Article

An Automated Approach for Sub-Pixel Registration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) Imagery

Lin Yan *, David P. Roy, Hankui Zhang, Jian Li and Haiyan Huang

Geospatial Sciences Center of Excellence, South Dakota State University, Brookings, SD 57007, USA;
David.Roy@sdstate.edu (D.P.R.); hankui.zhang@sdstate.edu (H.Z.); Jian.Li@sdstate.edu (J.L.);
Haiyan.Huang@sdstate.edu (H.H.)
* Correspondence: lin.yan@sdstate.edu; Tel.: +1-605-688-6591

Abstract: Moderate spatial resolution satellite data from the Landsat-8 OLI and Sentinel-2A MSI sensors together offer 10 m to 30 m multi-spectral reflective wavelength global coverage, providing the opportunity for improved combined sensor mapping and monitoring of the Earth's surface. However, the standard geolocated Landsat-8 OLI L1T and Sentinel-2A MSI L1C data products are currently found to be misaligned. An approach for automated registration of Landsat-8 OLI L1T and Sentinel-2A MSI L1C data is presented and demonstrated using contemporaneous sensor data. The approach is computationally efficient because it implements feature point detection across four image pyramid levels to identify a sparse set of tie-points. Area-based least squares matching around the feature points with mismatch detection across the image pyramid levels is undertaken to provide reliable tie-points. The approach was assessed by examination of extracted tie-point spatial distributions and tie-point mapping transformations (translation, affine and second order polynomial), dense-matching prediction-error assessment, and by visual registration assessment. Two test sites over Cape Town and Limpopo province in South Africa that contained cloud and shadows were selected. A Landsat-8 L1T image and two Sentinel-2A L1C images sensed 16 and 26 days later were registered (Cape Town) to examine the robustness of the algorithm to surface, atmosphere and cloud changes, in addition to the registration of a Landsat-8 L1T and Sentinel-2A L1C image pair sensed 4 days apart (Limpopo province). The automatically extracted tie-points revealed sensor misregistration greater than one 30 m Landsat-8 pixel dimension for the two Cape Town image pairs, and greater than one 10 m Sentinel-2A pixel dimension for the Limpopo image pair. Transformation fitting assessments showed that the misregistration can be effectively characterized by an affine transformation. Hundreds of automatically located tie-points were extracted and had affine-transformation root-mean-square error fits of approximately 0.3 pixels at 10 m resolution and dense-matching prediction errors of similar magnitude. These results and visual assessment of the affine transformed data indicate that the methodology provides sub-pixel registration performance required for meaningful Landsat-8 OLI and Sentinel-2A MSI data comparison and combined data applications.

Keywords: registration; Landsat-8; Sentinel-2; area-based least squares matching; feature based matching

1. Introduction

Moderate spatial resolution satellite data from the similar polar-orbiting sun-synchronous Landsat-8 and Sentinel-2 sensors together provide the opportunity for improved mapping and

monitoring of the Earth's surface [1]. Landsat-8 carries the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) that sense 11 spectral bands including eight 30 m reflective wavelength bands, one 15 m panchromatic band, and two 100 m thermal wavelength bands [2]. The Landsat-8 swath is approximately 185 km (15° field of view from an altitude of 705 km) and provides a global coverage of the Earth's surface every 16 days [2]. Sentinel-2A carries the Multi Spectral Instrument (MSI) that has 13 spectral bands ranging from 0.433 μm to 2.19 μm, including four 10 m visible and near-infrared bands, six 20 m red edge, near-infrared and short wave infrared bands, and three 60 m bands [3]. The Sentinel-2A swath is approximately 290 km (20.6° field of view from an altitude of 786 km) and provides a global coverage every 10 days and with the planned launch of a follow on identical Sentinel-2B sensor will provide 5-day global coverage [3]. Combined, the Sentinel-2 and Landsat-8 sensors will provide 10 m to 30 m multi-spectral reflective wavelength global coverage approximately every 3 days.

This paper describes the automated registration of Landsat-8 and Sentinel-2A reflectance data into the same common coordinate system. The geometrically corrected sensor data are available for Sentinel-2A as L1C top-of-atmosphere (TOA) tiles [4,5] and for Landsat-8 as L1T TOA images defined in a Worldwide Reference System (WRS) path/row coordinate system [6,7]. The geolocation performance specification for Sentinel-2A is 12.5 m (3σ) [8] and for Landsat-8 is 12 m (90% circular error) [9]. However, the Sentinel-2A L1C and Landsat-8 L1T data are currently misaligned relative to each other by more than several 10 m pixels [10]. This is because although both sensor geolocation systems use parametric approaches, whereby information concerning the sensing geometry is modeled and the sensor exterior orientation parameters (attitude and position) are measured, they use different ground control and digital elevation models to refine the geolocation [8,9]. The Landsat geolocation uses a global sample of ground control points derived from the Global Land Survey (GLS) cloud-free single-date Landsat images that are defined for each WRS path/row for different decades [11,12]. The Sentinel-2A geolocation will be improved by using a global reference image derived from an orthorectified set of Sentinel-2A cloud-free images [13,14]. Unfortunately, a relative misalignment of the Landsat-8 L1T and Sentinel-2A L1C data has been observed that varies among Landsat WRS path/row locations due primarily to variable GLS path/row locational accuracies [9]. Consequently, the GLS data are being reprocessed to provide a better match with the ground control used for operational geolocation of the Sentinel-2A data [10]. A more detailed description of the causes of the sensor misregistration is provided in [15].

A large body of research has been published concerning the registration of satellite images [16]. Methods are divided broadly into area-based matching methods, whereby a small region of one image is moved systematically across the other image and the location that provides the highest reflectance correlation provides a tie-point, and feature-based methods where tie-points are found by locating the positions of high-contrast features common to both images [17]. Given a sufficient number of tie-points, a transformation function, usually expressed as two polynomial functions that map the x and the y pixel locations of one image to the other image, is derived, often by least-squares regression analysis. For highly distorted data such as airborne imagery, local transformation functions are needed [18–20]. Different satellite data registration methods have been refined and proposed, for example, with respect to the initial knowledge of the relative orientation and scale of the images, to include computational efficiencies required to improve the matching speed, and to robustly handle cloud occlusion, land surface changes and differences between sensors [16].

In this study, a hierarchical image matching approach, which was originally developed for registration of High-Resolution Imaging Science Experiment (HiRISE) single-band stereo images to derive digital elevation models of Mars [21], was refined for registration of Landsat-8 L1T and Sentinel-2A L1C data. This approach was adopted because it has been proven for operational automated processing. Moreover, it uses an efficient feature- and area-based matching approach, is robust to noise, generates a large number of dense matches in a computationally efficient manner, and works well when the relative orientation and location of the images to be registered are

known approximately. In this study, the approach was refined and applied to the 10 m Sentinel-2A near-infrared (NIR) and the 30 m Landsat-8 NIR bands. The NIR bands were selected because NIR reflectance has a greater range over soil, vegetation, and water, than visible wavelengths [22], and so usually the NIR provides high spatial contrast suitable for area- and feature-based matching of images acquired with similar dates (as in this study). In addition, the NIR is less sensitive to atmospheric contamination than at visible wavelengths [23]. We note that the NIR is commonly used for registration of moderate and high spatial resolution satellite data [17,24,25]. The Landsat-8 15 m panchromatic band, which covers predominantly the green and red wavelengths (503 to 676 nm) [26], was not used in this study because there is no spectrally similar Sentinel-2A band.

The paper is organized as follows. First, the Sentinel-2A L1C and Landsat-8 L1T geometric data characteristics and the common projection and tiling scheme used to reproject the data are described and illustrated for the test data that are selected in South Africa where approximately contemporaneous Sentinel-2A and Landsat-8 were available. The registration methodology and assessment approach are then described. Quantitative and qualitative results are presented to illustrate the registration performance and examples of the misregistration between Landsat-8 and Sentinel-2A data. The paper concludes with a discussion on the results and implications for combined Landsat-8 and Sentinel-2A data applications.

2. Overview of Landsat-8 L1T and Sentinel-2A L1C Geometric Data Structure, the Common Map Projection, and the Study Test Data

2.1. Landsat-8 L1T and Sentinel-2A L1C Geometric Data Structure

The Sentinel-2A geometrically corrected L1C products are defined in approximately 110×110 km tiles referenced to the U.S. Military Grid Reference System [4,5]. The Landsat-8 geometrically corrected L1T products are defined by approximately 180×180 km images referenced by WRS path/row [6,7]. Both sensor data are geolocated in the Universal Transverse Mercator (UTM) map projection in the World Geodetic System 84 (WGS84) datum and with specified UTM zones. Each UTM zone covers $6°$ of longitude, and so forms the basis of a separate map projection [27]. The Sentinel-2A swath is approximately 290 km and encompasses more than one UTM zone. The Sentinel-2A L1C data are defined in tiles that spatially overlap, and adjacent L1C tiles may be defined in different UTM zones [4,5]. The Sentinel-2A L1C tiled data structure, although complex, ensures that the geographic coordinates of each L1C pixel are fixed. The Landsat-8 185 km swath is sufficiently narrow for each L1T image to be defined in a single UTM zone. However, the geographic coordinates of each Landsat L1T pixel are not fixed.

2.2. Map Projection Used for Registration

Over large areas, users of derived satellite information require spatially-explicit map products defined in a single projection rather than different UTM projections. Therefore, in this study, the Sentinel-2A L1C and Landsat-8 L1T data were reprojected to a common coordinate system. The equal-area sinusoidal projection and tiling scheme used to store the global Web Enabled Landsat (WELD) products were utilized. The global WELD products define monthly and annual 30 m Landsat nadir BRDF-adjusted reflectance (NBAR) surface reflectance derived by the algorithms described in [22,27,28] and are available at [29]. The global WELD tiles are nested within the standard $10° \times 10°$ MODIS land product tiles and are defined in the sinusoidal equal projection used to store the MODIS land products [30]. Each global WELD tile covers about 159×159 km. There are 7×7 global WELD tiles within each MODIS land tile, and the filename includes the MODIS horizontal (0 to 35) and vertical (0 to 17) tile coordinates, and the nested WELD tile horizontal and vertical tile coordinates (0 to 6).

