MODIFICATIONS IN THE SIFT OPERATOR FOR EFFECTIVE SAR IMAGE MATCHING

Sahil Suri, Peter Schwind[1], Johannes Uhl and Peter Reinartz

German Aerospace Center DLR, Germany

German Aerospace Center DLR
Remote Sensing Technology Institute
Münchener Str. 20, 82234 Wessling, Germany

http://www.dlr.de/caf/

---

[1] Corresponding author: Peter.Schwind@dlr.de

Taylor & Francis
Taylor & Francis Group

Modifications in the SIFT operator for effective SAR image matching

Sahil Suri, Peter Schwind*, Johannes Uhl and Peter Reinartz

Remote Sensing Technology Institute (IMF), German Aerospace Centre (DLR), 82234 Wessling, Germany

With the increasing availability and rapidly improving the spatial resolution of synthetic aperture radar (SAR) images from the latest and future satellites like TerraSAR-X and TanDEM-X, their applicability in remote sensing applications is set to be paramount. Considering challenges in the field of point feature-based multisensor/multimodal SAR image matching/registration and advancements in the field of computer vision, we extend the applicability of the scale invariant feature transform (SIFT) operator for SAR images. In this article, we have analysed the feature detection, identification and matching steps of the original SIFT processing chain. We implement steps to counter the speckle influence, which deteriorates the SIFT operator performance for SAR images. In feature identification, we evaluate different local gradient estimating techniques and highlight the fact that giving up the SIFT's rotation invariance characteristic increases the potential number of matches when the multiple SAR images from different sensors have been acquired with the same geometrical acquisition parameters. In the feature matching stage, we propose to assist the standard SIFT matching scheme to utilise the SIFT operator capability for effective results in challenging SAR image matching scenarios. The results obtained for SAR images acquired by different sensors using different incidence angles and orbiting directions over both rural and semi urban land cover, highlight the SIFT operator's capability for point feature matching in SAR imagery.

Keywords: SIFT; multimodal; SAR; image matching/registration; edge detection

1. Introduction

With the increasing availability and rapidly improving the spatial resolution of remote sensing synthetic aperture radar (SAR) images from latest and future satellites like TerraSAR-X (Eineder *et al.* 2005) and TanDEM-X (Krieger *et al.* 2007), the applicability of SAR images in various applications is bound to experience a tremendous boost. In particular, the SAR sensor's active nature in the microwave range of the electromagnetic spectrum gives them the capability to see through clouds and to acquire images at night, which might be the only viable alternative during a catastrophic event. Already with the present state of technology, SAR imagery is useful for diverse applications (Eineder *et al.* 2005). However, SAR image matching (especially high resolution) still does not have a perfect solution. Leaving aside area based techniques (Suri *et al.* 2009), feature-based SAR image registration/matching is difficult due to various reasons with the presence of multiplicative speckle influence being one of them. Reviewing research in SAR

*Corresponding author. Email: Peter.Schwind@dlr.de

image matching, an active contour-based model was utilised in Li *et al*. (1995) to detect and match contours in SAR and optical imagery. A method utilising SAR road extraction algorithms to make crossroad and road junctions in dense urban areas serving as registration control points was presented by Dell'Acqua *et al*. (2004). A point feature-based approach was used successfully for matching SAR images in Bentoutou *et al*. (2005). Here, the authors utilised the Harris operator for the detection of strong corner points. Subsequently, a template matching algorithm using similarity measure based on invariants was used. The mentioned techniques have their own pros and cons, like extraction of linear features and their intersections in SAR images, which is not very straight forward and extraction of interest points using Harris operator, which does not yield ideal points as many points are generally a result of salt and pepper noise in SAR images. In this article, we evaluate the application of the scale invariant feature transform (SIFT) (Lowe 2004) operator for remote sensing SAR image matching applications. SIFT is a computer vision algorithm and the method developed for optical camera images provides a robust technique to extract and match distinctive features within different images.

2. Scale invariant feature transform

To begin with, we briefly review the SIFT operator processing chain, for details please refer to the work in Lowe (2004).

2.1 *Feature detection*

The process starts with feature detection, for this purpose, a difference of Gaussians (DoG) pyramid is constructed by subtracting Gauss-filtered images where the standard deviation σ differs by a factor k. Next, extrema are detected in the DoG images by comparing every pixel to its neighbourhood in the scale space. If a pixel value is larger or smaller than all of its neighbours, it is accepted as a preliminary feature candidate. Finally, every feature candidate is interpolated for localisation and is filtered using additional stability checks.

2.2 *Orientation assignment*

To achieve rotation invariant features, one or more multiple orientation of a feature are determined. First, the gradient magnitude and orientation is computed for all scales of the Gauss pyramid. To find the dominant orientations, a 36 bin histogram is created. The gradient orientations surrounding the feature, weighted by their gradient magnitude and a Gaussian window, centred at the feature location, are used to fill the histogram. Once the histogram is created, the highest peaks are selected as dominant orientations. The gradient magnitudes and coordinates are rotated according to the assigned orientation to achieve rotation invariance (RI).

