

DEM orientation based on local feature correspondence with global DEMs

Amin Sedaghat & Amin Alizadeh Naeini

To cite this article: Amin Sedaghat & Amin Alizadeh Naeini (2018) DEM orientation based on local feature correspondence with global DEMs, GIScience & Remote Sensing, 55:1, 110-129, DOI: [10.1080/15481603.2017.1364879](https://doi.org/10.1080/15481603.2017.1364879)

To link to this article: <https://doi.org/10.1080/15481603.2017.1364879>



Published online: 16 Aug 2017.



Submit your article to this journal [↗](#)



Article views: 25



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)



DEM orientation based on local feature correspondence with global DEMs

Amin Sedaghat ^{*a} and Amin Alizadeh Naeini^b

^aDepartment of Geomatics Engineering, Faculty of Civil Engineering, University of Tabriz, Tabriz, Iran; ^bDepartment of Geomatics Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan, Isfahan, Iran

(Received 9 March 2017; accepted 4 August 2017)

Thanks to rational polynomial coefficients (RPCs), which are provided by vendors to end users, digital elevation models (DEMs) can be simply derived from satellite stereo images. However, DEMs are influenced by systematic errors in the rational function model (RFM), known as RPC biases. Global DEMs (GDEMs), such as the Shuttle Radar Topography Mission (SRTM), which is the most inexpensive solution, can be applied to improve the accuracy of the relative RFM-derived DEMs. In this article, an automatic and robust local feature-based DEM matching and orientation approach is proposed in order to improve the accuracy of the relative RFM-derived DEMs without the use of ground control points (GCPs). The proposed approach consists of four main steps: (1) combined local feature extraction; (2) computation of the distinctive order-based self-similarity (DOBSS) descriptor; (3) a feature correspondence and local consistency checking process; and (4) a relative RFM-derived DEM orientation process using three-dimensional (3D) transformation models, including 3D rigid, 3D similarity and 3D affine transformations. This technique can avoid the sensitivity of conventional 3D DEM matching methods to initial values, monotonous areas and local distortions. Experimental results on two CARTOSAT-1 derived DEMs demonstrate the superior performance of the proposed DEM matching method over state-of-the-art methods, including SIFT, DAISY, LIOP, LBP, and BRISK descriptors, in terms of the number of correct matches (NCM) and DEM orientation accuracy. The results also show that the proposed method is able to significantly improve the geometric accuracy of the relative RFM-derived DEMs.

Keywords: DEM matching; local features; RPCs refinement; UR-SIFT; DOBSS

1. Introduction

A digital elevation model (DEM) is a raster of elevation values, which consists of an array of points of various elevations, sampled systematically at equally spaced intervals (Wright et al. 2006). High-resolution satellite stereo images represent one of the most important resources in DEM generation. For this purpose, a geometric relationship between the two-dimensional (2D) image space and the 3D ground space must be defined. The RFM is one of the most efficient models in this regard. The RFM typically contains the necessary information about the sensor, satellite platform, position and attitude. This model is actually a ratio of two third-order polynomials, which, in the direct form, relates each position in the image scene (line and sample) to its corresponding 3D ground coordinates (longitude, latitude, and height) (Fraser, Dial, and Grodecki 2006).

*Corresponding author. Email: a.sedaghat@tabrizu.ac.ir

Today, RFM coefficients, known as RPCs, are provided by most vendors to end users. These coefficients are calculated using the auxiliary information provided by on-board GPS receivers, inertial navigation systems and star trackers (Fraser, Dial, and Grodecki 2006). However, due to errors from various sources, such as GPS/INS drift, these RPCs and the DEMs generated from them (RFM-derived DEMs) are not sufficiently accurate (d'Angelo et al. 2008; Fraser and Hanley 2003; Kim and Jeong 2011).

To address this issue, it is necessary to modify RPCs using several GCPs (Aguilar, del Mar Saldana, and Aguilar 2013; Jannati and Valadan Zoej 2015; Li, Liu, and Deng 2016). However, this modification involves several essential problems. The first is that the collection of GCPs may be impossible or very difficult in some areas due to inaccessibility or danger. The second is related to the cost and time required for the collection of GCPs, while the third concerns the accuracy of these GCPs, which directly affects the reliability of the modified RPCs. Finally, finding the corresponding points of GCPs in satellite images is not always straightforward or accurate.

Common solutions to the problem of RPC refinement, without the use of GCPs, involve either orbital information (Toutin, Schmitt, and Wang 2012) or existing geospatial data (Oh and Lee 2015). However, since satellite orbital information is not usually available, existing geospatial data are generally applied. According to the literature, several types of geospatial data, including large-scale digital maps (Oh and Lee 2015), ortho-images (Oh, Lee, and Seo 2013), LIDAR (Teo and Huang 2013), and global DEMs (Kim and Jeong 2011), have been used for this purpose. However, since global DEMs, such as SRTM, are freely available worldwide, these forms of data are used in the current study.

To compensate for RPC biases in the ground space through global DEMs, a matching method is applied between the RFM-derived DEM and a global DEM. To do this, Gonçalves (2006) used a normalized cross-correlation (NCC) between the two above-mentioned DEMs. Ravanbakhsh and Fraser (2012, 2013) used mutual information-based matching and compared their results to the iterative closest point (ICP) (Besl and McKay 1992). Kim and Jeong (2011) and d'Angelo and Reinartz (2012) adopted the 3D least squares matching technique (Akca 2010), based on the 3D similarity and 3D affine transformations, respectively, to solve the absolute orientation of push-broom satellite images.

In addition to these methods for RPC biases compensation, several other DEM matching methods exist (Ravanbakhsh and Fraser 2013). Milledge, Lane, and Warburton (2009) proposed an optimization method to increase the performance of stereo-matching algorithms for digital topographic determination using existing DEM data. Streutker, Glenn, and Shrestha (2011) proposed a method for matching the elevation surfaces based on the comparison and statistical analysis of the local slope versus local elevation differences. Gil et al. (2014) proposed a DEM shading method for the correction of pseudoscopic effects on multi-platform satellite imagery.

The existing DEM matching methods mentioned above involve area-based matching (ABM) approaches, such as the NCC matching method, least squares matching and mutual information. For DEM pairs with small amount of local distortions, the ABM methods can generally provide high positional accuracies, but require good initial values. Accordingly, the input DEM pairs in area-based matching should initially be georeferenced in the same geodetic datum and are generally assumed to be rotationally aligned (Ravanbakhsh and Fraser 2013). Furthermore, the ABM methods suffer from monotonous textures, occlusions, local distortions and illumination differences in image matching. More importantly, this problem is intensified in DEM matching, where there are intrinsic errors and low elevation variations.

Feature-based matching (FBM) methods, on the other hand, extract salient features from the raster data pairs and establish the correspondence between them, using various similarity measures or local descriptors. FBM methods do not require initial values and are generally faster and more reliable than ABM methods (Sedaghat, Ebadi, and Mokhtarzade 2012). Recently, local invariant feature-based matching has received a great deal of attention from the remote sensing community (Castillo-Carrión and Guerrero-Ginel 2017; Sedaghat and Ebadi 2015a, 2015b; Yu et al. 2013).

The local features generally consist of two main parts: a feature detector and a feature descriptor. A feature detector, such as the Harris (Harris and Stephens 1988), scale-invariant feature transform (SIFT) (Lowe 2004), speeded-up robust features (SURF), or maximally stable extremal regions (MSER) (Matas et al. 2004) detector, extracts distinct features from data in two raster frames. A descriptor provides robust representations of the extracted features in order to characterize and match them. The most well-known descriptors are shape context (Belongie, Malik, and Puzicha 2002), SIFT (Lowe 2004), local self-similarity (LSS) (Shechtman and Irani 2007) and their extensions, such as adaptive binning SIFT (*AB-SIFT*) (Sedaghat and Ebadi 2015c) and distinctive order-based self-similarity (DOBSS) (Sedaghat and Ebadi 2015b).

In summary, Table 1 represents the characteristics of area-based and feature-based matching methods.