The Sentinel-2A L1C 10 m NIR (842 nm) and the Landsat-8 L1T 30 m NIR (864 nm) bands were reprojected independently into the global WELD tiles, each resulting tile was composed of

15885 × 15885 10 m pixels. The computationally-efficient inverse gridding approach [30,31] was used to systematically relate the locations of each 10 m global WELD tile pixel to the Sentinel-2A L1C and Landsat-8 coordinates. Care was taken to use the correct UTM zones defining the sensor data and to handle the different sensor pixel georeferencing schemes (Sentinel-2A uses pixel corner and Landsat-8 uses pixel center references). The General Cartographic Transformation Package (GCTP), developed by the United States Geological Survey, was used to transform coordinates between the sinusoidal and UTM map projections. The GCTP has been used to develop a number of applications including the MODIS global browse imagery [32] and the WELD products [28]. The Sentinel-2A L1C 10 m NIR data were resampled to 10 m by nearest neighbor resampling considering all the spatially overlapping L1C tile data. Although nearest neighbor resampling is computationally efficient as it assigns the closest pixel to any output coordinate location, it introduces local resampling shifts up to 0.5 of the input image pixel dimension. This was not an issue for the registration results because the sensor-to-sensor registration transformations were fitted globally, *i.e.*, using a large number of tie points selected across each study image. However, the scale difference between the 10 m output grid and the 30 m Landsat data precluded reliable Landsat nearest neighbor resampling. Therefore, the Landsat-8 30 m NIR data were interpolated to 10 m by bilinear resampling [33], specifically, by fitting a hyperbolic paraboloid through the four neighboring 30 m pixel values and then interpolating the 10 m pixel value.

2.3. Test Data

The orbit and sensing geometry of Landsat-8 and Sentinel-2A are different, and consequently they acquire images over the same location at different times and on different days. In this study, two sites in South Africa, specifically over Cape Town (Figure 1) and over Limpopo province (Figure 2), were selected. Over Cape Town, a Landsat-8 L1T image sensed 22 November 2015 (for brevity referred to as week 47) and two Sentinel-2A L1C images, which were sensed 16 and 26 days later on 8 December and 18 December 2015 (referred to as weeks 49 and 51), were used to examine the robustness of the matching algorithm to surface, atmosphere and cloud changes. Over the Limpopo Province, the Landsat-8 L1T and Sentinel-2A data images were sensed just four days apart on 5 and 9 December 2015, respectively (both week 49). All the images contained cloud and/or topographic shadows to provide confidence that the matching could handle these phenomena. Both test sites include considerable (more than 500 m) changes in relief *i.e.*, Table Mountain in Cape Town to the surrounding coastal plain, and the Southern parts of the Waterberg Massif to the Springbok Flats in the Limpopo study area.
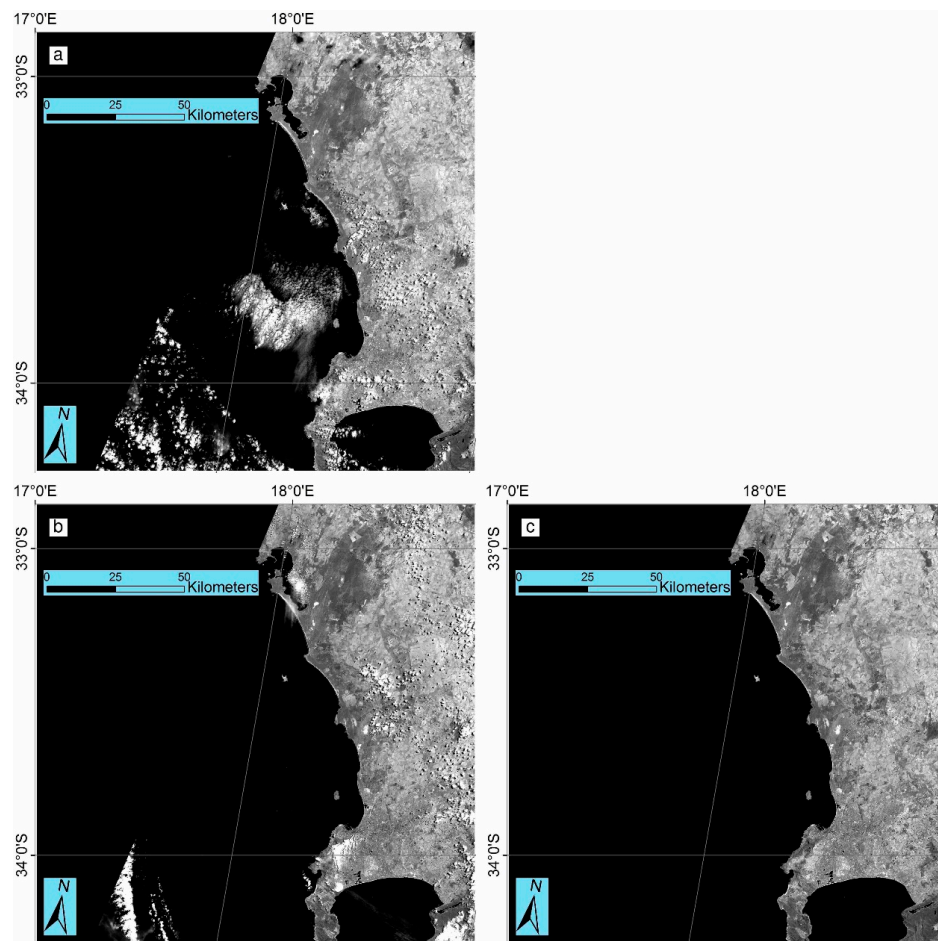
Figure 1. Cape Town, South Africa, test data showing (**a**) Landsat-8-L1T sensed 22 November 2015 (week 47); (**b**) Sentinel-2A L1C sensed 8 December 2015 (week 49); and (**c**) Sentinel-2A L1C sensed 18 December 2015 (week 51). The NIR (Sentinel-2: 842 nm and Landsat-8 864 nm band) TOA reflectance for each image is shown, which was reprojected to 10 m global WELD tile hh19vv12.h3v2 (sinusoidal projection, 15885 × 15885 10 m pixels).

Figure 2. Limpopo Province, South Africa, test data showing (**a**) Landsat-8-L1T sensed 5 December 2015 (week 49); (**b**) Sentinel-2A L1C sensed 9 December 2015 (week 49). The NIR TOA reflectance for each image is shown, which was reprojected to 10 m global WELD tile hh20vv11.h4v3 (sinusoidal projection, 15885 × 15885 10 m pixels).
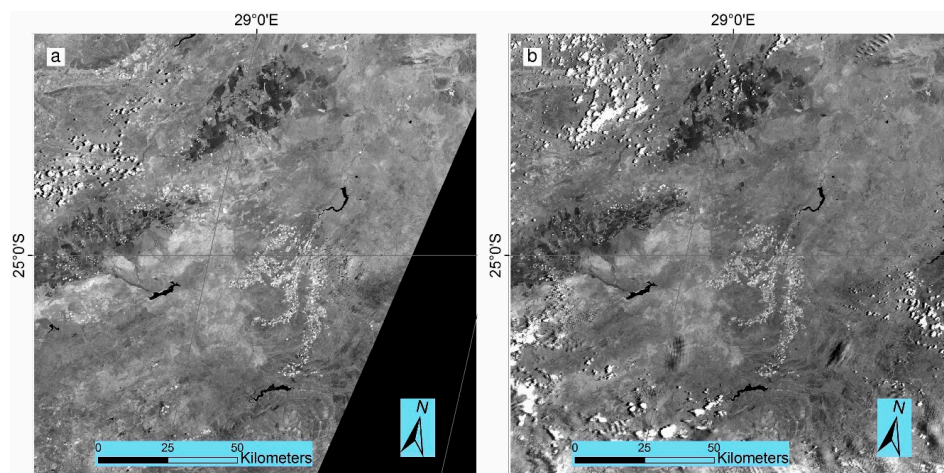
3. Registration Method

3.1. Overview

The registration method was adapted from the hierarchical matching approach developed for application to Martian High-Resolution Imaging Science Experiment (HiRISE) satellite stereo images [21]. The approach in [21] has four main steps: (1) image pyramids, *i.e.*, a hierarchy of low-pass filtered images from coarse spatial resolution to the native image resolution, are built for each image pair; (2) features are detected in one image at each pyramid layer; (3) area-based cross-correlation matching around the feature points is conducted sequentially from the coarse to the fine spatial resolution pyramid levels; (4) a set of tie-points defined at the native image spatial resolution are used, with knowledge of the sensor interior and exterior orientations, to derive a digital elevation model (DEM) using a dense grid-point matching guided by the tie-point locations.

The approach was adapted in this study (i) to handle the pairs of Sentinel-2A L1C and Landsat-8 L1T NIR data reprojected to the 10 m global WELD tiles (*i.e.*, illustrated in Figures 1 and 2, for brevity these are referred to below as Sentinel and Landsat image pairs); (ii) the area-based cross-correlation matching was replaced with the area-based least squares matching (LSM) that provides sub-pixel registration accuracy; (iii) a depth-first mismatch detection method was implemented that does not require knowledge of the sensors' interior and exterior orientation parameters; (iv) the dense matching approach was not used to generate a DEM but rather to examine the spatial pattern of the registration prediction errors. The approach produces a set of corresponding tie-point locations that are defined in the Sentinel and Landsat image pair. The tie-points are then used to derive transformation functions that are used to reproject the Landsat image into registration with the Sentinel image. Figure 3 illustrates the processing flow. The implementation details are described below.
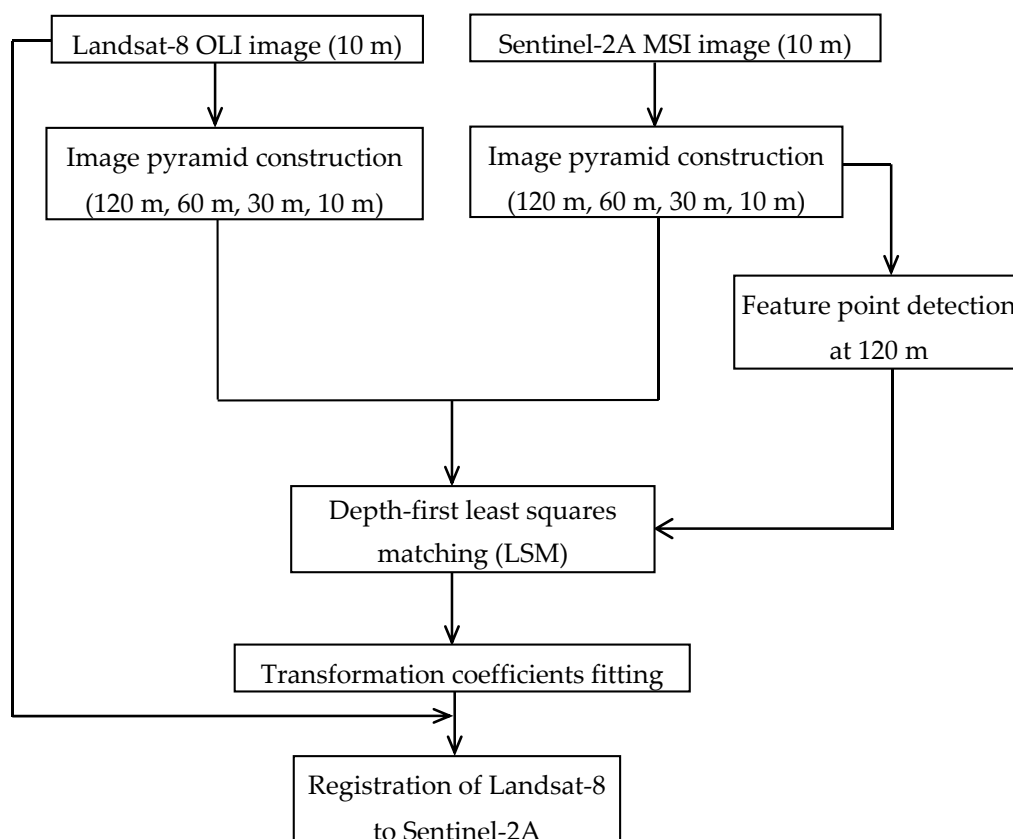

Figure 3. Automated workflow to register Landsat-8 OLI to Sentinel-2A MSI WELD tiles.

