

COMBINING MUTUAL INFORMATION AND SCALE INVARIANT FEATURE TRANSFORM FOR FAST AND ROBUST MULTISENSOR SAR IMAGE REGISTRATION

Sahil Suri, PhD Candidate
Peter Schwind, Team Member
Peter Reinartz, Group Leader
Johannes Uhl, Intern Student
Remote Sensing Technology Institute (IMF)
German Aerospace Center (DLR)
82234 Wessling, Germany
Sahil.Suri@dlr.de
Peter.Schwind@dlr.de
Peter.Reinartz@dlr.de
Johannes.Uhl@dlr.de

ABSTRACT

The Scale Invariant Feature Transform (SIFT) operator's success for computer vision applications makes it an attractive solution for the intricate feature based SAR image registration problem. For SAR images, SIFT feature matching results into lot of false alarms. To overcome the mentioned problem, we propose to use mutual information (MI) along with the SIFT operator for SAR image registration and matching applications. MI is an established multimodal registration similarity metric and has the capability to quickly estimate rough registration parameters from down-sampled images. The rough image registration parameters obtained using MI can be introduced for conjugate feature selection during the SIFT matching phase. Introduction of MI to the SIFT processing chain not only reduces the number of false alarms drastically but also helps to increase the number of matches as the operator detection and matching thresholds can be relaxed, relying on the available mutual information estimate. Further, the matching consistency of the SIFT matches especially for SAR images with various acquisition differences might not be up to the desired levels. To tackle the observed phenomenon, MI can further be utilized to refine the SIFT matches and to bring the matching consistency within desirable limits. We present our analysis based on multisensor, multitemporal and different view point SAR images acquired over plain and semi urban areas. The proposed registration methodology shows tremendous potential to become a fast and robust alternative for geometric SAR image registration as subpixel registration consistency has been achieved for diverse natured datasets.

INTRODUCTION

With the increasing availability and rapidly improving spatial resolution of remote sensing SAR images from latest and future satellites like TerraSAR-X (Roth 2003) and Tandem-X (Krieger et al. 2007), the applicability of SAR images to various on ground applications is bound to experience a tremendous boost. Moreover, the advantage of being acquired sunlight and weather independent make SAR images ideal for crisis and disaster mitigation, where data is required instantly and bad weather conditions might prohibit the use of optical sensors. Already with the present state of technology, SAR imagery has been found useful for diverse applications like DEM generation (Dupont et al. 1997), image fusion (Moghaddamet et al. 2002), soil moisture estimation (Hegar-Masclé et al. 2002), traffic related studies (Palubinskas et al. 2005), change detection (Bovolo and Bruzzone 2005) and many more (Eineder et al. 2005). Prior to most of these enlisted remote sensing applications, images need to be registered with sufficient accuracy depending upon the application demands (usually sub-pixel level).

Image registration refers to the task of aligning two or more images acquired at different times, from different sensors or from different view points. Mathematically, the problem of registering an input image (I) to a reference image (R) can be expressed as (Brown, 1992):

$$I_R(x, y) = g(I_I(T(x, y))) \quad (1)$$

where T is a transformation function which maps two spatial coordinates x and y , to the new spatial coordinates x' and y' (Equation 2) and g is a one dimensional (1D) intensity or radiometric interpolation function,

$$(x', y') = T(x, y) \quad (2)$$

The main objective of a registration process is to estimate the spatial transformation T and depending upon the method of resolution, image registration task can be divided into:

- i. Feature based techniques.
- ii. Intensity based techniques.

Feature based techniques depend on accurate identification of features or objects that describe important landmarks, sharp edges or shapes, which however may often be difficult to extract. The task of determining the best spatial transformation for the registration of the images can be broken down to feature detection and matching, transformation model estimation, image resampling and registration quality assessment.