The main goal of this article is to develop a feature-based DEM matching method for reliable DEM orientation. The main contribution of the proposed method is a combined local feature-based matching method adopted for raster DEM data. The method proposed here uses the Harris and SIFT detectors, a highly discriminative DOBSS descriptor and a local consistency check blunder detection process to extract well-distributed conjugate points between the DEM pairs. Following feature correspondence, the matched features

Table 1. Characteristics of area-based and feature based matching methods.

Category	Typical algorithms		Main characteristics
Area-based method (ABM)		- Sum of Absolute Differences (SAD) - Normalized cross-correlation (NCC) - Mutual Information (MI) - Least squares matching (LSM)	- Requirement of approximate values - High computational complexity - Sensitive to monotonous textures, occlusions, local distortions, and illumination differences - Large data volume must be handled
Feature-based method (FBM)	Detector	- Harris - SIFT - MSER	- High stability against radiometric and geometric differences - More flexible with respect to surface discontinuities - The accuracy of FBM method is limited by the accuracy of the feature extraction process - Requirement of mismatched elimination process - Relatively fast - Less sensitive to image noise
	Descriptor	- Shape Context - SIFT - LSS	

detected are effectively applied in order to orient the RFM-derived DEM accurately using three types of transformation models: namely, the 3D rigid, 3D similarity, and 3D affine transformations. In this way, an inaccurate RFM-derived DEM can also be modified. The steps of the proposed approach are discussed in the next section. This is followed by an analysis of the experimental results, with conclusions drawn in the final section.

2. Methodology

The details of the proposed DEM matching method are described in this section. The input DEM pairs are a relative RFM-derived DEM generated from satellite image pairs and a global DEM. In this study, the SRTM is used as global DEM.

Figure 1 indicates the flowchart of the proposed method. In the first step, the local features are extracted from both the SRTM and the RFM-derived DEM. For this purpose, the proposed algorithm uses an integrated uniform local feature extraction approach

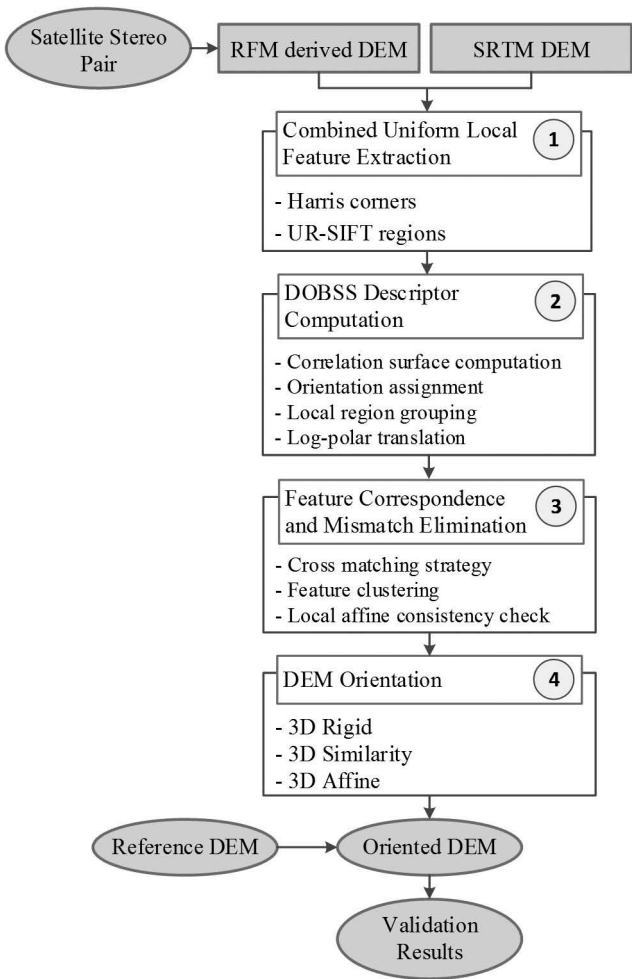


Figure 1. Flowchart of the proposed DEM matching method.

(Sedaghat and Ebadi 2015d), based on Harris corners (Harris and Stephens 1988) and uniform robust SIFT (UR-SIFT) regions (Sedaghat, Mokhtarzade, and Ebadi 2011) in order to increase the robustness and distribution of the DEM matching results. In the second step, the distinct and robust DOBSS descriptor (Sedaghat and Ebadi 2015b) is computed for each extracted feature using local self-similarity and correlation surface order-based grouping.

The third step consists of a local consistency check strategy for outlier rejection, based on k-means clustering and an affine transformation model (Sedaghat and Ebadi 2015d). Finally, to orientate the RFM-derived DEM, the geometric relationship between the two DEMs is defined using an appropriate transformation function. To investigate the impact of the transformation function on the accuracy of DEM orientation, the capability of three well-known 3D models (3D similarity, 3D rigid and 3D affine transformations) is also compared.

2.1. Combined local feature extraction

The proposed method begins with local feature extraction from two input SRTM and RFM-derived DEMs. Two DEMs are derived using different acquisition technologies (Shuttle Radar Topography Mission and CARTOSAT-1 satellite stereo images). These have different acquisition dates, resolution and accuracy. In addition, each DEM contains intrinsic errors due to the primary data acquisition technology and processing methodology used, in relation to the particular terrain and land cover type (Papasaika and Baltsavias 2010). Therefore, they may have different appearances with significant local distortions. To investigate these local differences, a combined local feature extraction approach is applied in this study; this is based on two types of features, namely, corner points and circular regions (Sedaghat and Ebadi 2015d).

To improve matching capability, for DEMs with different spatial resolutions, the DEM with the coarser resolution is up-sampled to that of the finer resolution. The combined feature extraction algorithm uses the well-known Harris operator (Harris and Stephens 1988) to extract corner features, while the UR-SIFT method is applied in order to extract circular “blob” features. To achieve feature stability, distinctiveness and distribution, a gridding selection strategy is also applied in the full distribution of the location and the scale space. The combinational extraction of different features with different properties offers an appropriate solution for the robust description of various structures in different DEMs with various scene types (Sedaghat and Ebadi 2015d).

For uniform combined local feature extraction, the number of required features for Harris corners and UR-SIFT regions should be initially determined in order to ensure features' density and distribution. In this article, based on our experiments on input DEM pairs, the number of the required Harris corner features, N_c , and the number of the required UR-SIFT region features, N_b , are set to 0.08% and 0.04% of the DEM pixel area, respectively. For example, for a sample DEM with a size of $1,000 \times 1,000$ pixels, the total number of required features of each type is set to $N_c = 8000$ and $N_b = 4000$. Additionally, to guarantee the quality of feature distribution, a regular gridding strategy is used. For this purpose, the input DEM is divided into a regular grid cells, with the feature selection process performed separately for each grid cell, based on initial feature qualities. In this article, a grid cell of 50×50 pixels is empirically used. Figure 2 indicates the extracted combined features in a sub sample of the DEM. As can be seen, the salient structures of the DEM scene, with different appearance and size, are extracted in one of the corners or blob region feature forms.

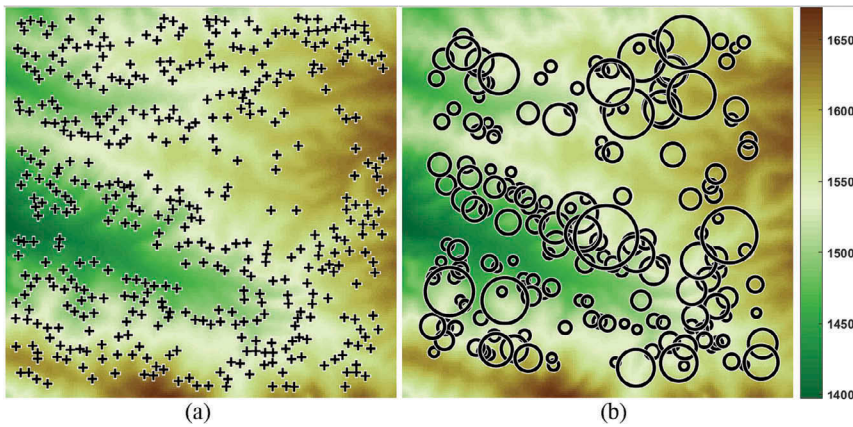


Figure 2. Integrated feature extraction in a sample DEM: (a) Harris points and (b) UR-SIFT regions.