3.2. Image Pyramid Construction

A number of registration methods have been developed using image pyramids to improve the computational matching speed and sometimes to improve the match reliability [16,21,24,34]. The image pyramid approach was used in this study for both reasons and in particular because the depth-first mismatch detection method (Section 3.5) uses the multiple image pyramid levels to reduce mismatches. In this study, a four-level (120 m, 60 m, 30 m, 10 m) image pyramid was derived for the Landsat and Sentinel image pairs, applying a conventional low-pass Gaussian filter [35] to smooth the 10 m data to 30 m, 60 m and 120 m. The coarsest spatial resolution was defined at 120 m because the misregistration between the Landsat-8 L1T and Sentinel-2A L1C data was assumed to be smaller than this [10]. Four levels were used to ensure reliable depth-first mismatch detection described below.

3.3. Coarse Resolution 120 m Feature Point Detection

Feature-based matching techniques detect features, such as the intersections of linear structures and the centroids of distinct geometric objects, are often invariant to spectral, spatial, radiometric and temporal variations [36,37]. Consequently, they are commonly utilized for matching of satellite and other remotely sensed imagery [17,21,38–40]. Following the Li *et al.* approach [21], feature points were detected using the locally adaptive Förstner operator that typically detects corners, distinct points and centers of ellipsoid features [36]. The feature points were detected in the Sentinel-2A 120 m image as the Sentinel-2A data has higher native spatial resolution (10 m) than the Landsat-8 data (30 m). Each detected feature point was defined to the closest 120 m Sentinel-2A pixel. There were usually several thousand feature points detected from a 120 m Sentinel-2A image.

3.4. Least-Squares Area Based Image Matching

Area-based matching was undertaken by comparing the NIR reflectance values over a small square image patch (n × n pixels) surrounding each feature point location in the Sentinel image with the corresponding sized patch in the Landsat image. Rather than using conventional area-based cross-correlation matching [41,42], the least squares matching (LSM) approach was used [43,44]. The LSM approach models both geometric and radiometric transformations, which makes it suitable for matching images acquired on different dates and/or from different sensors, and also for matching images where the misalignment is not a simple translation [17,45]. LSM is a form of area-based matching and is undertaken between two small image patches by least-squares fitting of their reflectance values. It can provide sub-pixel matching accuracy up to 0.02 pixels depending on the image content [46]. Conventionally, a correlation coefficient, such as the normalized cross-correlation coefficient [42], is used to measure the similarity between one image patch (in the Sentinel image) and the corresponding image patch (in the Landsat image) which is resampled using the fitted geometric and radiometric transformation parameters. The least-squares fitting of these parameters is conducted iteratively, generating a new resampled patch each time, and checking that the correlation between the two patches increases after each new resampling, otherwise the iteration is terminated. The Landsat matched position corresponding to the center point of the Sentinel image patch is derived using the final fitted geometric transformation parameters. This produced matched pairs of Sentinel and Landsat pixel locations.

In the LSM implementation, the correlation coefficient was replaced by the spectral angle mapper (SAM) measure [47,48]. This was found to be helpful because of the spectral and temporal differences between the Sentinel and Landsat images. The SAM is insensitive to exogenous reflectance brightness variations as demonstrated using hyperspectral data [48–50] and multi-spectral multi-temporal data [51]. The SAM is conventionally derived between the spectral reflectance values of two pixels by calculating the angle subtended between their points in spectral feature space and the feature space origin (*i.e.*, zero reflectance). In this implementation, the SAM was derived by comparing the n^2 Sentinel NIR image patch pixel values with the corresponding n^2 Landsat NIR values. For the South

African test data in this study, the SAM was found to be more robust against threshold selection than the conventional normalized correlation coefficient, and a constant SAM threshold was sufficient for LSM matching for all four pyramid image levels.

3.5. Depth-First Mismatch Detection

Most matching algorithms incorporate strategies to remove mismatches that can occur, for example, in regions with repetitive structured terrain. A depth-first mismatch detection method was implemented that does not require knowledge of the sensors' interior and exterior orientation parameters and that takes advantage of the sub-pixel matching accuracy provided by the LSM. The method is illustrated in Figure 4. Rather than undertaking mismatch detection independently on individual pyramid image levels, for example as in [21] or [24], the depth-first mismatch detection method utilizes the hierarchical pyramid image structure by considering the "depth" of a given feature point.

A scale-space approach [52] was used, based on the concept that correctly matched features should occur at all pyramid levels, and that at coarser spatial resolution, a smaller number of prominent features will be detected with less precise location than at higher spatial resolution. First, for each feature detected in the Sentinel 120 m data, the LSM matching was applied to find a corresponding location in the Landsat 120 m data. A check was undertaken to ensure a successful match. If the match was successful, then the Sentinel and Landsat (not shown in Figure 4) matched locations were projected to the 60 m level, and the LSM was repeated at these locations using the 60 m data. This matching-and-projection procedure was repeated to the 30 m and 10 m pyramid level data. If the match failed at any level, as illustrated by the red points in Figure 4, the locations were discarded. If the match was successful at all levels (illustrated by the two green points on the 10 m level in Figure 4), then the location in the Sentinel and Landsat data was recorded as a tie-point.
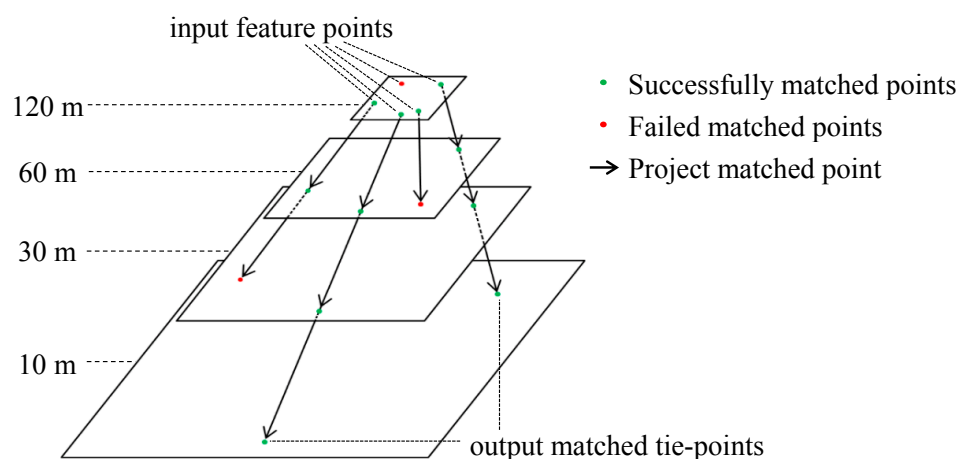

Figure 4. Illustration of depth-first LSM mismatch detection on the four-level image pyramid (shown for the Sentinel-2A image only).

A match was considered successful if the SAM value was $\geqslant 0.995$. This was a strict threshold as the SAM is bounded in the range [0, 1] and the SAM has a value of unity when the two image patches are identical. This was also helpful as cloudy 10 m pixels that are present in the 120 m, 60 m and 30 m pyramid levels will reduce the SAM values. Successful match locational criteria were also implemented. At the 120 m level, the difference between the Sentinel and Landsat locations was constrained to be less than 120 m to reduce matching blunders. At the 60 m, 30 m and 10 m levels, the Landsat LSM matched location was compared with the Landsat location projected from the level above, and if the Euclidean distance between the locations was greater than the one third of a pixel, it was also rejected. The LSM was undertaken considering 21 $\times$ 21 pixel patches in the Sentinel and

Landsat images for the 120 m, 60 m and 30 m levels and a slightly larger 35×35 pixel patch for the 10 m level because of the loss of detail in the Landsat 10 m data due to its native 30 m resolution.

3.6. Transformation Coefficient Fitting

The above processes result in a set of corresponding pixel locations (*i.e.*, tie-points) that are defined to the pixel in the Sentinel 10 m WELD tile and to sub-pixel precision in the Landsat 10 m WELD tile. These points were used to derive transformation functions used to reproject the Landsat WELD data into registration with the Sentinel WELD tile data. Conventionally, for moderately geometrically distorted imagery, polynomial functions that map the x and the y pixel locations of one image to the other image are derived by least-squares regression analysis of the tie-points. Conventional transformation functions were considered in this study to model translation, affine, and second order polynomial geometric distortions. Specifically: translations in both axes (Equation (1)), translation, rotation, and scaling in both axes and an obliquity (Equation (2)), and the same terms as Equation (2) with torsion and convexity in both axes (Equation (3)) [53].

$$
\begin{aligned}
x_L &= a_0 + x_s \\
y_L &= b_0 + y_s
\end{aligned}
\tag{1}
$$

$$
\begin{aligned}
x_L &= a_0 + a_1 x_s + a_2 y_s \\
y_L &= b_0 + b_1 x_s + b_2 y_s
\end{aligned}
\tag{2}
$$

$$
\begin{aligned}
x_L &= a_0 + a_1 x_s + a_2 y_s + a_3 x_s^2 + a_4 x_s y_s + a_5 y_s^2 \\
y_L &= b_0 + b_1 x_s + b_2 y_s + b_3 x_s^2 + b_4 x_s y_s + b_5 y_s^2
\end{aligned}
\tag{3}
$$

where x_L, y_L are the Landsat x and y WELD tile pixel coordinates, x_s, y_s are the Sentinel-2A WELD tile pixel coordinates, and a_i and b_i are the transformation coefficients derived by ordinary least squares regression of the tie-points found in the Landsat and Sentinel WELD tiles.

3.7. Image Registration

The pixel coordinates of each global WELD Sentinel tile pixel (x_s, y_s) were systematically projected into the global WELD Landsat tile coordinates (x_L, y_L), as Equations (1) and (2) or Equation (3), and then the coordinates (x_L, y_L) were reprojected using GCTP into the Landsat-8 L1T UTM map projection (Section 2.2). In this way, the Landsat-8 Level 1T data were only resampled once. The Landsat-8 L1T data were bilinear resampled to produce a new 30 m Landsat-8 WELD tile product that was registered with the corresponding Sentinel-2A WELD tile product.