2.3 *Descriptor formation*

Scale invariance is assured by using the image gradients (estimated through adjacent pixel differencing) calculated at the scale the feature belongs to. The gradient magnitudes surrounding the feature are weighted by a Gaussian window to weaken the influence of

gradients far away from the feature. The descriptor window is divided up into 4×4 sub-windows to calculate 16 gradient histograms, each consisting of eight bins. The final descriptor is a vector holding 128 elements (16 histograms composed of eight bins).

2.4 *Feature matching*

Euclidean distance between feature descriptors is the similarity measure used to match two features. The two closest features in the input image are determined for every reference image feature. If the Euclidean distance of the second-closest match is smaller than the threshold (0.6) times the distance of the closest match, the input feature is accepted as a match. Since comparing the distance of all features with each other would be too expensive, the approximate nearest neighbour best-bin-first (BBF) algorithm is used (Beis and Lowe 1997).

3. SIFT in context of SAR imagery

In our previous work (Schwind *et al*. 2010), we focussed mostly on improvements of the feature detection phase of the SIFT operator, while in this article changes to the feature identification and matching phases are discussed. In the following, we introduce and analyse previous as well as newly introduced changes in the SIFT chain (component wise) to be used for SAR images.

3.1 *Feature detection*

Initiating the analysis, here we show the SIFT features detected on an E-SAR (developed by DLR) (Schreiber *et al*. 1999) image of size 300×300 pixels (Figure 1). The application of the original SIFT operator on the image depicted in Figure 1 results in 1520 SIFT
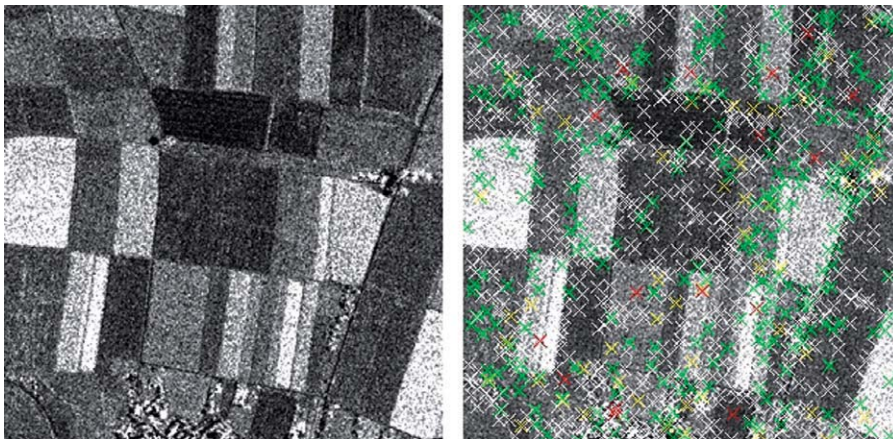

Figure 1. E-SAR imagery of size 300×300 pixels, SIFT operator produces 1520 features with around 74% features being detected at the first octave. The figure depicts 1119 octave 0, 315 octave 1, 63 octave 2 and 23 octave 3 SIFT features marked by white, green, yellow and red crosses, respectively.

features being detected. On analysing the SIFT feature distribution in different octaves, it becomes evident that maximum features have been found in first octave (set of highest scales in the scale space pyramid created for feature detection) and on their visual analysis it is inferred that most of them are a result of the multiplicative speckle influence. Based on this initial observation, in our previous work (Schwind *et al.* 2010), we studied the SIFT detection phase to propose the following:

(1) Skipping the features detected at the highest octave in the scale space pyramid: The idea here was to show that stable features in SAR imagery are detected lower down the scale space pyramid. We termed this detection as SIFT-OCT to represent exclusion of all the SIFT features detected at the first octave.

(2) Image smoothing by filtering the SAR images using the infinite size symmetrical exponential filters (ISEF) filter (Shen and Caston 1992): In the cited reference, the authors highlighted that an increase in Gaussian filter size is useful to reduce noise influence but the increase size has an adverse effect on edge localisation precision. To overcome this problem, the ISEF filter with an infinite window size and desired sharpness at the window centre was proposed. We utilised the ISEF filter for image smoothing and also skipped the first octave features and referred this detection scheme as SIFT + ISEF-OCT.

The original SIFT operator detection scheme along with two described minor variants (SIFT-OCT, SIFT + ISEF-OCT) were compared on the basis of feature detector's repeatability (Mikolajczyk and Schmid 2005), found number of correct matches, match consistency and registration turn around time for multimodal SAR images. In this study, by multimodal SAR images we refer to multimodal analysis of SAR images acquired by different sensors using different acquisition parameters at different points of time (Zitová and Flusser 2003). Specifically, we found that the capability of the SIFT operator to detect and match features in multimodal SAR images improved significantly by utilising the above-mentioned schemes. Further, changes in the detection, led to a 10 times speedup (approx) for 1000×1000 pixel images. Here, we utilise the SIFT-OCT detection scheme although we expect similar results after image smoothing using ISEF or any other filter as well.