Alternatively, in intensity based techniques, images are registered based on a relation between pixel intensity values of two images. In this method of resolution, the problem of registration is generally mapped as an optimization problem. Where the spatial transformation function T is the argument of the optimum of some similarity metric S , applied to reference image I_R and transformed input image I_{TI} . This can be expressed as:

$$T = \arg(\text{opt}(S(I_R, I_{TI}))) \quad (3)$$

Both the above mentioned contrasting techniques have their advantages and disadvantages. For example, feature based techniques are generally faster and might be better suited for multi-temporal cases where scenes have undergone some great changes but fixed and permanent features are available for extraction and matching. Intensity based techniques might become sensitive to the changes in intensity values, introduced for instance by noise, by varying the illumination, and/or by using different sensor types. Intensity based registration techniques definitely have an edge over the feature based techniques for scenarios where feature detection and matching becomes difficult (e.g. SAR-Optical registration scenario).

In the past, both intensity and feature based registration approaches have been utilized successfully for SAR image registration. Specifically, mutual information can be used to register both multisensor SAR and SAR-optical image pairs (Hua et al. 2003, Chen 2003). Feature based SAR image registration is a difficult task due to the presence of multiplicative speckle influence (Touzi et al. 1988, Bovik 1988). Still, some work in the field of feature based SAR image registration can be found in the remote sensing literature (Li et al., 1995; Borghys et al., 2001). Here, we evaluate a processing chain combing the advantages of both feature and intensity based techniques for fast and robust registration of multimodal SAR images. To detect and match features in SAR images the SIFT (Scale Invariant Feature Transform) operator has been utilized. The SIFT operator is a feature detector introduced in 1999 and improved in 2004 by Lowe (1999, 2004). Since its introduction it has proven its effectiveness for numerous applications especially in the field of computer vision (Lowe 1999, Se et al., 2001; Brown and Lowe, 2003). SIFT operator descriptors have been accredited to be very distinctive in the field of computer vision and therefore their evaluation in the field of remote sensing image registration becomes an interesting application. In this paper, we incorporate mutual information both before and after the SIFT processing chain. The advantage of using MI before the SIFT processing is to increase the number of matches and reduce the number of false alarms based on the rough registration parameters obtained using MI based registration of down-sampled images. Further, we also show refinement of SIFT matches using MI when desired registration consistency levels are not reached.

In the following sections we explain mutual information, a slightly modified SIFT processing chain and finally the results and conclusions drawn after testing the proposed registration chain for multisensor SAR images.

MUTUAL INFORMATION

Mutual information has evolved from the field of information theory. MI describes a statistical dependence between two random variables (e.g. A and B) expressed in terms of variable entropies. In

case Shannon entropy (additive in nature) is selected to represent the individual variable information, mutual information between two variable A and B is defined as (Wachowiak, 2003)

$$MI(A, B) = H(A) + H(B) - H(A, B) \quad (4)$$

Above, $H(A)$ and $H(B)$ are the Shannon entropies of A and B respectively, $H(A, B)$ is the joint entropy of B and A. Registration of two images A and B is based on maximization of $MI(A, B)$ (Equation 4). The marginal entropies and the joint entropy can be computed from the estimated joint histogram according to formulations described in (Chen et al., 2003). A vigilant observation of the results obtained for registration of Landsat images by Cole Rhodes et al. (2003) highlights the capability of MI to obtain rough registration parameters even from down sampled images. The idea here is to quickly estimate rough registration parameters using down sampled images and then select only those SIFT matches where conjugate features are within a user defined threshold from the approximated rough registration parameters.

SCALE INVARIANT FEATURE TRANSFORM

In this section, we briefly review the entire SIFT operator processing chain to detect and match features in remote sensing images. The entire processing can be broken up into three main steps of feature detection, descriptor formation and matching. All the three components are being summarized here briefly. For more conceptual and implementation details, interested readers are referred to the original work by David Lowe (Lowe, 2004).