4. Registration Assessment

4.1. Tie-Point Misregistration Assessment

The WELD tile x and y axis differences between the Landsat and Sentinel tie-point locations were derived as Equation (4). The differences were summarized quantitatively (minimum, maximum, mean and standard deviation) to characterize the sensor misregistration in units of 10 m pixels.

$$
\begin{aligned}
\Delta_{x,m} &= x_{S,m} - x_{L,m} \\
\Delta_{y,m} &= y_{S,m} - y_{L,m}
\end{aligned}
\tag{4}
$$

where $\Delta_{x,m}$ and $\Delta_{y,m}$ are the WELD tile x and y axis differences between the Landsat tie-point location $x_{L,m}$, $y_{L,m}$ and the corresponding Sentinel tie-point location $x_{S,m}$, $y_{S,m}$ for tie-point m. The differences calculated by Equation (4) were visualized as vectors to examine the spatial distribution of the tie-point locations and the spatial pattern of the sensor misregistration. Note that, following standard convention, the image origin is defined as the top-left, *i.e.*, the North-West, of each WELD tile, with

+*x* to the East (*i.e.*, to the right in the sample direction) and +*y* to the South (*i.e.*, downwards in the line direction).

4.2. Transformation Coefficient Fitting Assessment

The capability of the different transformation functions (Section 3.6) to accurately fit the tie-point locations was assessed by consideration of the root-mean-square error (RMSE) as Equation (5).

$$RMSE = \sqrt{\frac{\sum\limits_{m=1}^{n} (x_{L,m} - \hat{x}_{L,m})^2 + (y_{L,m} - \hat{y}_{L,m})^2}{n - t}} \qquad (5)$$

where RMSE is the root-mean-square error (units 10 m pixels), $x_{L,m}$, $y_{L,m}$ is the Landsat x and y WELD tile pixel coordinate for tie-point *m*, and $\hat{x}_{L,m}$, $\hat{y}_{L,m}$ are the transformed Landsat x and y WELD tile pixel coordinate for tie-point *m* calculated by Equations (1) and (2) or Equation (3), *n* is number of tie-points, and *t* is the number of coefficients in the transformations, *i.e.*, set as two for Equation (1), six for Equation (2) and twelve for Equation (3). Given *n* >> *t*, then smaller RMSE values are indicative of more accurate transformation function fitting.

For the Cape Town data (Figure 1), the temporal stability of the transformations was assessed by comparison of the transformations calculated for the two image pairs. Specifically, the pixel coordinates of the Cape Town WELD tile were systematically projected using the transformations derived from the Landsat-8 (week 47) and Sentinel-2A (week 49) tie-points, and also using the transformations derived from the Landsat-8 (week 47) and Sentinel-2A (week 51) tie-points, and then the Euclidean distance between them was computed as Equation (6).

$$d_k = \sqrt{\left(x_k^{47 \to 49} - x_k^{47 \to 51}\right)^2 + \left(y_k^{47 \to 49} - y_k^{47 \to 51}\right)^2} \qquad (6)$$

where d_k is the Euclidean distance (units 10 m pixels) for projected WELD tile pixel *k*, $(x_k^{47 \to 49}, y_k^{47 \to 49})$ is the transformed pixel *k* location calculated as Equations (1) and (2) or Equation (3) for Landsat-8 (week 47) to Sentinel-2A (week 49) tie-points, and $(x_k^{47 \to 51}, y_k^{47 \to 51})$ is the location calculated with Equations (1) and (2) or Equation (3) for Landsat-8 (week 47) to Sentinel-2A (week 51) tie-points. Summary statistics of d_k for all the 15885 × 15885 10 m WELD tile pixels were derived independently for each of the three transformations as Equations (1) and (2) or Equation (3).

4.3. Qualitative Visual Registration Assessment

Visual comparison of the registered Landsat and Sentinel data was undertaken using an approach similar to that used to visualize MODIS geolocation performance [30]. False color images composed of the NIR bands of the registered data were generated for spatial subsets containing high contrast features. The Sentinel NIR data were shown in red and the registered Landsat NIR data were shown in the blue and green bands. In this way, any sensor misregistration is exhibited by red and cyan tones over high contrast features and as greyscale tones elsewhere. The 10 m NIR Sentinel data were bilinear resampled from 10 m to 30 m, so they could be compared directly with the registered 30 m Landsat data.

4.4. Dense Grid-Point Matching Registration Assessment

The tie-point fitted transformations could be used to guide the least squares matching on a dense grid, and then the match results were used to locally register the Landsat and Sentinel data. This is not proposed as a practical registration solution, however, as it is computationally expensive and because spatial gaps between densely matched locations may remain due to, for example, clouds in either image. However, this approach was used to examine the Landsat and Sentinel misregistration in detail. Grid points were sampled every six 10 m pixel locations in the x and y axes across the 15885 × 15885 10 m

WELD tile. At each grid point location, the corresponding location in the Landsat 10 m WELD tile was transformed as Equations (1) and (2) or Equation (3) and then used to guide the least squares matching between the Landsat and Sentinel WELD 10 m NIR data. Similar mismatch rejection as described in Section 3.5 was undertaken, using the SAM 0.995 threshold and setting successful match locational criteria as 1.0 pixel for transformations as Equations (2) and (3) and as 1.6 pixels for transformation Equation (1). These successful match locational criteria thresholds were set as approximately three times the transformation RMSE values (*i.e.*, Equation (5) results described in Section 5). Maps of the prediction error defined as Equation (7) were derived for the three transformations.

$$ e_i = \sqrt{\left(x_{L,i} - \hat{x}_{L,i}\right)^2 + \left(y_{L,i} - \hat{y}_{L,i}\right)^2} \tag{7} $$

where for each WELD tile dense-matching grid point i, e_i is the prediction error (units 10 m pixels), $(\hat{x}_{L,i}, \hat{y}_{L,i})$ is the transformed Landsat x and y WELD tile pixel coordinate calculated by Equations (1) and (2) or Equation (3), and $(x_{L,i}, y_{L,i})$ is the least squares matched Landsat x and y WELD tile pixel coordinate.

5. Results

5.1. Cape Town

5.1.1. Tie-Point Misregistration Assessment

A total of 1346 and 2351 feature points were detected from the Sentinel-2A 120 m week 49 and 51 images, respectively. Fewer feature points were available in the Sentinel-2A week 49 120 m image because of the greater cloud cover (Figure 1 bottom row) and because of the Gaussian smoothing of the cloud edges that reduced the availability of high contrast features. After the depth-first mismatch detection process, a total of 116 tie-points were defined between the Landsat-8 week 47 and the Sentinel-2A week 49 image pair, and 797 tie-points were defined between the Landsat-8 week 47 and Sentinel-2A week 51 image pair. Figure 5 illustrates the locations of the tie-points for the two pairs of images (red and green vectors). The tie points were found across the tile except in regions where cloud occurred in one or both images, and the vectors were highly consistent. For both image pairs, the Landsat-8 image was misaligned in a similar south-west direction relative to the Sentinel-2A images.

Table 1 summarizes the tie-point differences illustrated in Figure 5. For both image pairs, the x-axis and y-axis mean shift magnitudes are greater than 5.2 and 2.1 pixels with standard deviations of about 0.4 and 0.3 pixels, respectively. These 10 m pixel shifts are not insignificant and even at the 30 m Landsat pixel resolution will limit the ability to meaningfully compare Landast-8 and Sentinel-2A data.

5.1.2. Transformation Coefficient Fitting Assessment

Tables 2 and 3 summarize the transformation coefficients derived by ordinary least squares regression fitting of the 797 and 116 tie-points extracted from the Cape Town image pairs (Figure (5)) for Equations (1)–(3). The transformation fittings' RMSE values (Equation (5)) are also tabulated. As there are an order of magnitude more tie-points than transformation coefficients, the RMSE values are indicative of the transformation fitting accuracy. The translation transformation (Equation (1)) is the least accurate with RMSE values greater than 0.5 pixels for both image pairs. The affine (Equation (2)) and second order polynomial (Equation (3)) transformations have small RMSEs ranging from 0.286 to 0.302 pixels, which indicate a high level of reliability of the tie-points obtained by the depth-first LSM and also that these transformations are sufficient to model the Landsat-to-Sentinel misregistration. The RMSEs are not much different for the affine and second order polynomial transformations. This suggests that the geometric complexity provided by the second order polynomial transformation will not provide a notable improvement to the registration accuracy.
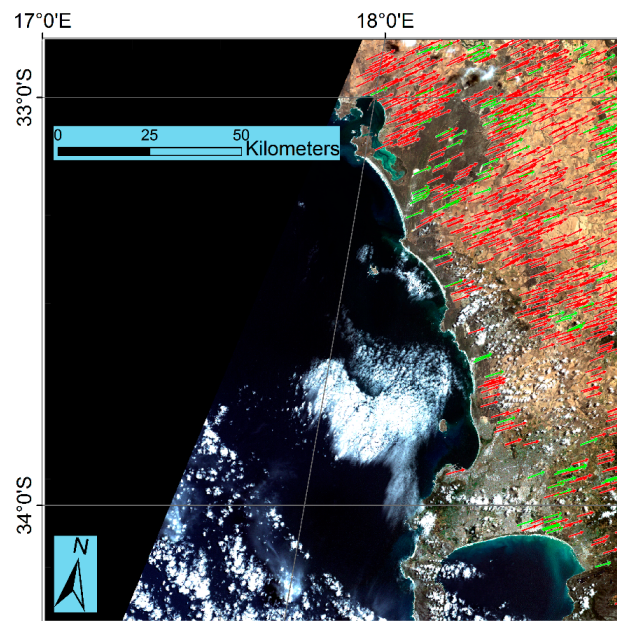
Figure 5. Illustration of the tie-points and misregistration vectors obtained from the Cape Town data (Figure 1). The 116 green vectors point from the Landsat-8 week 47 image tie-point locations to the corresponding Sentinel-2A week 49 tie point locations. The 797 red vectors point from the Landsat-8 week 47 image tie-point locations to the corresponding Sentinel-2A week 51 tie-point locations. The vector lengths are enlarged by 80 times for visual clarity. To provide geographic context, the background image shows the Landsat-8 week 47 30 m true color image.

Table 1. Summary statistics (units 10 m pixels) of the Cape Town (Figure 1) Landsat to Sentinel WELD tile x and y axis misregistration quantified by the tie-points as Equation (4).

	Landsat-8 Week 47 Image and Sentinel-2A Week 49 Image Tie-Point Differences (116 Tie-Points)		Landsat-8 Week 47 Image and Sentinel-2A Week 51 Image Tie-Point Differences (797 Tie-Points)	
	Δ_x	Δ_y	Δ_x	Δ_y
minimum	4.305	−2.707	4.270	−3.102
maximum	6.321	−1.484	6.326	−1.354
mean	5.445	−2.119	5.263	−2.132
standard deviation	0.402	0.301	0.400	0.334

Table 2. Cape Town transformation coefficients and RMSE (Equation (5)) derived by fitting the 116 tie-points extracted from the Landsat-8 week 47 image and Sentinel-2A week 49 image.