3.2 Feature identification

In this phase, the main step of estimating the local gradients through simple image differencing (SD) of neighbouring pixels assigns strong gradient magnitudes to speckle pixels and this might disrupt proper feature descriptor formation. Therefore, we attempt to use the following gradient estimation techniques.

3.2.1 Operator based upon image recursive differencing

Proposed for detecting step edges in noisy SAR images (Paillau 1997), the utilised operator is derived from Deriche filter and is implemented recursively. The operator specifically designed for SAR images according to the edge detection criteria's proposed in Canny (1983):

(1) Insensibility to noise.
(2) Good edge localisation.
(3) Unique response to one edge,

was shown to be more effective than classical differential operators like Shen–Castan operator (Shen and Castan 1986), first derivative of a Gaussian and the original Deriche operator (Deriche 1987).

3.2.2 *Operator based upon image ratio*

Considering the fact that due to the multiplicative nature of speckle, difference-based edge detectors detect more false edges in areas of high reflectivity (this might lead to wrong feature descriptor formation), we also test an edge detector with a constant false alarm ratio (CFAR) (Touzi *et al.* 1988). For this comparison, we evaluate ratio of exponential weighted averages (ROEWA) operator proposed in Fjørtoft *et al.* (1995). This operator was found to be much more efficient for SAR image edge detection in comparison to the ratio of averages (Touzi *et al.* 1988), ratio of order statistics (Bovik *et al.* 1988) and ratio-based version of Deriche's optimal edge detector. The SIFT descriptor formation or feature identification is a function of a region of pixels (edge magnitude and orientation) selected for feature identification. This article is intended to extend the application of the SIFT operator for SAR images and so we have compared techniques to estimate $\mathrm{d}x$ and $\mathrm{d}y$ (suited to SAR imagery) for feature identification.

3.2.3 *Simple differencing (used in the original SIFT implementation)*

The original SIFT implementation supports the simple first order derivatives calculated using the following equations:

$$\mathrm{d}x(i,j) = I(i+1,j) - I(i-1,j) \tag{1}$$

$$\mathrm{d}y(i,j) = I(i,j+1) - I(i,j-1)) \tag{2}$$

where I is the image involved and i, j represent the location of a pixel in image I.

3.2.4 *Recursive differencing*

This linear edge detector based on the Deriche filter uses the following recursive equations to obtain derivatives in x and y direction.

$$\mathrm{d}x(i,j) = I_{xp}(i,j) + I_{xm}(i,j) \tag{3}$$

$$I_{xp}(i,j) = a_{0p}I_{x0}(i,j) + a_{1p}I_{xo}(i-1,j) - b_1I_{xp}(i-1,j) - b_2I_{xp}(i-2,j) \tag{4}$$

$$I_{xm}(i,j) = a_{1m}I_{x0}(i+1,j) +_{2m} I_{x0}(i+2,j) - b_1I_{xm}(i+1,j) - b_2I_{xm}(i+2,j) \tag{5}$$

$$I_{x0}(i,j) = a_1(X_p(i,j) - X_m(i,j)) \tag{6}$$

$$X_p(i,j) = I(i,j-1) - b_1X_p(i,j-1) - b_2X_p(i,j-2) \tag{7}$$

$$X_m(i,j) = I(i,j+1) - b_1X_m(i,j+1) - b_2X_m(i,j+2) \tag{8}$$

In Equations (3)–(8), $i = 1, \ldots, S_x$, $j = 1, \ldots, S_y$, $a_{0p}, a_{1p}, a_{1m}, a_{2m}, b_2, b_1$ are operator parameters and S_x and S_y represent the image size. A similar recursive formulation is also utilised to obtain the gradient in y direction dy.

3.2.5 *Ratio of exponential weighted averages (ROEWA) operator*

Placing a step edge between two homogeneous regions degraded by speckle, this operator, intends to give an unbiased estimate of the ratio of the mean values of the two regions. Based on the ISEF filter which is realised as a cascade of two recursive filters, $h_1(x)$ (causal) and $h_2(x)$ (anticausal)

$$h(x) = c \cdot b^{|x|} \equiv h_1(x) * h_2(x) \tag{9}$$

where

$$h_1(x) = a \cdot b^x u(x) \quad h_2(x) = a \cdot b^{-x} u(-x) \tag{10}$$

where, $c = a/(2 - a)$, $b = 1 - a$, $0 < a < 1$ and $u(x)$ is the Heaviside function and $*$ denotes the convolution operator. The first derivative of the smoothing filter $h(x)$ is given by

$$f(x) = \frac{d}{dx} h(x). \tag{11}$$

For ratio-based implementation proposed in Fjørtoft *et al.* (1995)

$$f(x) = k \left[\frac{h_1(x)}{h_2(x)} \right] \quad k = \frac{\ln(1 - a)}{2 - a}. \tag{12}$$