2.2. Computation of the DOBSS descriptor

After feature extraction, the feature descriptor is generated using the DOBSS method. The DOBSS descriptor is an advanced extension of the LSS descriptor (Shechtman and Irani 2007) and has been proposed for highly reliable multi-sensor remote sensing image matching. It has several properties, such as invariance to rotation and illumination differences (in this case, elevation differences) with high descriptor discriminability; it is also computationally simpler than the SIFT algorithm (Sedaghat and Ebadi 2015b).

The self-similarity representation introduces a level of abstraction, which allows patterns in intensity, texture and edges to be treated in a single unified way (Chatfield, Philbin, and Zisserman 2009). Unlike gradient magnitudes and orientations, which are generally used in conventional descriptors, such as SIFT and DAISY, the applied correlation value in the DOBSS descriptor is a very robust measure against non-linear differences, due to the fact that it captures the internal geometric layouts of the region and represents the local shape property. It is therefore very applicable to feature correspondence for data in two DEMs coming from different sources, which may have differences in appearance with changes in local elevation. For more details about local LSS theoretical properties, see (Sedaghat and Ebadi 2015b; Shechtman and Irani 2007).

To obtain geometric invariance, the extracted UR-SIFT features are normalized to a canonical circular region (a 41×41 patch as suggested in Fan, Wu, and Hu (2012)). It should be noted that since the corner features have no shape parameters (circle or ellipse), a constant 41×41 patch is used. As mentioned earlier, in the proposed matching method, the DEM with a coarser resolution is resampled to that of the finer resolution. Therefore, the selected regions for conjugate feature pairs cover the same areas, such that the proposed DOBSS descriptor can easily compute the invariant of scale differences.

As illustrated in Figure 3, the steps of the DOBSS descriptor computation are as follows (Sedaghat and Ebadi 2015b):

- (1) Computation of the correlation surface: all surrounding small circular patches of radius R_{Patch} are compared with the central patch, using the sum of squared differences (SSD) between patch elevation values. The SSD values are then normalized and transformed into a correlation surface (Figure 3(b)).

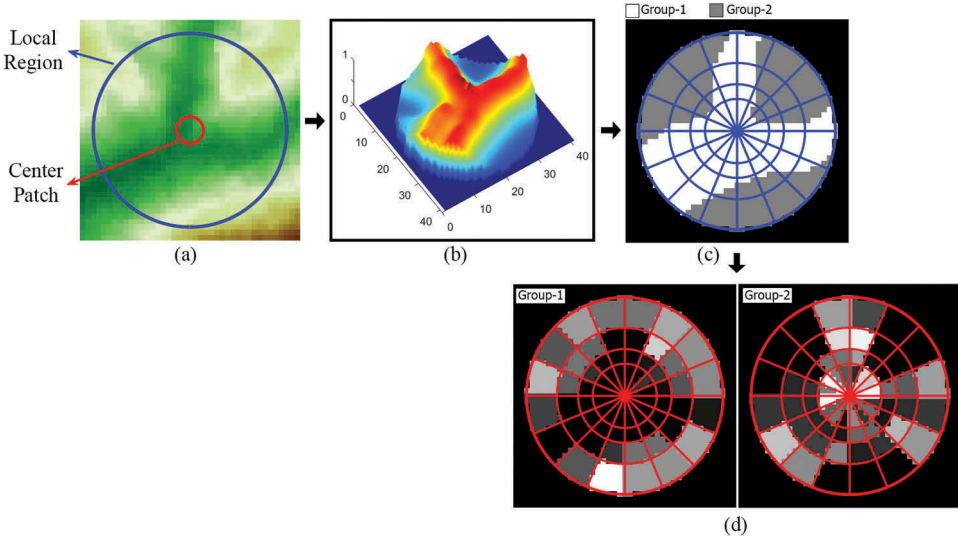


Figure 3. The DOBSS descriptor computation schema: (a) local sample DEM region, (b) computation of the correlation surface, (c) grouping process, and (d) log-polar translation.

- (2) Orientation assignment: an orientation histogram with n_{Hist} bins (covering 360°) is created based on weighted correlation values. Following this, the generated correlation surface is rotated in the direction of the local peak of the histogram.
- (3) Local region grouping: the pixels in the local region are adaptively partitioned into B_{Ord} ordinal groups, based on their correlation values (Figure 3(c)).
- (4) Log-polar translation and descriptor normalization: each pixel in the local region contributes to the log-polar descriptor structure (with n_θ and n_r for angular and radial bins, respectively). Then, the differences of the maximum and the median of the correlation value in each bin in the log-polar structure are concatenated (Figure 3(d)).

In the proposed method, the DOBSS descriptor parameters are set to $R_{Patch} = 3$, $n_{Hist} = 36$, $B_{Ord} = 2$, $n_\theta = 16$ and $n_r = 4$, as suggested by Sedaghat and Ebadi (Sedaghat and Ebadi 2015b).

2.3. Feature correspondence and blunder rejection

The computed descriptors of the extracted features from the two DEMs are compared, based on the Euclidean distance ratio, using a cross-matching strategy to achieve feature correspondence (Sedaghat and Ebadi 2015c). After feature correspondence, several mismatches generally exist, which should be removed. In general, the geometric consistency of the initial matched features is examined with a transformation model, whereby features with high levels of errors are identified as mismatches and eliminated. In this research, to overcome significant local DEM differences, a local consistency check process is applied, based on k-means clustering and the affine transformation model (Sedaghat and Ebadi 2015d). The main concept of this method is based on the assumption that DEM pairs from different sources with significant changes in local elevation can be locally approximated with a 2D affine model.

In summary, the local mismatch elimination process is performed as follows (Sedaghat and Ebadi 2015d):

- (1) Feature clustering: initial matched features are clustered into NC clusters using the k-means method (Figure 4(a)).
- (2) Local affine consistency check: the initial matched point pairs for each cluster are examined using an affine model. Those with the highest errors are rejected one by one until a desired root mean square error ($RMSE < T_A$) is achieved (Figure 4(b)). T_A is the affine local consistency check threshold for assessing DEM matching accuracy.

The number of clusters, NC, is defined as the ratio between the average number of initially matched features of each cluster, N_m , and the total number of the extracted initially matched

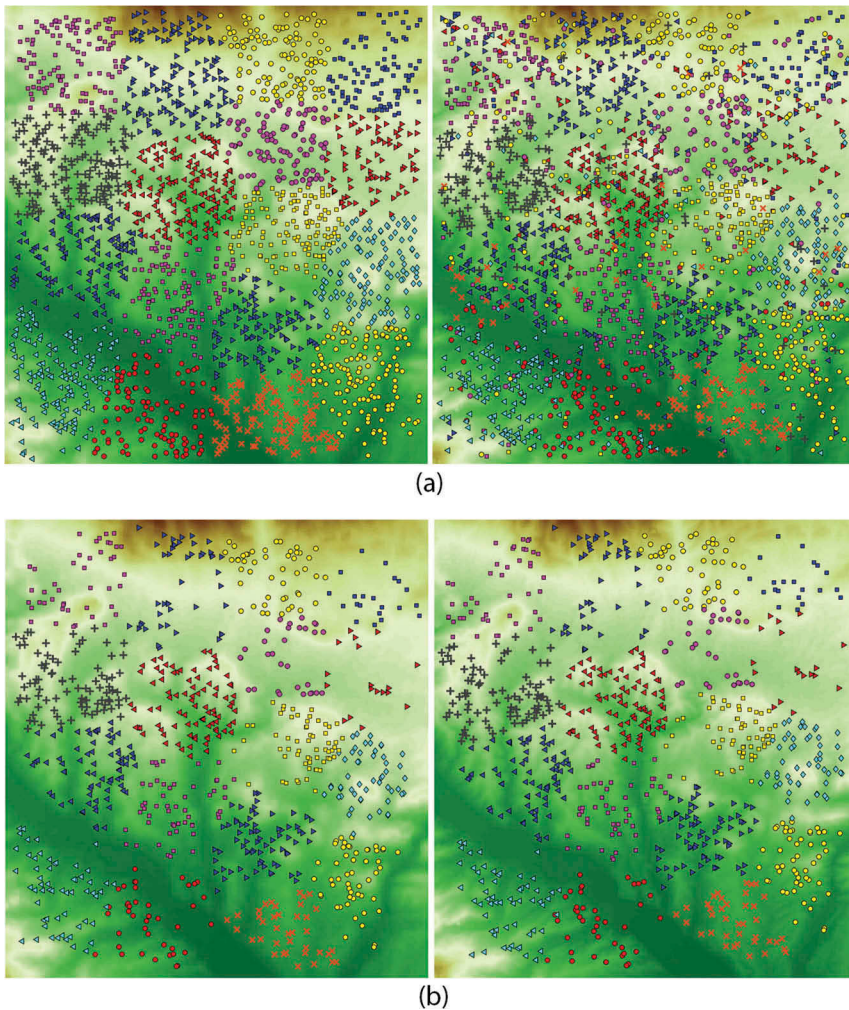


Figure 4. Mismatched elimination process based on a local consistency check: (a) feature clustering and (b) refinement clusters after the rejection of outliers.