Translation Transformation (Equation (1))		Affine Transformation (Equation (2))		2nd Order Polynomial Transformation (Equation (3))			
a_0	−5.479373097	a_0	−7.167002618	a_0	−8.803776971	a_3	−0.000000007
		a_1	1.000156515	a_1	1.000413024	a_4	−0.000000011
		a_2	−0.000061958	a_2	−0.000010265	a_5	0.000000004
b_0	2.131139260	b_0	3.642703772	b_0	3.871263951	b_3	−0.000000000
		b_1	−0.000105759	b_1	−0.000129013	b_4	0.000000000
		b_2	0.999971229	b_2	0.999933409	b_5	0.000000001
RMSE: 0.504 pixels		RMSE: 0.286 pixels		RMSE: 0.296 pixels			

Table 3. Cape Town transformation coefficients and RMSE (Equation (5)) derived by fitting the 797 tie-points extracted from the Landsat-8 week 47 image and Sentinel-2A week 51 image.

Translation Transformation (Equation (1))		Affine Transformation (Equation (2))		2nd Order Polynomial Transformation (Equation (3))			
a_0	−5.264385177	a_0	−7.294159359	a_0	−8.803776971	a_3	−0.000000009
		a_1	1.000180379	a_1	1.000413024	a_4	−0.000000006
		a_2	−0.000069438	a_2	−0.000010265	a_5	0.000000002
b_0	2.130857209	b_0	3.678436109	b_0	3.871263951	b_3	0.000000001
		b_1	−0.000106223	b_1	−0.000129013	b_4	0.000000002
		b_2	0.999962709	b_2	0.999933409	b_5	0.000000001
RMSE: 0.536 pixels		RMSE: 0.302 pixels		RMSE: 0.301 pixels			

The transformation coefficients for the two image pairs are quite similar (Tables 2 and 3). To investigate this, the Euclidean distances (Equation (6)) between the projected WELD tile pixel locations projected using the Tables 2 and 3 coefficients were assessed for all the 15885 × 15885 10 m WELD tile pixels (Table 4) as described in Section 4.2. For the translation transformation, there is a constant 0.215 pixels difference because the translational differences are constant relative to each other across the WELD tile. The affine and second order polynomial transformations provide spatially variable Euclidean distances. Compared with the second order polynomial transformation, the affine transformation has smaller minimum, maximum, mean and standard deviation values of 0.000, 0.266, 0.114 and 0.055 pixels *versus* 0.002, 1.332, 0.245 and 0.257 pixels, respectively. Evidently, the second order polynomial transformation fitting is less stable than the affine transformation fitting for the image pairs and tie-points considered. This and the similar transformation fittings' RMSE values (Tables 2 and 3) suggest that the affine transformation is appropriate for registration of the sensor data.

Table 4. Summary statistics (units 10 m pixels) of the Euclidean distances (Equation (6)) between the projected WELD tile pixel locations projected using the same transformation type but the coefficients in Tables 2 and 3. Derived considering all 15885 × 15885 10 m WELD tile pixels.

	Translation Transformation (Equation (1))	Affine Transformation (Equation (2))	2nd Order Polynomial Transformation (Equation (3))
minimum	−0.215	0.000	0.002
maximum	−0.215	0.266	1.332
mean	−0.215	0.114	0.245
standard deviation	0	0.055	0.257

5.1.3. Qualitative Visual Registration Assessment

Figure 6 illustrates the result of the Landsat-8 to Sentinel-2A registration undertaken using the affine transformation (Table 3). Sensor misregistration effects are quite apparent in Figure 6a, especially in the WELD x-axis orientation, which is reflected by the relatively larger absolute mean Δ_x than Δ_y tie-point values (Table 1). After the registration (Figure 6b), the sensor misregistration effects are considerably less pronounced. Surface changes that occurred between the Landsat-8 and Sentinel-2A acquisition dates are apparent in the registered data notably, such as differences in the beach profile on the south-east side of the harbor prominentary, the presence of clouds, and perhaps ship tracks in one image and not the other.
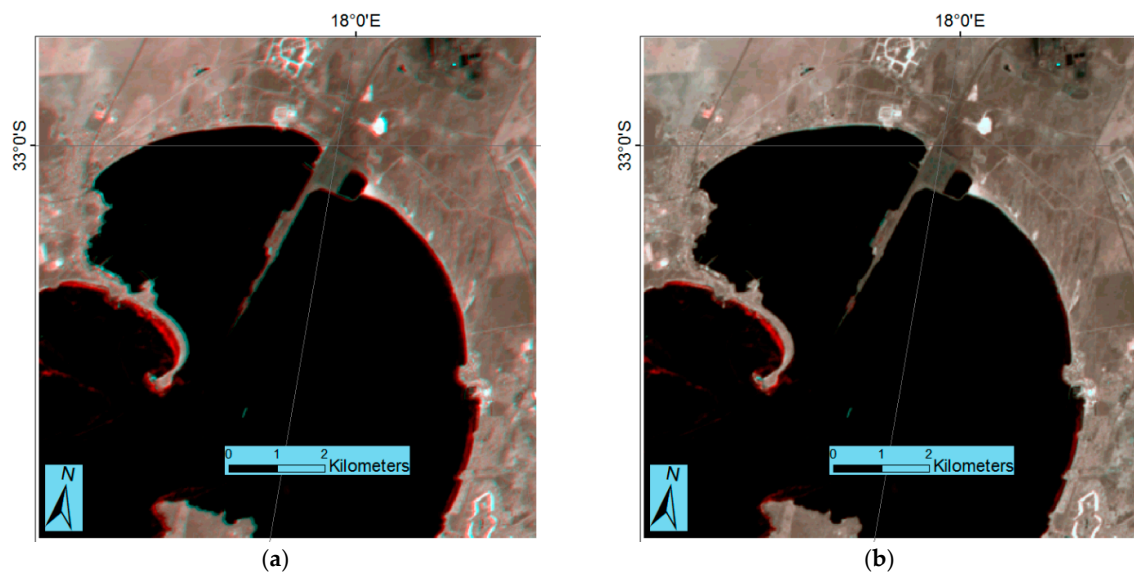
(**a**)   (**b**)

Figure 6. False color images illustrating Landsat-8 week 47 and Sentinel-2A week 51 images (**a**) before registration and (**b**) after registration. The Sentinel NIR data are shown as red and the Landsat NIR data are shown as blue and green. A 350 × 350 30 m pixel subset over Saldanha Bay, Cape Town (northern side of the WELD tile, Figure 1) is shown. The registration was undertaken using the affine transformation coefficients (Table 3).

5.1.4. Dense-Matching Prediction-Error Assessment

Figure 7 shows the dense-matching prediction-error maps for the translation, affine and second order polynomial transformations between the Landsat-8 week 47 and Sentinel-2A week 51 image pair. The areas with no matched points or satellite observations are shown in black. The black "holes" are predominantly due to water bodies and clouds in either image, and also due to surface changes that caused matches to be rejected. The prediction errors for the translation (Figure 7a) are greater than those for the affine (Figure 7b) and second order polynomial (Figure 7c) transformations. The affine and second order polynomial transformations prediction errors have similar magnitude and spatial pattern. Their greatest differences occur in the south-west and this is likely due to the lack of tie-points in this region (Figure 5). This and the stripes that are aligned approximately parallel to the Sentinel-2A and Landsat-8 track directions are discussed further in Section 6. The prediction-error maps obtained considering the other Cape Town image pair are not illustrated but were visually similar except for the spatial distribution of the "holes" due to image differences (cloud cover and surface changes).

Table 5 summarizes the dense-matching prediction errors illustrated in Figure 7. The mean and standard prediction errors are tabulated. The maximum error is the same as the successful match locational criteria (Section 4.4) and the minimum error was zero pixels (to decimal places). The prediction errors are greater for the translation transformation, which is expected due to the relatively less accurate translation model fits (Tables 2 and 3) and is reflected in the prediction error maps illustrated in Figure 7. The prediction errors are similar for the affine and second order polynomial transformations with means and standard deviations of less than 0.3 and 0.2 pixels, respectively. The errors for the Landsat-8 week 47 and Sentinel-2A week 49 image pair are higher and more variable than for the Landsat-8 week 47 and Sentinel-2A week 51 image pair, and this is likely due to the smaller number of tie-points used.
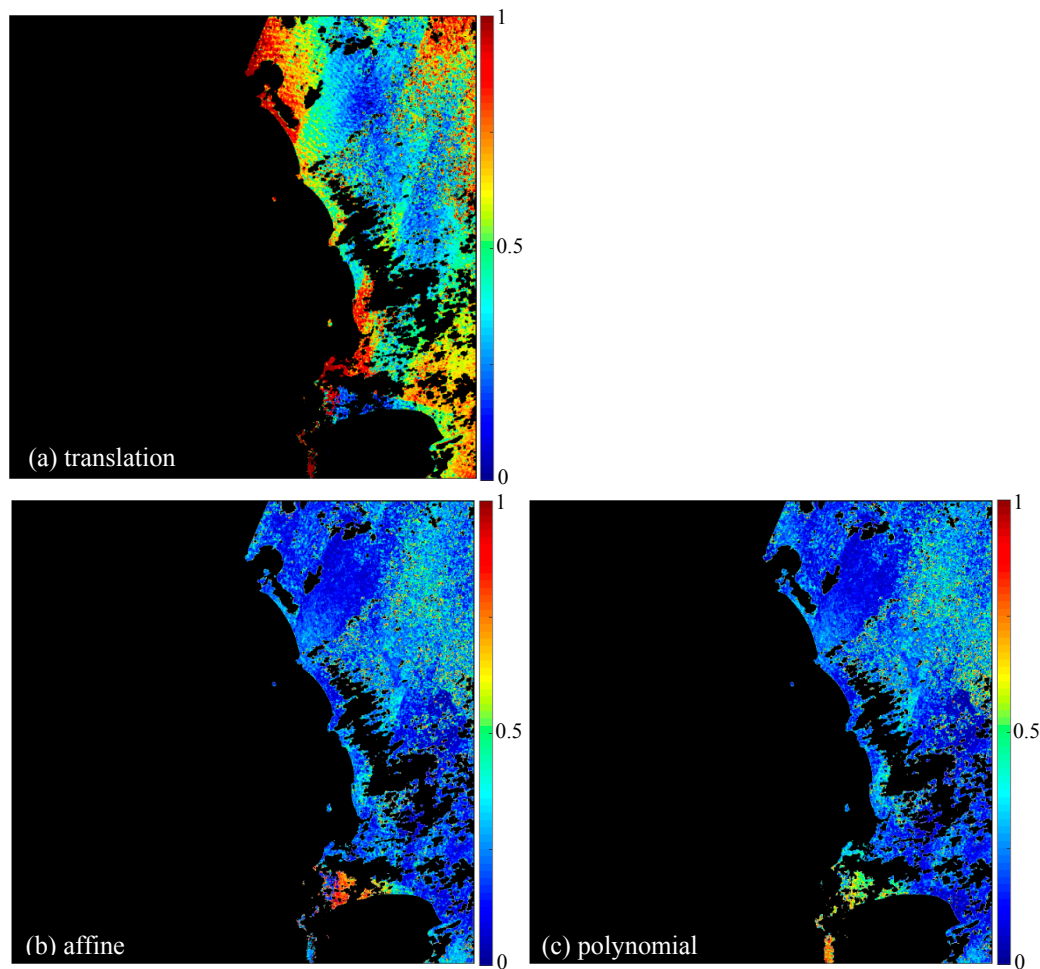
Figure 7. Dense-matching prediction-error maps for the translation (**a**); affine (**b**) and second order polynomial (**c**) transformations between the Cape Town Landsat-8 week 47 and Sentinel-2A week 51 image pair. Dense-matching grid points were sampled every six 10 m pixel across the 15885 × 15885 10 m WELD tile, generating 2647 × 2647 prediction-error maps (Equation (7)). Locations where there are no matches are shown as black.