For this case, horizontal and vertical components of the gradient of an image are given by

$$\Delta_x[(h(x)h(y)) \otimes I(i,j)] = f(x) * [h(y) \bullet I(i,j)] \tag{13}$$

$$\Delta_y[(h(x)h(y)) \otimes I(i,j)] = f(y) * [h(x) \bullet I(i,j)] \tag{14}$$

where $\otimes$ denotes bidimensional convolution, $*$ denotes convolution in the horizontal direction and $\bullet$ represents convolution in the vertical direction. The discrete realisations of h_1 and h_2 are obtained using the z-transform of the discrete versions of Equation (9)

$$H_1(z) = \frac{a}{1 - bz^{-1}}, \quad H_2(z) = \frac{a}{1 - bz}. \tag{15}$$

For discrete recursive implementation (one-dimensional), in terms of a spatial index n, this corresponds to the following simple recursions:

$$s_1(n) = a(e_1(n) - s_1(n-1)) + s_1(n-1) \quad n = 1, \ldots, N \tag{16}$$

$$s_2(n) = a(e_2(n) - s_2(n+1)) + s_2(n+1) \quad n = N, \ldots, 1 \tag{17}$$

where $e_1(n)$ and $e_2(n)$ are the inputs, and $s_1(n)$ and $s_2(n)$ are the outputs of h_1 and h_2 respectively. Adapting the edge detector to radar images, ratio $r(n)$ can be defined as

$$r(n) = \min \left\{ \frac{s_1(n)}{s_2(n)}, \frac{s_2(n)}{s_1(n)} \right\}. \tag{18}$$

It is important to note that according to formulation represented in Equation (18), a stronger edge response would have value near to zero. Further, the Equations (16) and (17) are realised both for x and y directions to obtain the gradients dx and dy.

3.3 *Feature matching*

Matching of the feature descriptors based on the ratio of the descriptor distance between the two best matches of a reference feature seems to be a reasonable choice for both SAR and optical images. The matching threshold is calculated as the ratio between Euclidean descriptor distance between the second-closest and the closest match of a feature. Recommended matching ratio (MR) is 0.6 but here we test the performance at MR 1.0. For this we utilise the rough registration parameters to compute an approximate match region for every match and filter out those matches where the corresponding feature does not lie within a user defined window size. The idea behind this scheme is to increase the number of feature matches as due to the MR threshold there would always be some correct matches left out in the process of keeping out most of the wrong matches. Rough registration parameters (rotation and translation) between multimodal images can be quickly estimated automatically using mutual information on downsampled images (Suri *et al.* 2009). As the BBF matching scheme utilised by the SIFT processing chain is not completely error free, we tend to get a significant number of bad matches. Therefore, we utilise RANSAC algorithm to filter all the obtained matches to obtain a subset of matches referred to as correct matches in further text.

In accordance to the discussion to this point, the main objectives of this article might be enumerated as:

(1) To compare SD, image recursive differencing (RD) and operator based on image ratio (RA) as local gradient estimation schemes to generate rotation and scale invariant feature descriptors for SAR image matching applications.

(2) Considering that normally remote sensing images might not necessarily have a big rotation difference, we also present a version of the SIFT operator which does not have the RI property. As aforementioned in Section 2.2, the SIFT operator estimates one or multiple dominant orientations around a keypoint according to which the descriptors are modified (basically estimated descriptors are rotated by the found dominant orientation(s) to make rotation invariant descriptors). However, with experimentation on a lot of datasets we have found that in multisensor/multitemporal images the orientations of the same features might be estimated to different values further resulting into different descriptors. Finally, these flawed descriptors lead to these correctly detected features not being matched in the matching phase. Practically in various applications, the remote sensing SAR images would generally have similar orientations so for these cases we do not require rotation invariant descriptors and there the presented approach becomes useful as we can increase the number of correct matches. The advantage of this version of the SIFT operator is realised by the significant increase in the number of matches (with similar accuracy) obtained within SAR images having complex acquisition differences (different sensor, incidence angle, speckle characterisitcs, orbit directions and various land covers).

4. Experimental datasets

For the described analysis, we have selected the following SAR image pairs (all 1000×1000 pixel amplitude images).

4.1 *Dataset 1*

Two E-SAR (Experimental SAR: an airborne SAR-sensor developed and used by DLR (Schreiber *et al.* 1999)) scenes acquired over Oberpfaffenhofen, Germany (Figure 2a).