features, N_t . In this article, N_m was set empirically to 120 features and T_A was set to 1.5 pixels. Figure 4 illustrates the results of the feature pairs' clustering and mismatched elimination process, based on a local consistency check approach for a sample of DEM pairs.

2.4. DEM orientation

In this section, the parameters of a 3D transformation function are estimated by means of the established corresponding features. Then, the RFM-derived DEM is transformed into the geometry of the existing global SRTM DEM by means of the estimated transformation functions. It should be noted that the transformed pixels of the elevation values in non-integer coordinates are computed by using a bilinear interpolation method. Various types of transformation functions can be used to establish DEM correspondence (Ravanbakhsh and Fraser 2013). In this research, three types of transformation functions are applied and compared: the 3D rigid, 3D similarity and 3D affine transformations. The 3D rigid and similarity transformations are used for shape-invariant mapping, as per Equations (1) and (2):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{R}_{\kappa\phi\omega} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix}, \quad (1)$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda \mathbf{R}_{\kappa\phi\omega} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix}, \quad (2)$$

where (X, Y, Z) are the coordinates of a 3D point in the SRTM DEM (x, y, z) are the coordinates of a 3D point in the input RFM-derived DEM, $\mathbf{R}_{\kappa\phi\omega}$ is a 3D rotation matrix (the product of rotations ω , ϕ and κ around the X , Y , and Z axes), $(\Delta X, \Delta Y, \Delta Z)$ is a translation vector between the two DEMs, and λ is a scale factor. In addition to the three rotation and three translation parameters, the 3D affine transformation requires three scaling factors and three additional shearing parameters for each coordinate axis; these are presented in Equation (3):

$$\begin{aligned} X &= a_0 + a_1x + a_2y + a_3z \\ Y &= b_0 + b_1x + b_2y + b_3z, \\ Z &= c_0 + c_1x + c_2y + c_3z \end{aligned} \quad (3)$$

where a_0 to c_3 are the 12 coefficients of the 3D affine model.

3. Experiments

In this section, the input DEM datasets, the evaluation criteria and experimental results are introduced.

3.1. Data sets

To evaluate the DEM matching method, two data sets were used, generated from two CARTOSAT-1 satellite stereo pairs. The first pair was captured from the Bomehen region in the north east of Tehran, Iran. This area consists of various landscapes, such as mountainous and hilly regions, as well as several flat urban regions. The second stereo pair was captured from

Tazehabad, an area near the west of Kermanshah, Iran. This area also includes various landscapes, such as flat and hilly regions, as well as some mountainous areas. Figure 5 shows the input stereo pairs, while Table 2 presents their characteristics.

To generate the RFM-derived DEM of each dataset, the PCI Geomatica OrthoEngine was used. To do this, tie points were first automatically generated using a fast Fourier transform phase matching method. The DEM was then generated using tie points and vendor-provided RPCs through a block adjustment procedure. The RFM-derived DEMs generated from two satellite stereo pairs are illustrated in Figure 6(a–d).

As a reference DEM for evaluation, a more accurate DEM was also generated from digitized 1:2000 maps for a limited area in both regions, as shown in Figure 6(b–e). These reference DEMs were produced from 3D points and contours using a triangulated irregular network (TIN) model. This model was then converted into the grid elevation data at a grid spacing of 10 m.

As mentioned above, the SRTM is used in this study, as a freely existing global DEM. These DEM data are provided by NASA, using the SAR interferometry technique, as shown Figure 6(c–f). Due to the inclined orbit of the satellite, the coverage of this DEM is principally limited to a latitude range of 56° S to 60° N (Nikolakopoulos,

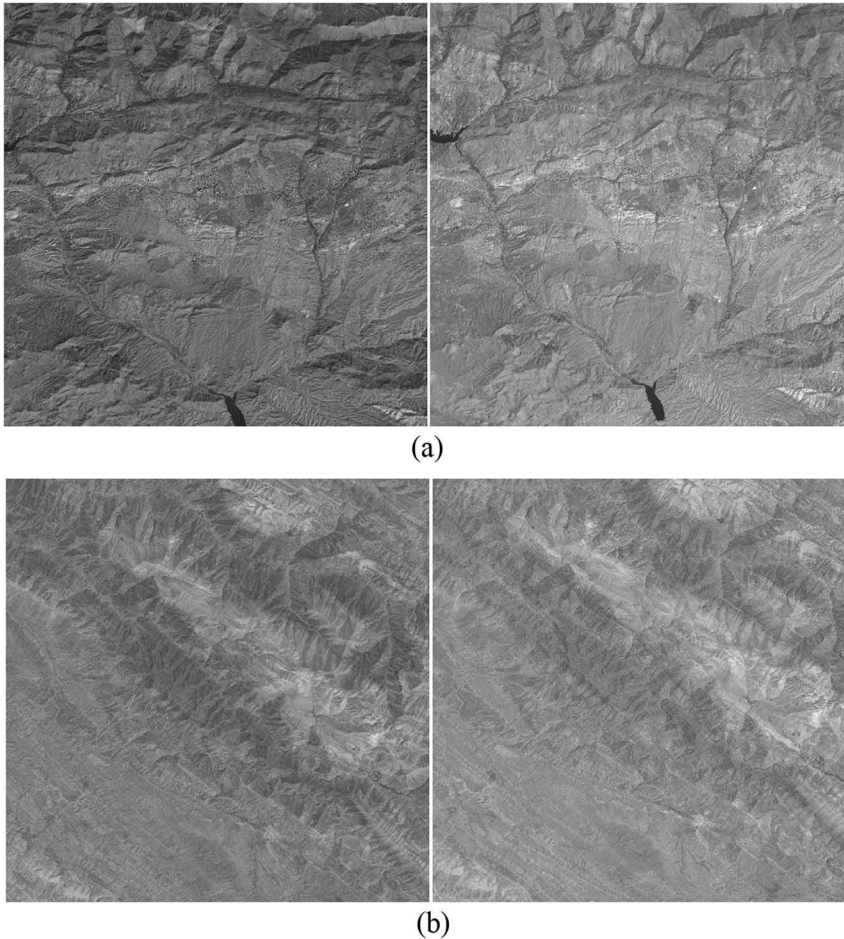


Figure 5. CARTOSAT-1 image pairs over the (a) Bomehen and (b) Tazehabad data sets.

Table 2. Characteristics of the input CARTOSAT-1 image pairs (see also Figure 5).