Table 5. Summary statistics (units 10 m pixels) of the dense-matching prediction errors (Equation (7)) for the translation, affine and polynomial transformations obtained from the Cape Town image pairs, for Landsat-8 week 47 to Sentinel-2A weeks 49 and 51.

	Translation Transformation (Equation (1))		Affine Transformation (Equation (2))		2nd Order Polynomial Transformation (Equation (3))	
	47 —> 49	47 —> 51	47 —> 49	47 —> 51	47 —> 49	47 —> 51
mean	0.508	0.494	0.270	0.252	0.260	0.247
standard deviation	0.272	0.239	0.184	0.177	0.180	0.174

5.2. S.W. Limpopo

5.2.1. Tie-Point Misregistration Assessment

For the Limpopo study site, only one pair of images was considered (Figure 2). A total of 1840 feature points were detected from the Sentinel-2A 120 m week 49 image, and after the depth-first mismatch detection process, 180 tie-points were defined (Figure 8). The tie-points were most densely distributed in less cloudy agricultural and urban areas where the image intensity contrast was pronounced. The Landsat-8 image was misaligned in a south-east direction relative to the Sentinel-2A

image. Not only was the misalignment orientation different from the Cape Town image pairs, but also the magnitude of the tie-point differences was different (Table 6). The x-axis and y-axis mean tie-point shifts were smaller with magnitudes greater than 1.5 and 1.7 pixels, respectively, but with similar standard deviations (Table 6) to those found for Cape Town (Table 1).
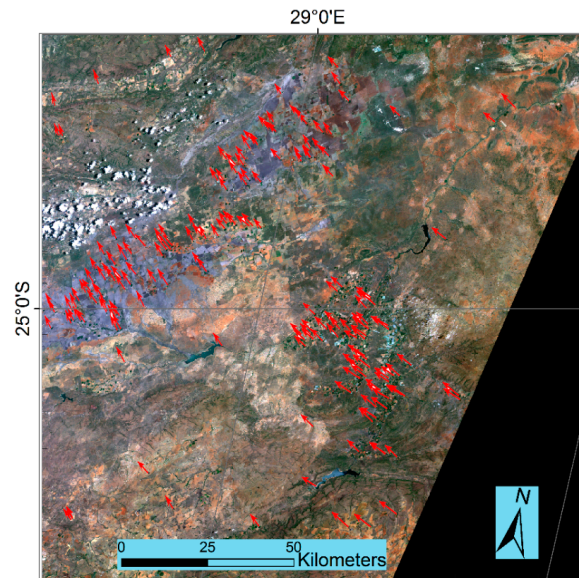

Figure 8. Illustration of the tie-points and misregistration vectors obtained from the Limpopo data (Figure 2). The vectors point from the Landsat-8 week 49 image tie-point locations to the Sentinel-2A week 49 (180 red vectors) tie-point locations. The vector lengths are enlarged by 200 times for visual clarity. To provide geographic context the background image shows the Landsat-8 week 49 30 m true color image.

Table 6. Summary statistics (units 10 m pixels) of the Limpopo (Figure 2) Landsat-8 week 49 and Sentinel-2A week 49 WELD tile x and y axis misregistration quantified by the tie-points as Equation (4).

	Δ_x	Δ_y
minimum	−2.551	−2.305
maximum	−0.512	−1.035
mean	1.498	−1.741
standard deviation	0.477	0.229

5.2.2. Transformation Coefficient Fitting Assessment

Table 7 summarizes the transformation coefficients and RMSE values. As for the Cape Town image pairs, the translation transformation is the least accurate (RMSE > 0.5 pixels) and the affine and second order polynomial transformations have similar RMSEs of 0.303 and 0.309 pixels, respectively, and so the affine transformation is deemed appropriate for registration of the sensor data.

Table 7. Limpopo transformation coefficients and RMSE (Equation (5)) derived by fitting the 180 tie-points extracted from the Landsat-8 week 49 image and Sentinel-2A week 49 image.

Translation Transformation (Equation 1)		Affine Transformation (Equation (2))		2nd Order Polynomial Transformation (Equation (3))			
a_0	1.487925133	a_0	0.408828226	a_0	0.611637495	a_3	0.000000004
		a_1	1.000101868	a_1	1.000011088	a_4	0.000000006
		a_2	0.000064412	a_2	0.000089111	a_5	−0.000000005
b_0	1.743876575	b_0	1.934822089	b_0	2.094640816	b_3	0.000000003
		b_1	−0.000036935	b_1	−0.000086720	b_4	0.000000002
		b_2	1.000004653	b_2	0.999997931	b_5	−0.000000001
RMSE: 0.534 pixels		RMSE: 0.303 pixels		RMSE: 0.309 pixels			

5.2.3. Dense-Matching Prediction-Error Assessment

As for the Limpopo image pair, the dense-matching prediction errors for the translation transformation (Figure 9a) are greater than those for the affine (Figure 9b) and second order polynomial (Figure 9c) transformations. The affine and second order polynomial transformations prediction errors have similar magnitude and spatial pattern. There are stripes aligned parallel to the sensor track directions, especially evident in Figure 9a, that are discussed further in Section 6. Similar to the Cape Town prediction error statistics (Table 5), the prediction errors are greater for the translation transformation, and are similar for the affine and second order polynomial transformations (Table 8) that have means and standard deviations of about 0.21 and 0.17 pixels, respectively.
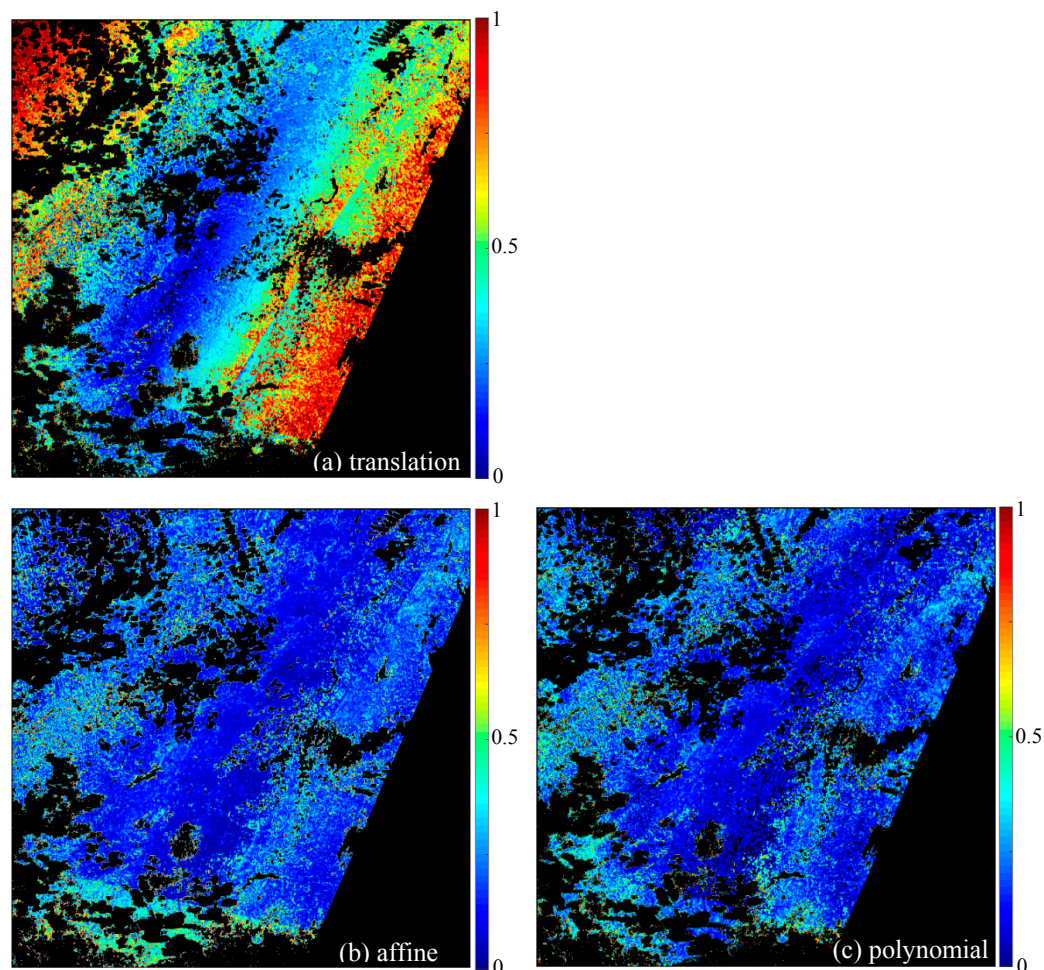

Figure 9. Dense-matching prediction-error maps for the translation (**a**); affine (**b**); and second order polynomial (**c**) transformations, between the Limpopo Landsat-8 week 49 and Sentinel-2A week 49 image pair.

Table 8. Summary statistics (units 10 m pixels) of the dense-matching prediction error (Equation (7)) for the translation, affine and polynomial transformations obtained from the Limpopo image pair.

	Translation Transformation (Equation (1))	Affine Transformation (Equation (2))	2nd Order Polynomial Transformation (Equation (3))
mean	0.467	0.215	0.210
standard deviation	0.274	0.173	0.170

6. Discussion

Registration of satellite data to sub-pixel precision is a pre-requisite for meaningful data comparison and surface change detection [54,55]. The registration tie-points were derived by least squares matching, which provides sub-pixel location precision and handles the sensor spectral band differences and the non-linear geometric distortions present between the Landsat-8 and Sentinel-2A data. The tie-points were used to fit translation, affine and second order polynomial transformation functions, and for the three pairs of sensor data (different dates and locations), the translation transformation was less accurate with RMSE fit values greater than 0.5 pixels and greater dense-matching prediction errors. The affine and second order polynomial transformations had RMSE fit values of approximately 0.3 pixels and dense-matching prediction errors of similar magnitude. However, the tie-points derived from the two Sentinel-2A Cape Town images and the same Landsat-8 Cape Town image (Figure 5) provided less stable transformations relative to each other when the second order polynomial rather than the affine transformation was used. This is likely because of the greater sensitivity of the polynomial model to the different numbers and spatial distributions of tie-points [18,56] (Figure 5). These results suggest that an affine transformation is sufficient to register Landsat-8 L1T and Sentinel-2A L1C data when there are considerably more tie-points than transformation coefficients.