4.2 *Dataset 2*

Multi-sensor registration, two scenes acquired using different sensors (Radarsat-1 and ERS-2, the European remote-sensing satellite) at a time difference of 15 days, featuring a rural land cover class with well-defined features. Here, both scenes were acquired at different orbiting directions (Figure 2b).

4.3 *Dataset 3*

One of the E-SAR images of dataset-1 is matched with a TerraSAR-X image. The chief characteristics of this selected dataset can be summarised as (Figure 2c):

(1) Different sensor nature (airborne and spaceborne).
(2) Time difference of 4 years.
(3) Different incidence angle.
(4) High-resolution imagery acquired over semi urban land cover.

5. Experimental results

For a comparative evaluation, we show the gradient images and pixel orientation histogram (Figure 3) obtained by the three gradient estimation schemes for the E-SAR image shown in Figure 2(a).

The magnitude images representing edge strengths are linearly mapped from their respective value range to the range [0,255] (for display purposes). The edge orientations are linearly mapped from the initial range $[-\pi, +\pi]$ to $[0, 255]$ to plot a histogram of edge orientations of all the image pixels. The key observation is that for difference-based gradient estimation schemes, the feature descriptors have orientations assigned in the complete viable range of $[-\pi, +\pi]$, but that is not the case for the ROEWA operator where all the image pixels obtain orientations from the range $[0, \pi/2]$. This behaviour is attributed to the fact that none of the x and y gradient has any sign associated which decides an edge pixels orientation. This narrow range can influence the detector's performance for images having a rotation difference.

5.1 *Dataset-1*

To continue with the analysis we run the same SIFT-OCT detection and feature matching scheme and utilise the three different gradient estimators for matching images in Figure 2(a). For one matching scenario, we also rotate the reference image by 90° to
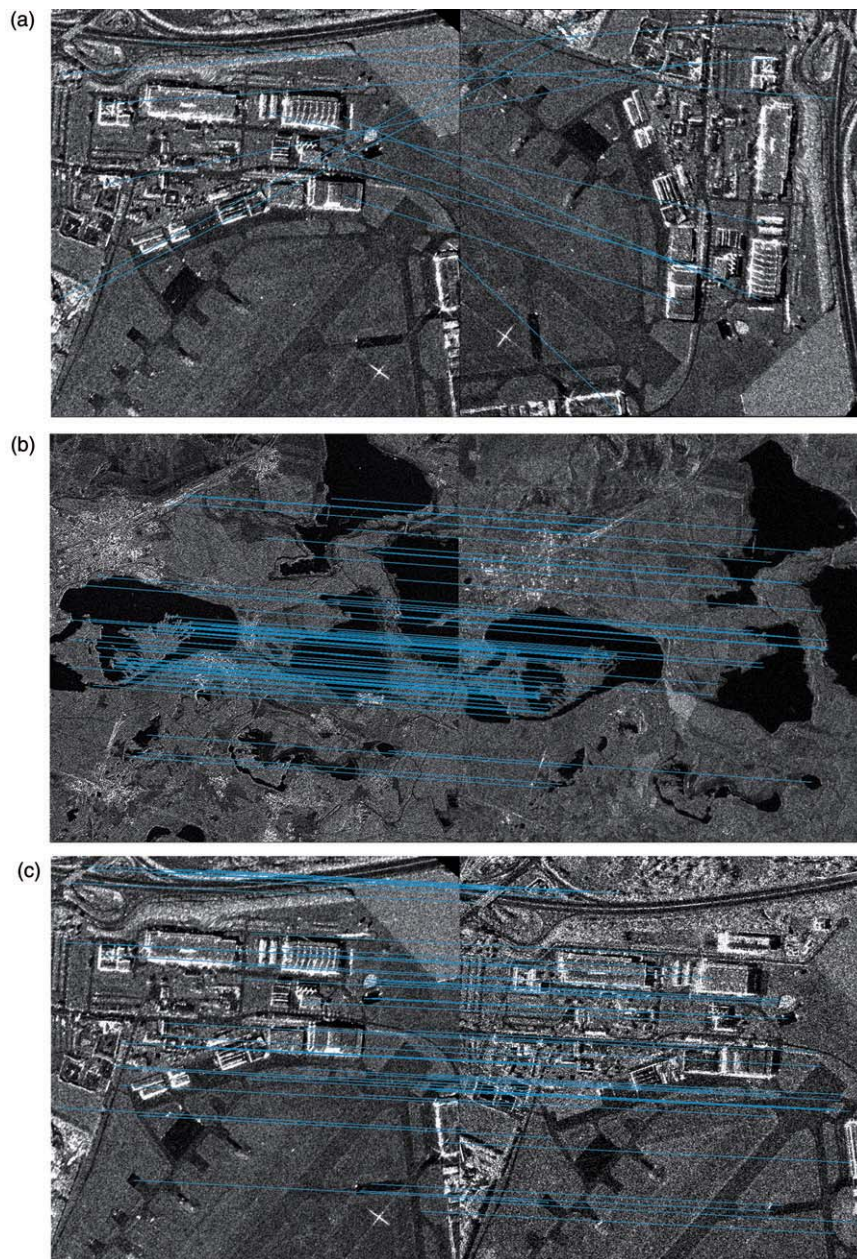
Figure 2. (a) E-SAR images acquired at 1 h interval and 90° rotation difference (match results: RA descriptors, MR 0.4). (b) Radarsat (left) and ERS (right) images. (c) E-SAR image (left) and TerraSAR-X image (right).