Feature Detection

The process starts with keypoint detection (Lowe, 2004). 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 :

$$D(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (5)$$

Next, extrema are detected in the DoG images by comparing every pixel to the eight neighboring pixels and the nine pixels in the scales above and below. If a pixel value is larger or smaller than all of its neighbors, it is accepted as a preliminary keypoint candidate. In the following keypoint localization phase all keypoint locations are interpolated with subpixel accuracy, using an iterative method developed by Brown and Lowe (2002). After the interpolation, two additional checks are performed to remove unstable keypoints. First the value at the extremum $D(\hat{x})$ is computed. Keypoints with a value below a certain threshold are eliminated thereby removing points with low contrast. Then, points lying on edges are removed, making use of a Hessian matrix H computed at the keypoint location:

$$H = \begin{pmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{pmatrix} \quad (6)$$

The derivatives D_{xx} , D_{xy} and D_{yy} are determined by calculating the differences between neighboring points. Rather than solving the Eigen value problem, keypoints are selected from the trace $Tr(H)$ and the determinant $Det(H)$ by requiring

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (7)$$

Lowe suggests using a value of 10 for r (Lowe 2004)

Feature Descriptor Formation

A SIFT feature detector is a 128 length vector containing information about local gradients around the detected feature. A 16×16 window around the detected feature is selected to compute a histogram of gradient location and orientation. The selected window for descriptor formation is broken into 4×4 location grids (16 in number) and the gradient angles are categorized into 8 orientations. After that, the gradient magnitudes surrounding the keypoint are weighted by a Gaussian window to weaken the influence of gradients far away from the keypoint. The size of the Gaussian filter is set to half the size of the descriptor window, which itself has a size of 16×16 . By doing this, the descriptor is robust to small shifts, since gradients close to the center have a larger impact. To make the descriptor invariant to linear brightness changes, the elements of the vector are normalized to unit length. The

influence of non-linear brightness changes (e.g. illumination changes that affect some surfaces more than others) is reduced by introducing thresholds in the vector elements. After that, all the values are normalized once again to unit length.

Feature Matching

Even though the main SIFT operator objective is to detect stable keypoints, Lowe (1999) also proposed a matching strategy for the keypoints. To compare two descriptors with each other, the Euclidean distance of the descriptor vectors is calculated. However, simply matching the keypoints with the smallest Euclidean distance might not produce adequate results if no additional checks are applied. Therefore instead, the two closest matches in the other image are determined for every keypoint. If the Euclidean distance of the second-closest match is smaller than 0.8 times the distance of the closest match, the point is accepted as a match. Since comparing the distance of all keypoints with each other is expensive, an approximate algorithm called Best-Bin-First (BBF) (Beis and Lowe, 1997) is used. The matching results presented in this paper are based on the source code provided on David Mount's homepage (<http://www.cs.umd.edu/~mount/>). For more information about approximate nearest neighbor algorithm and technique the interested readers are referred to list of publications mentioned on the author's homepage.

EXPERIMENTAL DATASETS

The objective here is to present an evaluation of mutual information and the SIFT operator combined for registration of SAR images acquired:

- i. Using different sensors (multimodal registration)
- ii. With different incidence angles and sensor orbiting directions (view point registration)
- iii. At different times (temporal registration)
- iv. Over different scene conditions (semi urban and rural areas)

Dataset 1: In this dataset we test two scenes acquired using different sensors (Radarsat-1 and ERS-2 (European Remote Sensing Satellite)) at a time difference of 14/15 days, featuring a rural land cover class. The chosen scene has prominently recognizable features (several lakes of the Lausitzer Seenkette near Senftenberg, Germany) that might lead to good detector performance. To evaluate the registration chain performance for scenes with different aspect angles, the ERS-2 image was matched once with an ascending Radarsat-1 image (dataset 1a) and once with a descending Radarsat-1 image (dataset 1b) (Table 1). In both the scene pairs the incidence angle of the sensor differs by approximately 20° . The scenes were initially registered using a combination of manual and intensity based technique and so for evaluation purposes the ERS scene has been subjected to 5° rotation, 10 pixels x translation and -5 pixels y translation. The transformed ERS image along with the two Radarsat scenes can be visualized in Figure 1.

