SAR image, as well as significant differences in the intensities between the two images.

The registration accuracy of the AISIR method remain comparable to the registration accuracy shown in Table 2 under scale distortions, while the ARRSI and RCWF methods both showed a noticeable reduction in registration accuracy and the M-SSD method failed to produce reasonable registration results as shown in Table 4. An image example of inter-sensor registration achieved using the AISIR system under scale distortions is shown in Fig. 5. Successful registration was achieved between the images despite large scale distortions in the LIDAR image, as well as significant differences in the intensities between the two images. The higher registration accuracy of the AISIR system compared to the other tested methods under both test scenarios is largely due to the fact that both scale and rotation considerations are taken into account during the control point matching and control point refinement

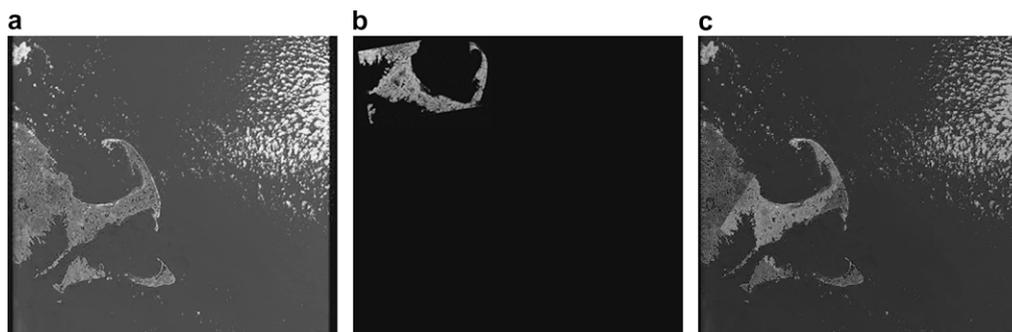


Fig. 4. Image registration from IS1 under orientation distortion of 45°: (a) Landsat 7 ETM+, Band: 3; (b) SAR; (c) registration result.

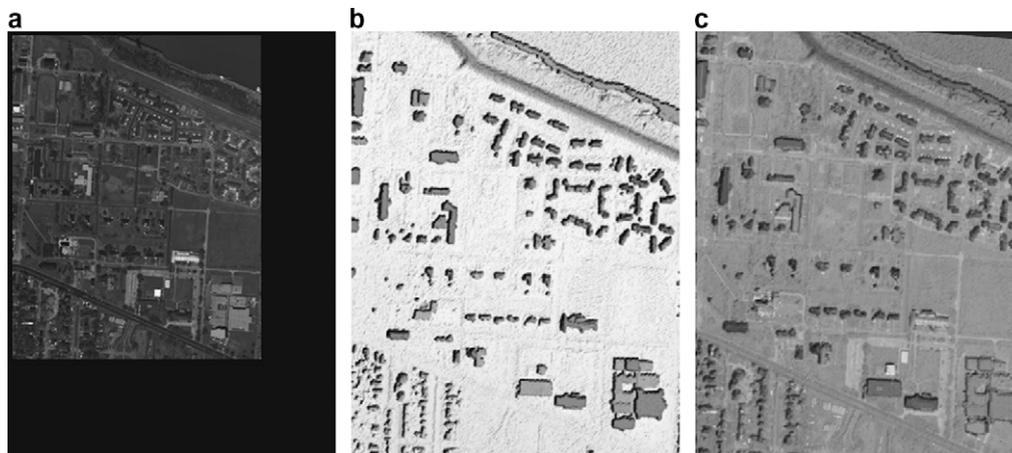


Fig. 5. Image registration from IS3 under scale distortion of 1.3: (a) air-photo; (b) LIDAR; (c) registration result.

processes of the AISIR system. The experimental results demonstrate the robustness of the proposed AISIR system to large distortions of scale and orientation.

6. Conclusions

A novel automated system name AISIR was introduced for the registration of inter-band and inter-sensor remote sensing images. A novel modified Geman–McClure control point matching scheme based on robust complex wavelet feature representations was used to address issues associated with contrast non-uniformities, noise, and inter-sensor and inter-band intensity mapping differences. An iterative control point pair refinement method was introduced for improving control point pair localization. Finally, a robust mapping function estimation scheme was introduced to further improve robustness to outlier data. Experimental results indicate that AISIR is capable of outperforming the state-of-the-art M-SSD and ARRSI registration methods for both inter-band and inter-sensor images.

Acknowledgement

This research has been sponsored by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Geomatics for Informed Decisions (GEOIDE). The authors would also like to thank USGS and NASA/JPL for the test data.