Dataset		Bomehen	Tazehabad
Spatial resolution (m)		2.5	2.5
Image size		12,000 × 12,000	12,000 × 12,000
Incidence angle (degree)	Left image	−5°	−5°
	Right image	26°	26°
No. of GCPs (#)		27	30

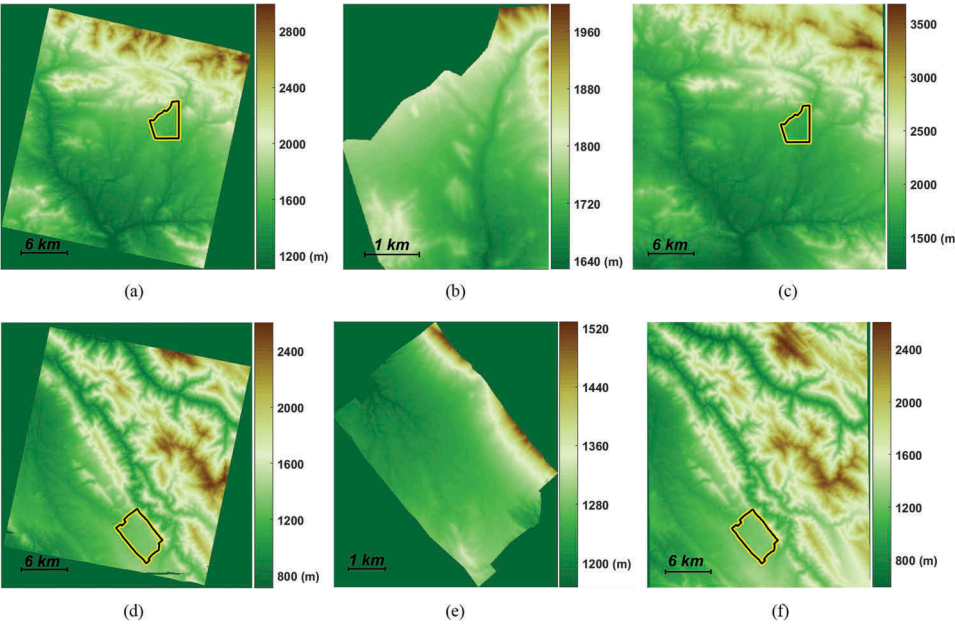


Figure 6. Input DEM datasets: (a) Bomehen RFM-derived DEM, (b) Bomehen reference DEM, (c) Bomehen SRTM DEM, (d) Tazehabad RFM-derived DEM, (e) Tazehabad reference DEM, and (f) Tazehabad SRTM DEM (the specified regions are covered by reference DEMs).

Kamaratakis, and Chrysoulakis 2006). The SRTM elevation data are distributed in both 30 and 90 m cell sizes. In the current research, the latest version with 30 m resolution was used.

As additional reference DEMs for evaluation, absolute DEMs generated by the PCI Geomatica OrthoEngine in the presence of GCPs were also used. It should be noted that GCPs were extracted from a 1:2000 map prepared by the National Cartographic Center of Iran. Table 3 summarizes the relevant details for all the input DEM data sets, including the RFM-derived, SRTM, reference and absolute DEMs.

3.2. Evaluation criteria

The quality of the proposed DEM matching method for the orientation process was evaluated using two criteria:

Table 3. Characteristics of the input DEMs (see also [Figures 5 and 6](#)).

Dataset	DEM	Grid size (m)	DEM size	Generation technology
Bomehen	RFM-derived DEM	10	3441 × 3292	CARTOSAT-1 satellite stereo pairs with PCI software and without GCPs
	SRTM	30	1164 × 1119	Space-borne IfSAR
	Reference DEM	10	487 × 377	1:2000 Topographic Map
	Absolute DEM	10	3488 × 3296	CARTOSAT-1 satellite stereo pairs with PCI software and GCPs
Tazehabad	RFM-derived DEM	10	3467 × 3280	CARTOSAT-1 satellite stereo pairs with PCI software and without GCPs
	SRTM	30	1164 × 978	Space-borne IfSAR
	Reference DEM	10	719 × 607	1:2000 Topographic Map
	Absolute DEM	10	3482 × 3282	CARTOSAT-1 satellite stereo pairs with PCI software and GCPs

- (1) Number of correct matches (NCM): the number of correct matched points in the DEM matching process are counted.
- (2) RMSE: the RMSEs of the height differences of all pixels in the overlapping area in two RFM-derived and reference DEMs were used as evaluation criteria. To compute these height differences, the heights obtained from the RFM-derived DEM were compared with the derived heights, which had been bilinearly interpolated from the reference DEM. The interpolation process is necessary due to pixel location differences between an RFM-derived DEM and a reference DEM.

To evaluate the capability of the adapted DOBSS descriptor, the results from the proposed approach were also compared with the results obtained using the standard SIFT descriptor.

3.3. Results and discussion

The proposed DEM matching and orientation approach was implemented in MATLAB 2016a. The details and values of the input parameters for the proposed method were set as discussed above. In the proposed method, the RFM-derived DEM is transformed to a global DEM space using an appropriate transformation model. The results of the proposed matching and orientation method, which are given in [Table 4](#), demonstrate the efficacy of the proposed method in the DEM matching process, where DEMs with a different resolution and accuracy are involved. According to the NCM and RMSE criteria, the proposed method, based on the DOBSS descriptor outperforms the standard SIFT descriptor, as shown in [Table 4](#). The main reason for DOBSS descriptor superiority is its robustness against non-linear local differences in the DEM pair.

[Figure 7](#) illustrates the results of the DEM matching by the proposed method for two data sets, which demonstrate the well distribution quality of the extracted feature matches. The proposed feature-based matching method is fully automatic. Furthermore, unlike existing area-based matching methods, the proposed method does not require initial values.

[Table 4](#) compares the accuracy of the matching and orientation process. As can be seen, the RMSE of the height differences between the RFM-derived DEM and the

Table 4. Final DEM matching and orientation results.

Dataset	Compare DEM	RMSE before orientation (m)	RMSE after orientation (m)										NCM (#)	
			3D Rigid		3D Similarity		3D Affine							
			SIFT	DOBSS	SIFT	DOBSS	SIFT	DOBSS	SIFT	DOBSS				
			SIFT	DOBSS	SIFT	DOBSS	SIFT	DOBSS	SIFT	DOBSS				
Bornehen	Reference DEM	131.5	6.7	4.8	5.3	4.2	5.1	3.9	656	3767				
	Absolute DEM	153.9	8.7	7.9	8.5	7.1	7.7	6.6						
Tazehabad	Reference DEM	73.5	5.9	3.3	4.5	2.7	4.2	2.5	457	3229				
	Absolute DEM	97.6	11.8	9.1	11.5	8.5	10.2	8.4						

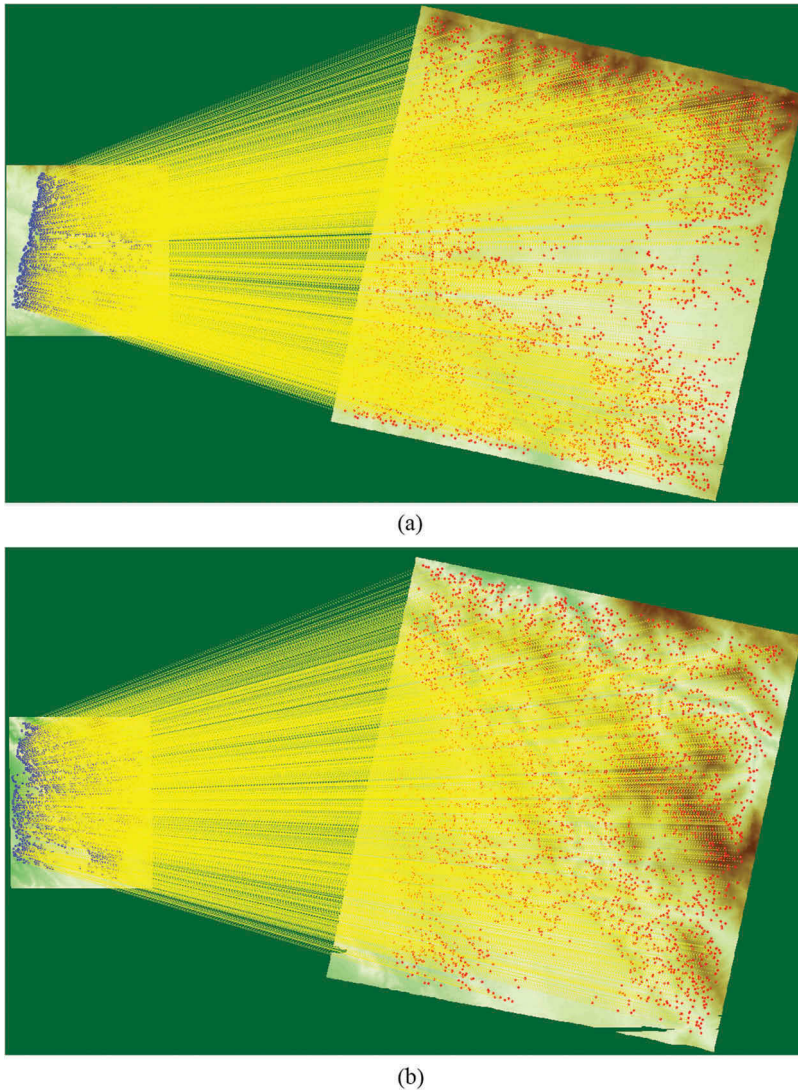


Figure 7. DEM matching results for input datasets: (a) Bomehen and (b) Tazehabad.

reference DEM after the orientation process is significantly smaller than that before orientation. In both data sets, when the orientation process is performed with all three transformation models, the accuracy of the oriented RFM-derived DEM is significantly improved with respect to the unoriented DEM.