Dataset-2: This dataset offers a challenging matching scenario with images acquired over semi urban area having significant differences in date of acquisition (4 years), incidence angle, and sensor geometries. The images have been taken over Oberpfaffenhofen (near Munich) by ESAR sensor and the latest German high resolution satellite, TerraSAR-X (Roth 2003). The acronym ESAR stands for Experimental Synthetic Aperture Radar and is an airborne SAR-sensor developed and used by the German Aerospace Center DLR (Schreiber et al. 1999). The different sensor geometries are expected to have a strong influence on the features in the urban areas and thus this selection of dataset offers an interesting evaluation of the SIFT operator capability to match and detect features in SAR images with complex acquisition differences. The details of the imagery selected are tabulated in Table 2 and the changes in the sub urban area in form of new constructions and deconstructed buildings in 4 years time span can be observed in Figure 2. In this case the images were not pre-registered and the spatial deformation has been modeled by two translations in x and y direction.

PROPOSED REGISTRATION METHODOLOGY

The hybrid registration scheme experimented and analyzed in this paper can be visualized in Figure 3. To counter the speckle influence the images are first subjected to image smoothing and also the features detected from the first octave of the scale space are not considered for further matching. In our previous work we demonstrated that filtering by the ISEF filter and skipping the features detected at

highest octave lead to comparable (sometimes better) and significantly faster SIFT operator performance for SAR image matching and registration applications (Schwind et al., in press).

Table 1: Details of the ERS-2 and Radarsat Imagery utilized for Dataset 1

	ERS-2	Radarsat-1	Radarsat-1
Mode	SAR-Image Mode	Standard Beam (mode 6)	Standard Beam (mode 6)
Radar Frequency	5.3 Ghz	5.3 Ghz	5.3 Ghz
Pixel Spacing	12.5m	12.5m	12.5m
Bits/Pixel	16 bit	16 bit	16 bit
Incidence Angle	22.97°	43°	43°
Data of Acquisition	20-April-06	05-April-06	06-April-06
Orbit	Ascending	Ascending	Descending
Image Size	1084 x 1085	1000 x 1000	1000 x 1000

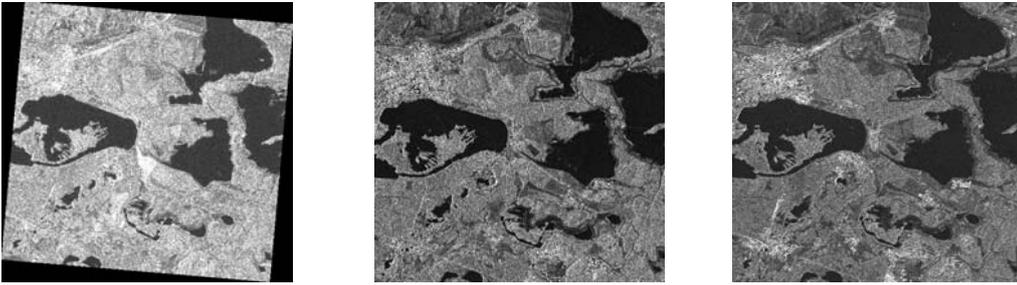


Figure 1: The images from Dataset 1 (left) Transformed ERS-2 (Ascending) (center) Radarsat-1 (Ascending) (right) Radarsat-1 (Descending) acquired over the lakes of Lausitzer SeenKette near Seftenberg, Germany

Table 2: Details of the ESAR and TerraSAR-X Imagery utilized for Dataset 2

	ESAR(Ref)	TerraSAR-X(Inp)
Mode	Multi-look image (4 looks)	High Resolution spot light
Radar Frequency	9.6 Ghz	9.6 Ghz
Pixel Spacing	1m	1m
Bits/Pixel	8 bit	16 bit
Incidence Angle	24.78°	35.14°
Data of Acquisition	20-April-04	20-May-2008
Orbit	-	Ascending
Image Size	1000 x 1000	1000 x 1000