The causes of the observed sensor misregistration are complex. For both Cape Town image pairs, the Landsat-8 image was misaligned in a similar south-west direction relative to the Sentinel-2A images, and for the Limpopo image pair, the Landsat-8 misalignment was in a relative south-east direction. The relatively constant geographical pattern of tie-point-characterized sensor misregistration (Figures 5 and 6) supports the hypothesis that the misregistration is due primarily to Landsat GLS path/row specific locational errors [9,15]. However, other sources of error may be present, including, for example, geometric relief distortion imposed by digital elevation model inaccuracies, although we observed no elevation related shifts, and inadequate knowledge and/or modeling of the sensor interior and exterior orientation. The detailed study on these error sources is beyond the scope of this paper.

The dense-matching prediction-error maps (Figures 7 and 9) exhibited stripes that are aligned approximately parallel to the Sentinel-2A and Landsat-8 track directions. To investigate this further, maps of the x and y axis shifts used to compute the dense-matching prediction errors (Equation (7)) were generated. The results revealed similar patterns for all the study data. Figure 10a,b show the dense-matching x and y shifts for the Limpopo image pair. The mean *x*-shift and *y*-shift values were −1.592 and −1.674 pixels, respectively, which is a similar magnitude to the fitted translation transformation coefficients (Table 7). However, the shifts were unevenly distributed across the tile, with *x*-shift and *y*-shift standard deviations of 0.473 and 0.237 pixels, respectively. This illustrates why the translation transformation had larger errors than the other transformation types. Figure 10c shows the Sentinel-2A L1C tile and detector boundaries and Figure 10d shows the Landsat-8 L1T image boundaries. The dense-matching x and y axis shift maps have apparent zones with edges aligned approximately parallel to the detector and image boundaries. The x and y shifts across the zone boundaries are quite small however, usually less than 0.35 pixels (manually measured). Their cause is likely due to a combination of factors. Small geometric misalignments and/or radiometric calibration differences between the detector banks combined with directional reflectance affects may result in least squares matching differences between the sensor data that will be pronounced along the detector bank boundaries. Along scan directional reflectance variations of several percent are present in 15° field of view Landsat data [57] and are expected to be greater in wider field of view (20.5°) Sentinel-2 data. Consequently, different sensor viewing geometry and forward and backward scatter sensing conditions may introduce along scan reflectance variations that cause small least squares matching shifts. However, we note that for the data considered, the solar zenith angles were only a few degrees different and the images were sensed in the same scattering direction. The along track shifts evident at

the bottom left of Figure 10 are coincident with the Landsat-8 L1T boundary, evident in Figure 10d, and may be caused by a different set of Landsat ground controls being used.
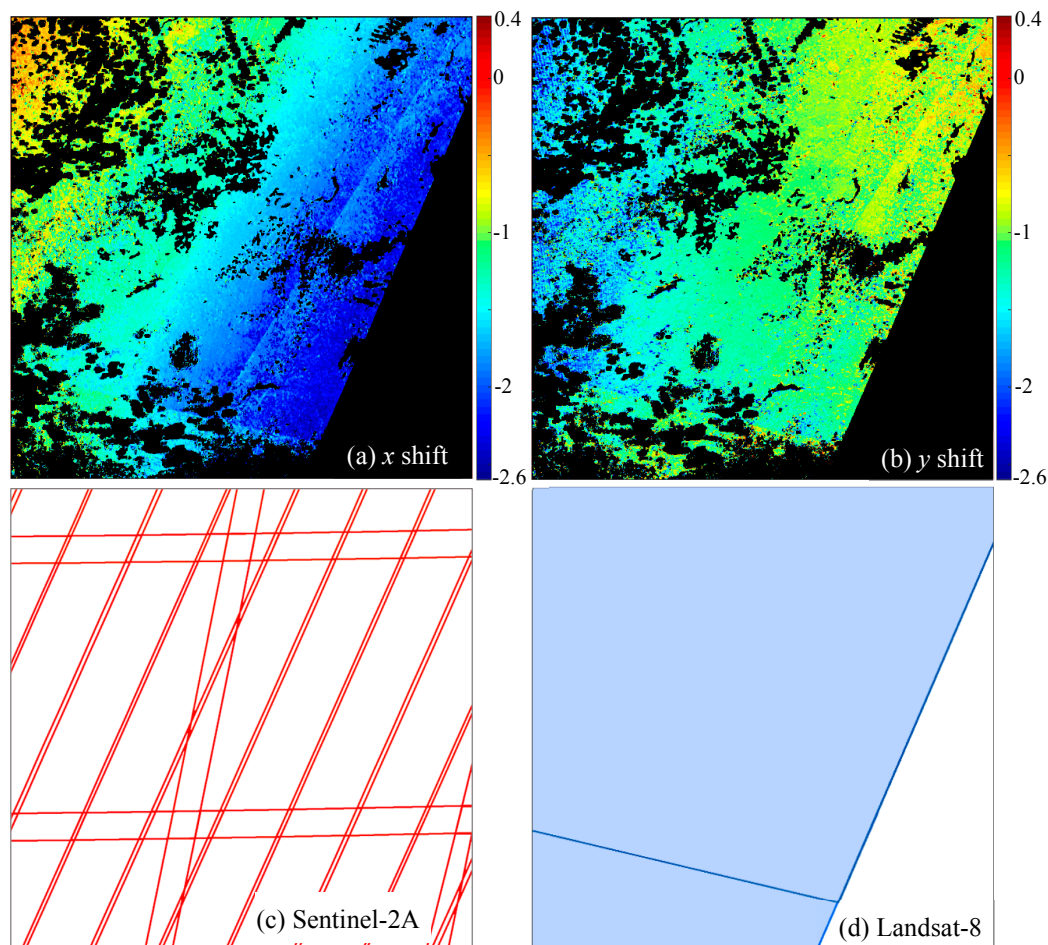

Figure 10. Dense-matching maps x and y axis shifts (units 10 m pixels) between the Limpopo Landsat-8 week 49 and Sentinel-2A week 49 image pair. The affine transformation was used to guide the dense matching (Section 4.4). The (**a**) x-shift $x_{S,i} - x_{L,i}$ and (**b**) y-shift $y_{L,i} - y_{L,i}$ are shown, where $(x_{S,i}, y_{S,i})$ is the Sentinel grid-point location and $(x_{L,i}, y_{L,i})$ is the corresponding least squares matched Landsat location for grid-point i. Locations where there are no matches are shown as black. Note that $(x_{L,i}, y_{L,i})$ is theoretically independent on the transformation type and so the translation and polynomial-based shift maps, which were very similar to the affine-based shift maps, are not shown; (**c**) shows the Sentinel-2A L1C tile and detector boundaries (red) and (**d**) shows the Landsat-8 L1T image boundaries (blue).

Finally, as with all registration methods, the accuracy will be dependent upon the availability of tie-points which will be limited in regions of unstructured terrain and where there are clouds and shadows. However, if the sensor misregistration is predominantly constant for each Landsat WRS path/row location, *i.e.*, due primarily to GLS path/row locational errors [9,15], then tie-points extracted from matching many pairs of approximately contemporaneous Landsat-8 and Sentinel-2A data through time may provide a reliable set of tie-points, from either a "good" pair or a combination of multiple pairs. This will require further assessment on sensor's multi-temporal intra-misregistration.

7. Conclusions

This study presented an approach for the automated registration of geolocated Landsat-8 OLI L1T and Sentinel-2A MSI L1C data. Registration errors between South African test data greater than the native sensor pixel resolutions were found and were effectively characterized by an affine transformation. Tie-points extracted from pairs of Landsat-8 L1T and Sentinel-2A L1C images revealed sensor misregistration greater than one 30 m Landsat-8 pixel dimension (for two Cape Town image pairs) and greater than one 10 m Sentinel-2A pixel dimension (for Limpopo image pair). This degree of registration error is not insignificant, and will limit the ability to meaningfully compare Landast-8 and Sentinel-2A data even at the 30 m Landsat-8 pixel resolution. As the degree of sensor misregistration is unknown at other global locations and times, further study considering a global distribution of Landsat-8 and Sentinel-2A contemporaneous data and considering the causes of the sensor misregistration is recommended.

The developed registration approach is computationally efficient because it implements feature point detection at reduced spatial resolution and then area-based least squares matching around the feature points with mismatch detection across four image pyramid levels to identify a sparse set of tie-points. Nevertheless, the registration of large spatial and temporal coverage of Landsat-8 and Sentinel-2A data will still be computationally challenging because of the considerable satellite data volume imposed by the 10 m Sentinel-2A bands. The performance of the approach on the selected test data, which contained clouds, shadows, land cover changes, and sensor acquisition date differences of up to 26 days, provide confidence in its robustness. The approach provided hundreds of automatically located tie-points that had least-squares fits with RMSE of 0.286, 0.302 and 0.303 10 m pixels under affine transformation for the three registered Landsat-8 and Sentinel-2A image pairs. Dense-matching prediction-error assessment considering every sixth pixel on a systematic grid revealed sub-pixel prediction errors (means < 0.3 pixels and standard deviations < 0.2 pixels at 10 m pixel resolution) for the fitted affine transformations. These results and visual assessment of the affine transformed data indicate the sub-pixel registration performance required for meaningful sensor data comparison and time series applications [54,55] including those illustrated in this special issue [58].

Author Contributions: L.Y. developed the algorithm and undertook the data analysis, and processed the data and developed the graphics with assistance from H.K., J.L. and H.H.; D.R. helped with the data analysis and structured and drafted the manuscript with assistance from L.Y.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R.; *et al.* Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172. [CrossRef]

2. Irons, J.R.; Dwyer, J.L.; Barsi, J.A. The next Landsat satellite: The Landsat data continuity mission. *Remote Sens. Environ.* **2012**, *122*, 11–21. [CrossRef]

3. Drusch, M.; Del Bello, U.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P.; *et al.* Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [CrossRef]

4. European Space Agency (ESA). *Sentinel-2 User Handbook, Issue 1, Rev 2, Revision 2*; ESA Standard Document; ESA: Paris, France, 2015.

5. European Space Agency (ESA). *Sentinel-2 Products Specification Document, Issue 13.1*; ESA REF: S2-PDGS-TAS-DI-PSD; ESA: Paris, France, 2015.