According to Table 4, the 3D affine model gives the highest accuracy, while the 3D rigid model gives the lowest accuracy. Accordingly, the model used for DEM orientation contains a scale parameter. On the other hand, given that there is about a 7% improvement in the 3D affine model compared to the 3D similarity model, the 3D affine model is the best model for orienting DEMs derived from CARTOSAT-1 images.

Since the reference DEMs cover a limited area of the RFM-derived DEM, another reference DEM, which was manually generated using PCI and GCP points, was also

applied. As can be seen in Table 4, the obtained RMSE from the second reference data set also confirms the capability of the proposed method for the accurate orientation of RFM-derived DEM data using an existing global base model. However, the results of the proposed method deteriorate in the case where absolute DEMs are used for evaluation. Assuming absolute DEMs are generated accurately, this means there are no mismatches, which could be explained by the fact that unlike the reference DEMs, the absolute DEMs cover various landscapes, including mountainous areas.

Figure 8 shows the pixel-to-pixel height differences between the unoriented and oriented RFM-derived DEMs and the accurate reference DEM for two data sets, both before and after the orientation process, with the 3D affine transformation function. Figure 8(a–c) shows the initial height differences, while Figure 8(b–d) shows the height differences after the RFM-derived DEMs have been oriented. The histograms of the height differences for the two data sets are also shown in Figure 9.

It can be concluded that automatic RFM-derived DEM matching with an existing global DEM, based on local features and the 3D transformation model, can be utilized for accurate RFM-derived DEM orientation without the use of any GCP points. It should be

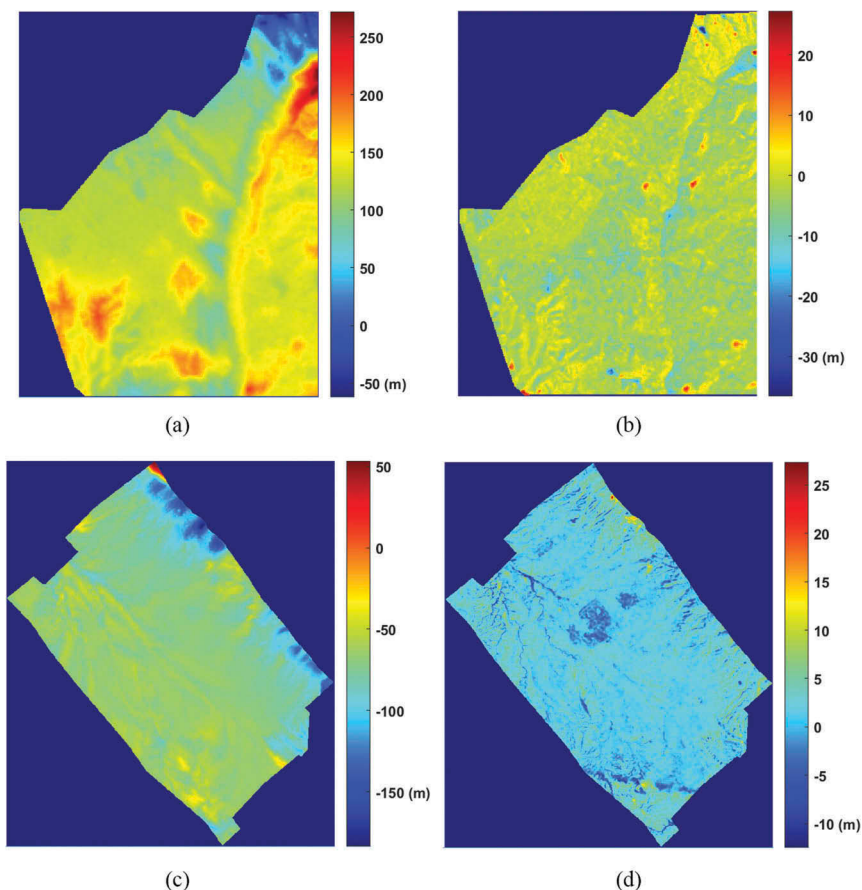


Figure 8. Height differences between oriented RFM-derived DEMs based on a 3D affine model and reference DEMs for two data sets: (a) and (b) before and after orientation for Bomehen; (c) and (d) before and after orientation for the Tazehabad data set.

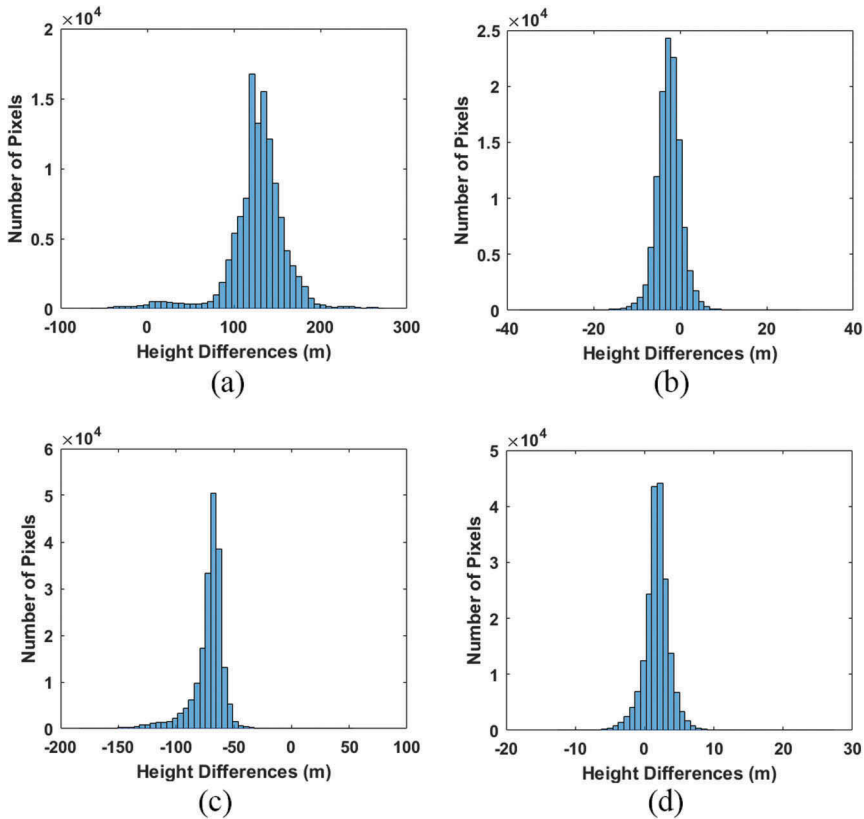


Figure 9. Histograms of the height differences between the oriented RFM-derived DEMs based on the 3D affine and reference DEMs for two data sets: (a) and (b) before and after orientation for Bomehen; (c) and (d) before and after orientation for the Tazehabad data set.

noted that the quality of the proposed DEM orientation method depends entirely on the accuracy of the global SRTM DEM. The proposed method, which is used to compensate for the bias caused by RFM errors, clearly cannot eliminate the internal errors in the RFM-derived DEMs for various reasons such as a dense matching process.