analyse the gradient estimation schemes. Due to the same detection scheme, the number of detected features (approximately 3500 features) and matching accuracy levels (sub pixel) are similar in all the SIFT versions evaluated here. However, we compare the three techniques on the basis of their correct matches (percentage and number). The observations with three
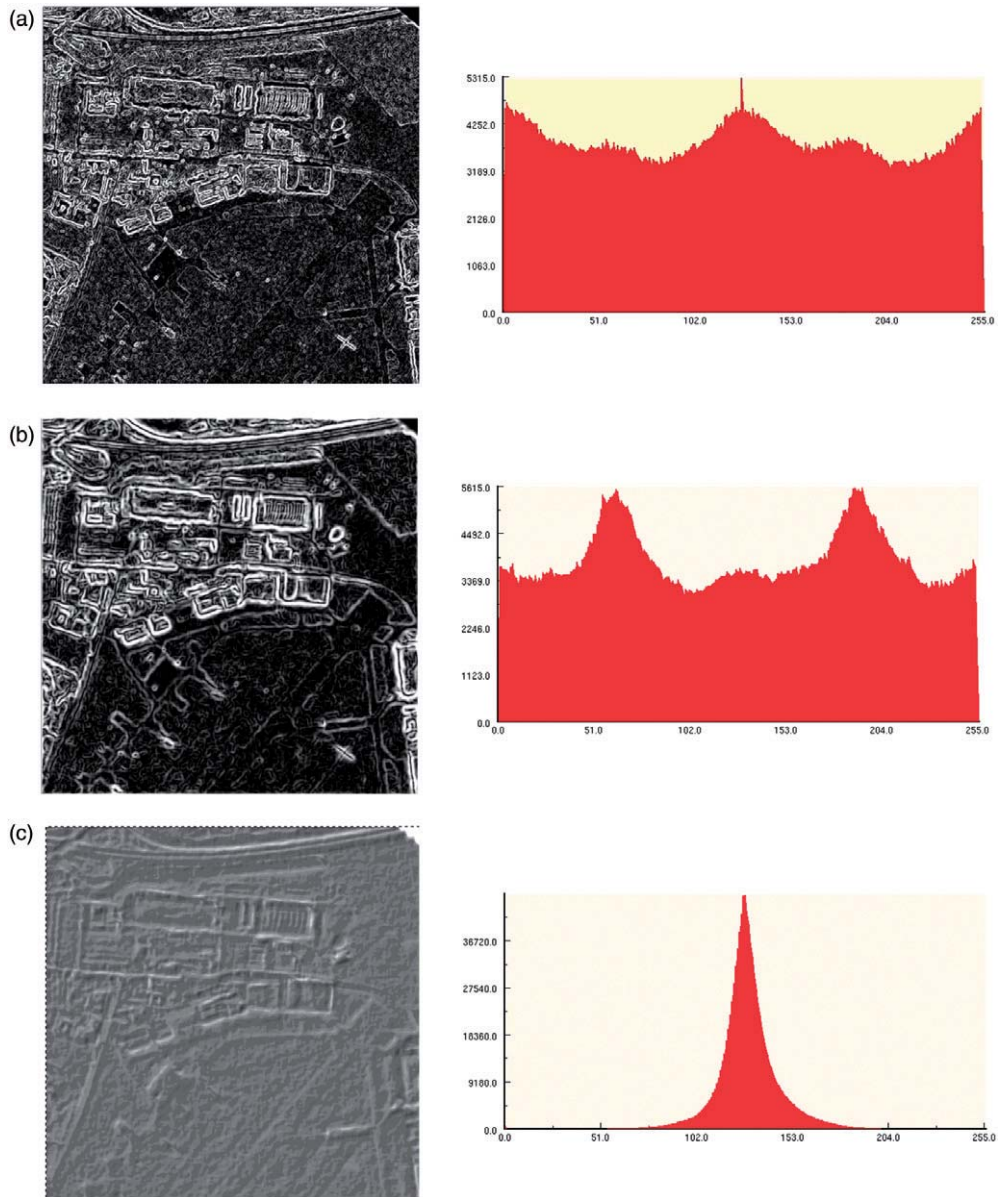
Figure 3. Edge Magnitude image with the corresponding edge orientation histogram (a) SD, (b) RD (c) RA.

different MR values (0.4, 0.6 and 0.8) are depicted in Table 1. Overall, it is seen that the simple differencing performs almost at par with the two more conceptually sophisticated techniques. Results indicate that the ROEWA-based descriptor formation (RA) scheme looses the maximum number of correct matches on rotating the reference image by 90° and this is due to the narrow range of feature orientations (Figure 3). This is not encountered in the gradient estimation schemes based on image differencing (SD and RD).