References

- Ali, M., Clausi, D. A., 2002. Automatic registration of SAR and visible band remote sensing images. In: IEEE Internat. Geoscience and Remote Sensing Symp., vol. 3, pp. 1331–1333.
- Black, M., Rangarajan, A., 1994. The outlier process: Unifying line processes and robust statistics. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 15–22.
- Boggs, P., Tolle, J., 1995. Sequential quadratic programming. *Acta Numerica*, 151.
- Capel, D., Zisserman, A., 1998. Automated mosaicing with super-resolution zoom. In: Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 885–891.
- Chen, H., Varshney, P., Arora, M., 2003. Mutual information based image registration for remote sensing data. *Internat. J. Remote Sens.* 24 (18), 3701–3706.
- Chen, L., Chiang, T., Teo, T., 2005. Fusion of LIDAR data and high resolution images for forest canopy modelling. In: Proc. 26th Asian Conf. on Remote Sensing.
- Cole-Rhodes, A., Johnson, K., LeMoigne, J., Zavorin, I., 2003. Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *IEEE Trans. Image Process.* 12 (12), 1495–1511.
- Eugenio, F., Marques, F., Marcello, J., 2002. A contour-based approach to automatic and accurate registration of multitemporal and multisensor satellite imagery. In: Proc. IEEE Internat. Geoscience and Remote Sensing Symp., vol. 6, pp. 3390–3392.
- Fan, Y., Ding, M., Liu, Z., Wang, D., 2007. Novel remote sensing image registration method based on an improved SIFT descriptor. In: Proc. of SPIE 6790, 67903G.
- Field, D., 1987. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Amer. A* 4 (12), 2379–2394.
- Fischler, M., Bolles, R., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM* 24, 381–395.
- Geman, S., McClure, D., 1987. Statistical methods for tomographic image reconstruction. *Bull. Internat. Statist. Inst.* LII-4, 521.
- Hartley, R., Zisserman, A., 2001. *Multiple View Geometry in Computer Vision*. Cambridge University Press.
- Huber, P., 1981. *Robust Statistics*. Wiley-Interscience.
- Kovesi, P., 2003. Phase congruency detects corners and edges. In: Proc. Australian Pattern Recognition Society Conf., pp. 309–318.
- Morrone, M., Burr, D., 1988. Feature detection in human vision: A phase-dependent energy model. *Proc. Royal Soc. Lond. B* 235, 221–245.
- Orchard, J., 2005. Efficient global weighted least-squares translation registration in the frequency domain. In: Proc. Internat. Conf. on Image Analysis and Recognition, pp. 116–124.
- Orchard, J., 2007. Efficient least squares multimodal registration with a globally exhaustive alignment search. *IEEE Trans. Image Process.* 10 (16), 2526–2534.
- Orchard, J., 2008. Multimodal image registration using floating regressors in the joint intensity scatter plot. *Med. Image Anal.* 12 (4), 385–396.
- Roche, A., Malandain, G., Pennec, X., Ayache, N., 1998. The correlation ratio as a new similarity measure for multimodal image registration. In: Proc. of First Internat. Conf. on Medical Image Computing and Computer-Assisted Intervention, vol. 1496, pp. 1115–1124.
- Roche, A., Malandain, G., Ayache, N., 2000. Unifying maximum likelihood approaches in medical image registration. *Internat. J. Imaging Systems Technol.* 11 (1), 71–80.
- Rottensteiner, F., Jansa, J., 2002. Automatic extraction of building from LIDAR data and aerial images. In: Proc. Internat. Society for Photogrammetry and Remote Sensing, vol. 34, No. 4, pp. 295–301.
- Viola, P., Wells, W., 1997. Alignment by maximization of mutual information. *Internat. J. Comput. Vision* 24 (2), 137–154.
- Wong, A., 2008. Adaptive bilateral filtering of image signals using local phase characteristics. *Signal Process.* 88 (6), 1615–1619.
- Wong, A., 2008. An iterative approach to improved local phase coherence estimation. In: Proc. Canadian Conf. on Computer and Robot Vision, pp. 301–307.
- Wong, A., Clausi, D.A., 2007. ARRSI: Automatic registration of remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 45 (5), 1483–1493. Part II.
- Wong, A., Clausi, D.A., 2008. Automatic registration of inter-band and inter-sensor images using robust complex wavelet feature representations. In: Proc. of 5th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS 2008).
- Xiaolong, D., Khorram, S., 1999. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. *IEEE Trans. Geosci. Remote Sens.* 37 (5), 2351–2362. Part II.
- Yasein, M., Agathoklis, P., 2008. A feature-based image registration technique for images of different scale. *Circuits and Systems*. In: Proc. of ISCAS 2008, pp. 3558–3561.