To investigate the performance of the DOBSS descriptor in the DEM matching process, the SIFT, DAISY (Tola, Lepetit, and Fua 2010), LIOP (Wang, Fan, and Wu 2011), LBP (Heikkilä, Pietikäinen, and Schmid 2009), and BRISK (Leutenegger, Chli, and Siegwart 2011) descriptors are also evaluated and compared. For all descriptors, the proposed combined feature extraction process based on Harris and UR-SIFT algorithms is applied. Figure 10 indicates the comparative results of the six descriptors on two data sets in terms of the NCM, and the RMSE of the height differences in the 3D affine transformation model. It can be observed that the DOBSS descriptor outperformed the other descriptors in the two data sets. The DOBSS descriptor is based on the self-similarity measure with a distinctive structure and very robust in the face of significant local differences. The LIOP and LBP descriptors are based on intensity orders and provide the next levels of DEM matching results. The DAISY and SIFT descriptors, which are based on intensity gradients, are ranked in the next levels. The BRISK binary descriptor is sensitive to the DEM matching process and provides the weakest results.

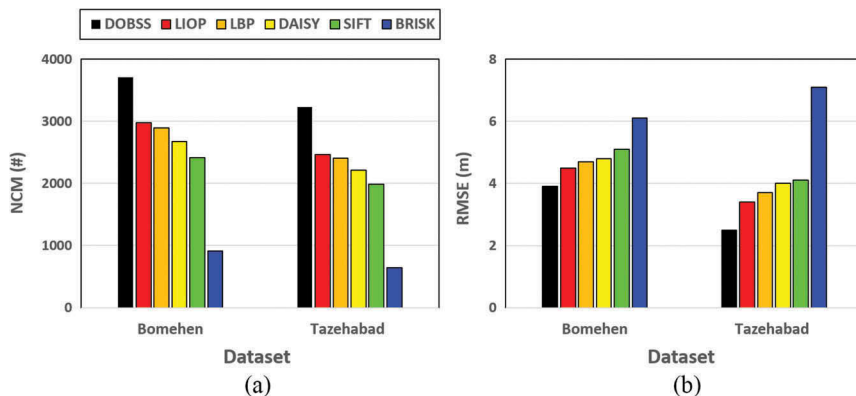


Figure 10. DEM matching results for various descriptors: (a) NCM and (b) RMSE in the 3D affine model.

Different from other distribution based descriptors, which are based on gradient or intensity order, the DOBSS descriptor is characterized by comparing the self-similarity of a central patch with its neighboring patches in a distinctive grouping manner. Theoretically, the self-similarity measure is resistant to non-linear differences because it captures the internal geometric layouts of the region and represents the local shape property. Two different DEMs from different sources may have differences in appearance with changes in local elevation. The self-similarity property is interesting for DEM matching with local distortions because their local shape is similar, but their elevations or texture may be different. Therefore, it can be effectively used to match a DEM local region with other differently textured DEM regions, as long as they have similar layouts and shapes. The experimental results prove that DOBSS is more robust than other distribution based descriptors to the non-linear height differences for multi-source DEM matching process.

To evaluate the computational complexity of the descriptor, the average running time of the descriptor was measured for each feature. Using an Intel Core i7 CPU 2.4 GHz with 8 GB of memory, the average running time for each feature is 2.7 ms for DOBSS, 2.9 ms for SIFT, 10.6 ms for DAISY, 1.8 ms for LBP, 2.6 ms for LIOP, and 0.4 ms for BRISK. As can be seen, the DOBSS descriptor provides a moderate level of computational complexity.

As another evaluation, a comparative experiment was also conducted using the mutual information (MI) ABM matching method (Ravanbakhsh and Fraser 2012). A template window with the size of 1×1 km was selected, as suggested by Ravanbakhsh and Fraser (2012), after which the entire RFM-derived DEM was sampled at the template size interval. The initial corresponding location of these sampled templates in the global SRTM DEM were selected manually, and then introduced into the MI ABM matching method. Afterwards, MI similarity measures between the sampled template from the RFM-derived DEM and the global SRTM DEM were computed. The maximum value of the MI measure in a 10×10 pixel search window was determined as the final correspondence of each sampled template. The RMSE of the height differences in the case of 3D affine transformation were 6.1 and 7.3 m, respectively, for the Bomehen and Tazehabad areas. This result demonstrates a superior performance of the proposed FBM DEM matching method over the MI ABM method. It should be noted that, unlike the MI ABM method, the proposed method, does not require initial values.

4. Conclusions

This article introduces a robust local feature-based DEM matching approach for the accurate orientation of a RFM-derived DEM with no GCPs. For this purpose, the RFM-derived DEM generated from a CARTOSAT-1 satellite stereo image pair was matched to a global SRTM DEM. The matching method used the Harris and UR-SIFT feature detectors, a DOBSS descriptor and the k-means-based local consistency check process to establish accurate feature correspondence between the RFM-derived DEM and the global SRTM DEM. After feature correspondence in the DEMs, a 3D transformation function was used for RFM-derived DEM orientation. Three well-known transformation functions (3D rigid, 3D similarity, and 3D affine), were used to evaluate the effect of transformation type. The experimental results for two DEM data sets derived from two CARTOSAT-1 satellite image pairs demonstrated the advantages of this algorithm in terms of the NCM and positional accuracy (RMSE). The results of the proposed DEM matching method indicated that it can be used for reliable DEM orientation.

Since the proposed feature-based DEM matching method is based on finding salient conjugate features in the same area of a DEM pair, the difference between the acquisition times of the two data sets (stereo image pairs and global SRTM DEM) will not significantly affect the DEM orientation results. However, investigating the impact of the acquisition times on the DEM matching and orientation process is suggested as future work. As with any new approach, there are certain unresolved issues, which may present themselves as challenges in due course. One of the most important of these challenges is the type of image: in this work, CARTOSAT-1 stereo images were used, whereas different stereo images such as SPOT, could also be considered and evaluated.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Amin Sedaghat  <http://orcid.org/0000-0002-0414-5800>

References

- Aguilar, M. A., M. del Mar Saldana, and F. J. Aguilar. 2013. "Assessing Geometric Accuracy of the Orthorectification Process from GeoEye-1 and WorldView-2 Panchromatic Images." *International Journal of Applied Earth Observation and Geoinformation* 21: 427–435. doi:10.1016/j.jag.2012.06.004.
- Akca, D. 2010. "Co-Registration of Surfaces by 3D Least Squares Matching." *Photogrammetric Engineering & Remote Sensing* 76: 307–318. doi:10.14358/PERS.76.3.307.
- Belongie, S., J. Malik, and J. Puzicha. 2002. "Shape Matching and Object Recognition Using Shape Contexts." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24: 509–522. doi:10.1109/34.993558.
- Besl, P. J., and N. D. McKay. 1992. "A Method for Registration of 3-d Shapes." *Review of. IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (2): 239–256.
- Castillo-Carrión, S., and J.-E. Guerrero-Ginel. 2017. "SIFT Optimization and Automation for Matching Images from Multiple Temporal Sources." *International Journal of Applied Earth Observation and Geoinformation* 57: 113–122. doi:10.1016/j.jag.2016.12.017.
- Chatfield, K., J. Philbin, and A. Zisserman. 2009. "Efficient Retrieval of Deformable Shape Classes Using Local Self-similarities." Paper presented at the 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, Kyoto, October 2009.