Table 1. SIFT performance in terms of number of correct matches (NOM) and percentage (%) for dataset-1.

		Number of rotation			Reference image rotated by 90°		
	RI	Yes			No		
Gradient type	MR	0.4	0.6	0.8	0.4	0.6	0.8
SD	NOM (%)	89	289	539	67	263	512
		97	91	77	93	88	77
RD	NOM (%)	102	328	569	84	290	540
		95	91	66	93	89	70
RA	NOM (%)	64	247	506	12	115	364
		100	90	81	100	95	79

Notes: Rotation Invariance (RI), Matching Ratio (MR), local feature gradients estimated through Simple Differencing (SD), Recursive Differencing (RD) and image ratio (RA).

Table 2. Matching results in terms of number of correct matches (NOM) and percentage (%) for dataset-2 using different gradient estimation techniques.

	RI	Yes			No		
Gradient type	MR	0.6	0.8	1.0	0.6	0.8	1.0
SD	NOM (%)	10	34	56	30	76	85
		83	33	100	83	57	100
RD	NOM (%)	13	33	47	29	67	77
		62	18	100	78	34	100
RA	NOM (%)	7	23	47	26	63	74
		88	30	100	70	38	100

Notes: Rotation Invariance (RI), Matching Ratio (MR), local feature gradients estimated through Simple Differencing (SD), Recursive Differencing (RD) and image ratio (RA).

5.2 Dataset-2

Results for this dataset are presented in Table 2. Further, we test the SIFT variant without RI property. This is achieved by simply skipping the orientation assignment step in the SIFT processing chain. Here, we detect approximately 2600 and 2000 features in ERS-2 and Radarsat images, respectively. At MR 1.0, the maximum number of detections and accuracy is expected as we utilise rough registration parameters to compute an approximate match region for every feature and filter out those matches where corresponding feature is not within a predefined window size (5 pixels). Images in dataset-2 offer a much more complicated matching scenario as compared to dataset-1. This is reflected in the reduced number of matches obtained for all the six operator variants. In general, it is observed that gradients estimated through SD perform at par (or marginally better) with the more sophisticated local gradient estimation techniques (RD and RA). Here, the SIFT version without the RI characteristic outperforms the SIFT version with the RI competence (only in terms of the number of matches not in

Table 3. Matching results in terms of number of correct matches (NOM) and percentage (%) for dataset-3 using different gradient estimation techniques.

	RI	Yes			No		
Gradient type	MR	0.6	0.8	1.0	0.6	0.8	1.0
SD	NOM (%)	0	3	39	6	9	49
		–	3	78	60	9	60
RD	NOM (%)	1	3	63	4	15	64
		25	1	63	25	6	64
RA	NOM (%)	0	4	22	0	14	49
		–	5	59	–	9	60

Notes: Rotation Invariance (RI), Matching Ratio (MR), local feature gradients estimated through Simple Differencing (SD), Recursive Differencing (RD) and image ratio (RA).

terms of percentage accuracy). The main idea of removing the RI characteristic is the fact that we already know the images have similar orientations.

Further, it looks likely, that determining orientations of certain conjugate features might not yield similar results as the process is heavily dependent on feature neighbourhood, which might be very different in multimodal SAR images. For all different match results, we obtain a match consistency of 1–2 pixels in both x and y direction for 12.5-m pixel spacing images.

5.3 *Dataset-3*

Here, we have 2800 and 3500 features detected in TerraSAR-X and ESAR images, respectively. The number of correct matches for all the variants has gone down drastically (Table 3), further asserting the adverse influence of present complex image acquisition differences. Again, the SIFT version without the RI characteristic has done much better (both in terms of absolute number and percentage of correct matches) compared to its other counterpart. For MR 1.0, we initially utilised a 16 pixel window and then used RANSAC to filter the matches using a threshold of 5 pixels; therefore some of the matches are rejected as bad matches. Using approximate registration parameters significantly improves the number of correct matches for these very different SAR images. This highlights the fact that the SIFT operator has excellent same feature detection capability in multimodal SAR imagery, but very few of those are generally reported as successful matches. The average match consistency after building a first-order polynomial (only taking the on ground matches) is in the range 1–3 pixels for both x and y direction at 1-m pixel spacing.

6. Conclusions

In this article, we have presented and analysed certain modifications in the SIFT operator processing chain for effective SAR image matching. On basis of our previous works and in accordance to the laid down objectives, the following can be concluded:

(1) For SIFT feature detection, skipping the features detected at the first octave improves the operator performance both in terms of accuracy and processing time.