6. Arvidson, T.; Gasch, J.; Goward, S.N. Landsat 7's long-term acquisition plan—An innovative approach to building a global imagery archive. *Remote Sens. Environ.* **2001**, *78*, 13–26. [CrossRef]

7.   Loveland, T.R.; Dwyer, J.L. Landsat: Building a strong future. *Remote Sens. Environ.* **2012**, *122*, 22–29. [CrossRef]

8.   Languille, F.; Déchoz, C.; Gaudel, A.; Greslou, D.; de Lussy, F.; Trémas, T.; Poulain, V. Sentinel-2 geometric image quality commissioning: First results. *Proc. SPIE* **2015**. [CrossRef]

9.   Storey, J.; Choate, M.; Lee, K. Landsat-8 Operational Land Imager on-orbit geometric calibration and performance. *Remote Sens.* **2014**, *6*, 11127–11152. [CrossRef]

10. Storey, J. Sentinel-2 On-orbit geometric analysis and harmonization plans. In Proceedings of the Landsat Science Team Meeting, Virginia Tech, Blacksburg, VA, USA, 12–14 January 2016.

11. Tucker, C.J.; Grant, D.M.; Dykstra, J.D. NASA's global orthorectified Landsat data set. *Photogramm. Eng. Remote Sens.* **2004**, *70*, 313–322. [CrossRef]

12. Gutman, G.; Huang, C.; Chander, G.; Noojipady, P.; Masek, J.G. Assessment of the NASA–USGS global land survey (GLS) datasets. *Remote Sens. Environ.* **2013**, *134*, 249–265. [CrossRef]

13. Briottet, X.; Lier, P.; Valorge, C. *Satellite Imagery, From Acquisition Principles to Processing of Optical Images for Observing the Sarth*; Cepadues Editions: Toulouse, France, 2012.

14. Dechoz, C.; Poulain, V.; Massera, S.; Languille, F.; Greslou, D.; de Lussy, F.; Gaudel, A.; L'Helguen, C.; Picard, C.; Trémas, T. Sentinel-2 global reference image. *Proc. SPIE* **2015**. [CrossRef]

15. Storey, J.C.; Roy, D.P.; Masek, J.G.; Dwyer, J.L.; Gascon, F.; Choate, M.J. A note on the temporary mis-registration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) imagery. *Remote Sens. Environ.* Submitted.

16. Zitova, B.; Flusser, J. Image registration methods: A survey. *Image Vis. Comput.* **2003**, *21*, 977–1000. [CrossRef]

17. Long, T.; Jiao, W.; He, G.; Zhang, Z. A fast and reliable matching method for automated georeferencing of remotely-sensed imagery. *Remote Sens.* **2016**, *8*, 56. [CrossRef]

18. Goshtasby, A. Piecewise cubic mapping functions for image registration. *Pattern Recognit.* **1987**, *20*, 525–533. [CrossRef]

19. Roy, D.P.; Devereux, B.; Grainger, B.; White, S. Parametric geometric correction of airborne thematic mapper imagery. *Int. J. Remote Sens.* **1997**, *18*, 1865–1887. [CrossRef]

20. Liu, D.; Gong, P.; Kelly, M.; Guo, Q. Automatic registration of airborne images with complex local distortion. *Photogramm. Eng. Remote Sens.* **2006**, *72*, 1049–1059. [CrossRef]

21. Li, R.; Hwangbo, J.; Chen, Y.; Di, K. Rigorous photogrammetric processing of HiRISE stereo imagery for Mars topographic mapping. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2558–2572.

22. Ustin, S.L.; Roberts, D.A.; Pinzon, J.; Jacquemoud, S.; Gardner, M.; Scheer, G.; Castaneda, C.M.; Palacios-Orueta, A. Estimating canopy water content of chaparral shrubs using optical methods. *Remote Sens. Environ.* **1998**, *65*, 280–291. [CrossRef]

23. Roy, D.P.; Qin, Y.; Kovalskyy, V.; Vermote, E.F.; Ju, J.; Egorov, A.; Hansen, M.C.; Kommareddy, I.; Yan, L. Conterminous United States demonstration and characterization of MODIS-based Landsat ETM+ atmospheric correction. *Remote Sens. Environ.* **2014**, *140*, 433–449. [CrossRef]

24. Gao, F.; Masek, J.G.; Wolfe, R.F. An automated registration and orthorectification package for Landsat and Landsat-like data processing. *J. Appl. Remote Sens.* **2009**, *3*, 033515.

25. Devaraj, C.; Shah, C.A. Automated geometric correction of multispectral images from High Resolution CCD Camera (HRCC) on-board CBERS-2 and CBERS-2B. *ISPRS J. Photogramm. Remote Sens.* **2014**, *89*, 13–24. [CrossRef]

26. Zhang, H.K.; Roy, D.P. Computationally inexpensive Landsat-8 Operational Land Imager (OLI) pansharpening. *Remote Sens.* **2016**, *8*, 180. [CrossRef]

27. Snyder, J.P. *Flattening the Earth: Two Thousand Years of Map Projections*; The University of Chicago Press: Chicago, IL, USA; London, UK, 1993.

28. Roy, D.P.; Ju, J.; Kline, K.; Scaramuzza, P.L.; Kovalskyy, V.; Hansen, M.; Loveland, T.R.; Vermote, E.; Zhang, C. Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sens. Environ.* **2010**, *114*, 35–49. [CrossRef]

29. Web-Enabled Landsat Data (WELD). Available online: http://globalweld.cr.usgs.gov/collections (accessed on 6 June 2016).

30. Wolfe, R.; Nishihama, M.; Fleig, A.; Kuyper, J.; Roy, D.; Storey, J.; Patt, F. Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sens. Environ.* **2002**, *83*, 31–49. [CrossRef]

31. Konecny, G. Mathematical models and procedures for the geometric registration of remote sensing imagery. *Int. Arch. Photogramm. Remote Sens.* **1976**, *21*, 1–33.

32. Roy, D.P.; Borak, J.; Devadiga, S.; Wolfe, R.; Zheng, M.; Descloitres, J. The MODIS land product quality assessment approach. *Remote Sens. Environ.* **2002**, *83*, 62–76. [CrossRef]

33. Shlien, S. Geometric correction, registration and resampling of Landsat imagery. *Can. J. Remote Sens.* **1979**, *5*, 75–89. [CrossRef]

34. Thevenaz, P.; Ruttimann, U.E.; Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **1998**, *7*, 27–41. [CrossRef] [PubMed]

35. Adelson, E.H.; Anderson, C.H.; Bergen, J.R.; Burt, P.J.; Ogden, J.M. Pyramid methods in image processing. *RCA Eng.* **1984**, *29*, 33–41.

36. Förstner, W.; Guelch, E. A Fast operator for detection and precise location of distinct points, corners and centers of circular features. In Proceedings of the ISPRS Intercommission Workshop on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, 2–4 June 1987; pp. 281–305.

37. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]

38. Di, K.; Xu, F.; Wang, J.; Agarwal, S.; Brodyagina, E.; Li, R.; Matthies, L. Photogrammetric processing of rover imagery of the 2003 Mars Exploration Rover mission. *ISPRS J. Photogramm. Remote Sens.* **2003**, *63*, 181–201. [CrossRef]

39. Di, K.; Liu, Z.; Yue, Z. Mars rover localization based on feature matching between ground and orbital imagery. *Photogram. Eng. Remote Sens.* **2011**, *77*, 781–791. [CrossRef]

40. Wu, B.; Guo, J.; Hu, H.; Li, Z.; Chen, Y. Co-registration of lunar topographic models derived from Chang'E-1, SELENE, and LRO laser altimeter data based on a novel surface matching method. *Earth Planet. Sci. Lett.* **2013**, *364*, 68–84. [CrossRef]

41. Pratt, W.K. *Digital Image Processing*, 2nd ed.; Wiley: New York, NY, USA, 1991.

42. Lewis, J.P. Fast normalized cross-correlation. *Vis. Interface* **1995**, *10*, 120–123.

43. Förstner, W. On the geometric precision of digital correlation. *Int. Arch. Photogramm. Remote Sens.* **1982**, *24*, 176–189.

44. Gruen, A. Adaptive least squares correlation: A powerful image matching technique. *S. Afr. J. Photogramm. Remote Sens. Cartogr.* **1985**, *14*, 175–187.

45. Gruen, A. Development and status of image matching in photogrammetry. *Photogram. Rec.* **2012**, *27*, 36–57. [CrossRef]

46. Remondino, F.; El-Hakim, S.; Gruen, A.; Zhang, L. Turning images into 3-D models. *IEEE Signal. Process. Mag.* **2008**, *25*, 55–65. [CrossRef]

47. Kruse, F.A.; Lefkoff, A.B.; Boardman, J.W.; Heidebrecht, K.B.; Shapiro, A.T.; Barloon, P.J.; Goetz, A.F.H. The spectral image processing system (SIPS)—Interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.* **1993**, *44*, 145–163. [CrossRef]

48. Keshava, N. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1552–1565. [CrossRef]

49. Van der Meer, F. The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. *Int. J. Appl. Earth Obs. Geoinform.* **2006**, *8*, 3–17. [CrossRef]

50. Yan, L.; Niu, X. Spectral-angle-based Laplacian eigenmaps for nonlinear dimensionality reduction of hyperspectral imagery. *Photogramm. Eng. Remote Sens.* **2014**, *80*, 849–861. [CrossRef]

51. Yan, L.; Roy, D.P. Improved time series land cover classification by missing-observation-adaptive nonlinear dimensionality reduction. *Remote Sens. Environ.* **2015**, *158*, 478–491. [CrossRef]

52. Witkin, A.P. Scale-space filtering: A new approach to multi-scale description. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'84), San Diego, CA, USA, 19–21 March 1984; Volume 9, pp. 150–153.

53. Toutin, T. Review article: Geometric processing of remote sensing images: Models, algorithms and methods. *Int. J. Remote Sens.* **2004**, *25*, 1893–1924. [CrossRef]

54. Townshend, J.R.; Justice, C.O.; Gurney, C.; McManus, J. The impact of misregistration on change detection. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 1054–1060. [CrossRef]

55. Roy, D.P. The impact of misregistration upon composited wide field of view satellite data and implications for change detection. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 2017–2032. [CrossRef]

56. De Leeuw, A.J.; Veugen, L.M.M.; Van Stokkom, H.T.C. Geometric correction of remotely-sensed imagery using ground control points and orthogonal polynomials. *Int. J. Remote Sens.* **1988**, *9*, 1751–1759. [CrossRef]

57. Roy, D.P.; Zhang, H.K.; Ju, J.; Gomez-Dans, J.L.; Lewis, P.E.; Schaaf, C.B.; Sun, Q.; Li, J.; Huang, H.; Kovalskyy, V. A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Remote Sens. Environ.* **2016**, *176*, 255–271. [CrossRef]

58. Atzberger, C. Remote Sensing Special Issue: First Experiences with European Sentinel-2 Multi-Spectral Imager (MSI). 2016. Available online: http://www.mdpi.com/journal/remotesensing/special_issues/sentinel-2_msi (accessed on 6 June 2016).