- d'Angelo, P., M. Lehner, T. Krauss, D. Hoja, and P. Reinartz. 2008. "Towards Automated DEM Generation from High Resolution Stereo Satellite Images." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, International Society for Photogrammetry and Remote Sensing* 37: 1137–1342.
- d'Angelo, P., and P. Reinartz. 2012. "DSM Based Orientation of Large Stereo Satellite Image Blocks." *International Archives Photogramm Remote Sens Spatial Information Sciences* 39: 209–214. doi:10.5194/isprsarchives-XXXIX-B1-209-2012.
- Fan, B., F. Wu, and Z. Hu. 2012. "Rotationally Invariant Descriptors Using Intensity Order Pooling." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34: 2031–2045. doi:10.1109/TPAMI.2011.277.
- Fraser, C. S., G. Dial, and J. Grodecki. 2006. "Sensor Orientation via RPCs." *ISPRS Journal of Photogrammetry and Remote Sensing* 60: 182–194. doi:10.1016/j.isprsjprs.2005.11.001.
- Fraser, C. S., and H. B. Hanley. 2003. "Bias Compensation in Rational Functions for IKONOS Satellite Imagery." *Photogrammetric Engineering & Remote Sensing* 69: 53–57. doi:10.14358/PERS.69.1.53.
- Gil, M., M. Arza, J. Ortiz, and A. Ávila. 2014. "DEM Shading Method for the Correction of Pseudoscopic Effect on Multi-Platform Satellite Imagery." *GIScience & Remote Sensing* 51: 630–643. doi:10.1080/15481603.2014.988433.
- Gonçalves, J. 2006. "Orientation of SPOT Stereopairs by Means of Matching a Relative DEM and the SRTM DEM." In *Proceedings of the internaional Calibration and Orientation Workshop-EuroCow2006*, Castelldefels, January 25.
- Harris, Chris, and Mike Stephens. 1988. "A combined corner and edge detector." In *Proceedings of Fourth Alvey Vision Conference*, Manchester. September 1988.
- Heikkilä, M., M. Pietikäinen, and C. Schmid. 2009. "Description of Interest Regions with Local Binary Patterns." *Pattern Recognition* 42: 425–436. doi:10.1016/j.patcog.2008.08.014.
- Jannati, M., and M. J. Valadan Zoej. 2015. "Introducing Genetic Modification Concept to Optimize Rational Function Models (RFMs) for Georeferencing of Satellite Imagery." *GIScience & Remote Sensing* 52: 510–525. doi:10.1080/15481603.2015.1052634.
- Kim, T., and J. Jeong. 2011. "DEM Matching for Bias Compensation of Rigorous Pushbroom Sensor Models." *ISPRS Journal of Photogrammetry and Remote Sensing* 66: 692–699. doi:10.1016/j.isprsjprs.2011.06.002.
- Leutenegger, S., M. Chli, and R. Y. Siegwart. 2011. "BRISK: Binary Robust Invariant Scalable Keypoints." Paper presented at the Computer Vision (ICCV), 2011 IEEE International Conference, Barcelona, November 2011.
- Li, C., X. Liu, and T. Deng. 2016. "A New Optimized RFM of High-Resolution Satellite Imagery." *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* 41: 65–69. doi:10.5194/isprsarchives-XLI-B3-65-2016.
- Lowe, D. G. 2004. "Distinctive Image Features from Scale-Invariant Keypoints." *International Journal of Computer Vision* 60: 91–110. doi:10.1023/B:VISI.0000029664.99615.94.
- Matas, J., O. Chum, M. Urban, and T. Pajdla. 2004. "Robust Wide-Baseline Stereo from Maximally Stable Extremal Regions." *Image and Vision Computing* 22: 761–767. doi:10.1016/j.imavis.2004.02.006.
- MATLAB. 2016a, *The Mathworks, Inc.*, Natick, MA.
- Milledge, D., S. N. Lane, and J. Warburton. 2009. "Optimisation of Stereo-Matching Algorithms Using Extant DEM Data." *Photogrammetric Engineering and Remote Sensing* 75: 323–333. doi:10.14358/PERS.75.3.323.
- Nikolakopoulos, K. G., E. K. Kamaratakis, and N. Chrysoulakis. 2006. "SRTM Vs ASTER Elevation Products. Comparison for Two Regions in Crete, Greece." *International Journal of Remote Sensing* 27: 4819–4838. doi:10.1080/01431160600835853.
- Oh, J., and C. Lee. 2015. "Automated Bias-Compensation of Rational Polynomial Coefficients of High Resolution Satellite Imagery Based on Topographic Maps." *ISPRS Journal of Photogrammetry and Remote Sensing* 100: 14–22. doi:10.1016/j.isprsjprs.2014.02.009.
- Oh, J., C. Lee, and D. C. Seo. 2013. "Automated HRSI Georegistration Using Orthoimage and SRTM: Focusing KOMPSAT-2 Imagery." *Computers & Geosciences* 52: 77–84. doi:10.1016/j.cageo.2012.09.026.
- Papasaika, H., and E. Baltsavias. 2010. "Quality Evaluation of DEMs." Paper presented at the International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Leicester, July 2010.

- Ravanbakhsh, M., and C. Fraser. 2012. "DEM Registration Based on Mutual Information." Presented paper at ISPRS Congress 2012, Melbourne, September 2012.
- Ravanbakhsh, M., and C. S. Fraser. 2013. "A Comparative Study of DEM Registration Approaches." *Journal of Spatial Science* 58: 79–89. doi:10.1080/14498596.2012.759091.
- Sedaghat, A., and H. Ebadi. 2015a. "Accurate Affine Invariant Image Matching Using Oriented Least Square." *Photogrammetric Engineering & Remote Sensing* 81: 733–743. doi:10.14358/PERS.81.9.733.
- Sedaghat, A., and H. Ebadi. 2015b. "Distinctive Order Based Self-Similarity Descriptor for Multi-Sensor Remote Sensing Image Matching." *ISPRS Journal of Photogrammetry and Remote Sensing* 108: 62–71. doi:10.1016/j.isprsjprs.2015.06.003.
- Sedaghat, A., and H. Ebadi. 2015c. "Remote Sensing Image Matching Based on Adaptive Binning SIFT Descriptor." *IEEE Transactions on Geoscience and Remote Sensing* 53: 5283–5293. doi:10.1109/TGRS.2015.2420659.
- Sedaghat, A., and H. Ebadi. 2015d. "Very High Resolution Image Matching Based on Local Features and K-Means Clustering." *The Photogrammetric Record* 30: 166–186. doi:10.1111/phor.12101.
- Sedaghat, A., H. Ebadi, and M. Mokhtarzade. 2012. "Image Matching of Satellite Data Based on Quadrilateral Control Networks." *The Photogrammetric Record* 27: 423–442. doi:10.1111/phor.12101.
- Sedaghat, A., M. Mokhtarzade, and H. Ebadi. 2011. "Uniform Robust Scale-Invariant Feature Matching for Optical Remote Sensing Images." *IEEE Transactions on Geoscience and Remote Sensing* 49: 4516–4527. doi:10.1109/TGRS.2011.2144607.
- Shechtman, E., and M. Irani. 2007. Matching local Self-Similarities Across Images and Videos. Paper presented at the Computer Vision and Pattern Recognition, 2007. Cvpr'07. IEEE Conference on, Minnesota, June 2007.
- Streutker, D. R., N. F. Glenn, and R. Shrestha. 2011. "A Slope-Based Method for Matching Elevation Surfaces." *Photogrammetric Engineering & Remote Sensing* 77: 743–750. doi:10.14358/PERS.77.7.743.
- Teo, T.-A., and S.-H. Huang. 2013. "Automatic Co-Registration of Optical Satellite Images and Airborne LiDAR Data Using Relative and Absolute Orientations." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6: 2229–2237. doi:10.1109/JSTARS.2012.2237543.
- Tola, E., V. Lepetit, and P. Fua. 2010. "Daisy: An Efficient Dense Descriptor Applied to Wide-Baseline Stereo." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32: 815–830. doi:10.1109/TPAMI.2009.77.
- Toutin, T., C. Schmitt, and H. Wang. 2012. "Impact of No GCP on Elevation Extraction from WorldView Stereo Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 72: 73–79. doi:10.1016/j.isprsjprs.2012.05.009.
- Wang, Z., B. Fan, and F. Wu. 2011. Local Intensity Order Pattern for Feature Description. Paper presented at the Computer Vision (ICCV), 2011 IEEE International Conference on, Barcelona, November 2011.
- Wright, R., H. Garbeil, S. M. Baloga, and P. J. Mougini-Mark. 2006. "An Assessment of Shuttle Radar Topography Mission Digital Elevation Data for Studies of Volcano Morphology." *Remote Sensing of Environment* 105: 41–53. doi:10.1016/j.rse.2006.06.002.
- Yu, X., Z. Lyu, D. Hu, and J. Xu. 2013. "Scale-Invariant Feature Transform Based on the Frequency Spectrum and the Grid for Remote Sensing Image Registration." *GIScience & Remote Sensing* 50: 543–561.