Infinite Symmetric Exponential Filter (ISEF) filter has been utilized to reduce the speckle influence, the SAR images are preprocessed using this smoothing filter before the scale space pyramid computation. The ISEF filter was shown to deliver good results for edge detection in SAR images (Fjortoft et al. 1995). ISEF was proposed by Shen and Castan (1992) as an "optimal low-pass filter as a preparation for edge detection". Shen and Castan illustrated that an increase in Gaussian filter size is useful to reduce noise influence but the increased size has an adverse effect on edge localization. To overcome this problem, the ISEF filter with an infinite window size and desired sharpness at the window centre was proposed. ISEF is mathematically expressed as

$$f(x) = \frac{P}{2} \cdot \exp^{-P \cdot |x|} \quad (8)$$

For the two dimensional case, an efficient, recursive function can be used (Shen and Castan 1992). A sample influence of utilizing ISEF filtering on a SAR image prior to SIFT feature detection can be visualized in Figure 4. The number of features detected is reduced significantly, leading to faster execution times (approx 20 sec for 1000x1000 pixel images on an Intel P4 Xeon machine) for the

processing chain as fewer descriptors are created, resulting in a faster matching process (Schwind et al., in press).



Figure 2: The images from Dataset 2 (left) ESAR (air borne) (right) TerraSAR-X acquired over Oberpfaffenhofen near Munich in Germany. The images have approximately 4 years of acquisition difference.

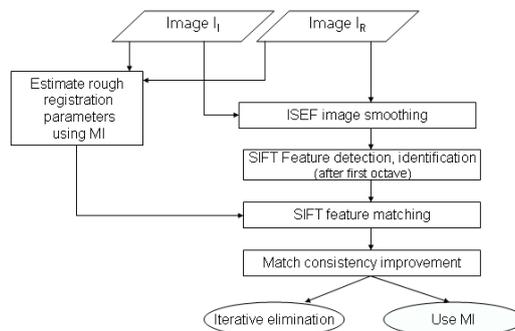


Figure 3: The proposed SAR image registration methodology combining the advantages of intensity based and feature based techniques

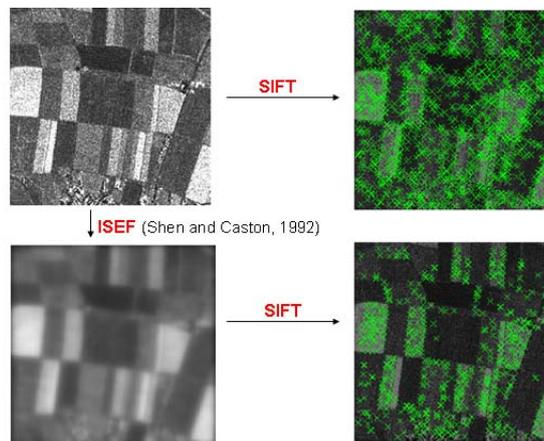


Figure 4: Influence of the ISEF smoothing filter on SIFT operator detection for a SAR image

Test Conditions: We analyze here the SIFT operator matching scheme performance both with and without using the rough registration parameters in the SIFT processing chain. The matches obtained at matching thresholds of 0.6, 0.8 and 1.0 (section 3.3) have been reported and compared for all the 3 analyzed datasets. As explained earlier, the matching threshold is calculated as the ratio between Euclidean descriptor distance between the second-closest and the closest match of a feature. At matching ratio 1.0 we utilize the rough registration parameters to compute an approximate match region for every match and filter out those matches where corresponding feature does not lie with in a user defined window size (16 pixels). In general, the SIFT operator matching scheme leads to large number of false alarms as the matching thresholds are relaxed and thus some kind of a filtering to

remove the outliers is absolutely mandatory. To remove the outliers automatically we use a simple iterative approach. Using the residuals from an initial first order builds polynomial using all the matched points, the most deviating point from the polynomial is removed and a new polynomial is computed. This process is iterated until all returned residuals are smaller than twice the standard deviation of all the points. As this iterative procedure is deterministic, it might fail to produce results in case the number of outliers is very large. To remove the outliers automatically, methods like RANSAC (Fischler and Bolles 1981) or the Hough transform (Hough 1962) can also be utilized. Here, we presented results using the iterative outlier elimination approach and for cases of its failure the RANSAC algorithm (robust but not deterministic) has been utilized for the elucidated task.