(2) For SIFT feature identification, the gradients estimated using SD seems to be almost as good as the two analysed more sophisticated techniques. There is no significant difference within the three compared techniques on basis of correct matches (percentage and absolute number).

(3) The number or matches can be increased by using the SIFT processing chain without the RI characteristic. As remote sensing SAR images utilised in certain applications like change detection can be expected to have similar orientation, we expect the SIFT operator without RI to increase the potential number of matches in most SAR image matching applications.

(4) The original SIFT processing chain does not make full use of its potential in matching similar features in SAR images, meaning that similar features are detected in both scenes but not matched by using BBF matching alone. This assertion is confirmed in all the three tested datasets but most profoundly for dataset-3, which presents a challenging SAR image matching scenario. As presented in this article, the SIFT capability to match conjugate point features in multimodal/multisensor SAR images can be enhanced by making use of approximate registration parameters.

The obtained results indicate that by incorporating the above suggestions, the originally proposed SIFT processing chain can be effectively refined for fast and robust multimodal SAR image matching applications. Concluding in terms of general applicability of the presented approach for multimodal SAR imagery, it has been found that SIFT operator performs optimally for same sensor images and its performance is influenced with increasing image differences such as incidence angle, orbit directions, time difference within images, etc. It needs a mention that in plain areas good operator performance has also been reported for images acquired by different sensors using different orbit directions and incidence angles. However, registration in mountainous areas with different orbit directions and also in urban cities with different flight paths still remains a challenge for such a point feature-based registration strategy.

References

Beis, J. and Lowe, D.G., 1997. Shape indexing using approximate nearest neighbour search in high dimensional spaces, *In*: *Proceedings of IEEE conference on computer vision pattern recognition*, June 1997, San Juan, Puerto Rico, 1000–1006.

Bentoutou, Y., *et al.*, 2005. An automatic image registration for applications in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 43, 2127–2137.

Bovik, A.C., 1988. On detecting edges in speckle imagery. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36 (10), 1618–1627.

Canny, J.F., 1983. Finding edges and lines in images, Massachusetts Institute of Technology Technical Report No. 720, Cambridge.

Deriche, R., 1987. Using Canny's criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 1 (2), 167–187.

Dell'Acqua, F., Gamba, P., and Lisini, G., 2004. Coregistration of multiangle fine spatial resolution SAR images. *IEEE Geoscience and Remote Sensing Letters*, 1 (4), 237–241.

Eineder, M., *et al.*, 2005. TerraSAR-X SAR products and processing algorithms. *In*: *Proceedings of IEEE international geoscience and remote sensing symposium*, 7, July 2005, Seoul, Korea, 4870–4873.

Fjørtoft, R., *et al.*, 1995. Edge detection in radar images using recursive filters. *In*: *Proceedings of ACCV'95*, 3, 5–8 December 1995, Singapore, 87–91.

Krieger, G., *et al.*, 2007. The TanDEM-X mission: a satellite formation for high resolution SAR interferometry. *In: Proceedings of international astronautical congress*, September 2007, Hyderabad, India.

Li, H., Manjunath, B.S., and Mitra, S.K., 1995. A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, 4 (3), 320–334.

Lowe, D.G., 2004. Distinctive image features from scale-invariant features. *International Journal of Computer Vision*, 60 (2), 91–110.

Mikolajczyk, K. and Schmid, C., 2005. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27, 1615–1630.

Paillau, P., 1997. Detecting step edges in noisy SAR images, a new linear operator. *IEEE Transactions on Geoscience and Remote Sensing*, 35, 191–196.

Schreiber, R., *et al.*, 1999. Overview of interferometric data acquisition and processing modes of the experimental airborne SAR system of DLR. *In*: *Proceedings of IEEE international geoscience and remote sensing symposium*, July 1999, Hamburg, Germany, 35–38.

Schwind, P., *et al.*, 2010. Applicability of the SIFT operator for geometrical SAR image registration. *International Journal of Remote Sensing*, 31 (8), 1959–1980.

Shen, J. and Castan, S., 1986. An optimal linear operator for edge detection. *In*: *Proceedings of IEEE international conference on computer vision and pattern recognition*, Miami, FL, 109–114.

Shen, J. and Castan, S., 1992. An optimal linear operator for step edge detection. *CVGIP: Graphical Models and Image Processing*, 54 (12), 112–133.

Suri, S., *et al.*, 2009. Combining mutual information and scale invariant feature transform for fast and robust multisensor SAR image registration. *In: Proceedings of 75th ASPRS annual conference*, March 2009, Baltimore, MD, USA.

Touzi, R., Lopes, A., and Bousquet, P., 1988. A statistical and geometrical edge detector for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 26, 764–773.

Zitová, B. and Flusser, J., 2003. Image registration methods: a survey. *Image and Vision Computing*, 21 (11), 977–1000.