Finally, depending upon the number of matches/match consistency and application demands it can be further decided to refine each of the matches individually by using MI locally around the matched features (chip matching technique). In the presented paper the images have been smoothed by ISEF filter. Though ISEF filter has been proven to produce good edge localization, still some influence might be propagated to the final matching results which might not be acceptable for certain accuracy critical applications. Therefore, some kind of local refinement of the final matched might be inevitable for certain applications. In the presented scenario, the performance of intensity based metrics like MI is expected to be good considering the fact that normally features detected by the SIFT operator would offer high entropy neighborhoods.

EXPERIMENTAL RESULTS

In this section, we discuss and highlight all the registration results obtained for the datasets detailed in the previous section.

Results from Dataset 1a

The chief characteristics of this selected dataset can be summarized as:

- i. Different sensor images
- ii. Incidence angle difference of 20°
- iii. Same sensor orbiting directions

First we present the performance of MI to register the two images using intensity based registration techniques. A sample MI performance in a multiresolution optimization framework has been tabulated in Table 3. The original images (Figure 1) can be compressed using a simple block mean filter. For all the experiments reported in this paper we have compressed the images to one-fourth of their original resolution to initiate the intensity based process. The optimization process has been initiated with the seed $(-3^\circ, 0, 0)$ and for mutual information computations, a joint histogram of bin size 64 has been estimated using the GPVE technique (Suri and Reinartz, 2008). In Table 3 we also provide the metric consistency measure (Holden et al., 2000) normally utilized for intensity based registration processes in absence of ground truth measurements. Normally values near to zero of the registration consistency measure (RC) represents a good metric performance. Important point to be kept in mind is that the value of the consistency measure might not be directly related to the ground level accuracy of registration achieved.

Table 3: Intensity based registration of ERS and Radarsat image pairs of dataset 1a

Level	ERS to RS	RS to ERS	RC	TAT (sec)
2	$(-4.99^\circ, -1.78, 1.20)$	$(5.05^\circ, 2.02, -1.19)$	0.1	200
1	$(-4.96^\circ, -4.25, 3.46)$	$(4.97^\circ, 4.68, -3.11)$	0.09	540
0	$(-4.98^\circ, -9.06, 6.20)$	$(5.03^\circ, 9.84, -6.02)$	0.39	1500

A point to be noted is that the registration parameters achieved by mutual information for coarser resolution images (Level 2) are strongly correlated to the parameters obtained for the original resolution images (Level 0). For example, registration parameters obtained for MI for level 2 images when multiplied by the scaling factor of 4 (only for translation parameters, rotation is scale invariant) give a fair estimate of the registration parameters obtained at the original resolution. Further, the turn around time listed for every resolution level individually (two way optimization) highlights the fact that a lot of computational time is invested with not much of improvement in registration parameters while utilizing the multiresolution framework of intensity based techniques.

The approximate registration parameters obtained from level 2 images can be utilized to significantly enhance the SIFT operator matching performance for SAR images. The SIFT matching

results for the matching thresholds of 0.6, 0.8 (without using the rough registration parameters) and 1.0 (while using the rough registration parameters) for dataset 1a are tabulated in Table 4. As expected it can be clearly seen that the number of false matches increase significantly with the matching threshold relaxed from 0.6 to 0.8. Even at the matching ratio of 0.6 more than 50% of the found matches were removed to reach the sub pixel match consistency. On the other hand, using the rough registration parameters obtained through MI and utilizing a matching threshold of 1.0 the number of matches increased to 114 with 2 to 3 pixel match consistency in both x and y directions.

Table 4: SIFT matching scheme performance for Dataset 1a.

Matching Ratio	0.6	0.8	1.0
Matches Found by SIFT	35	249	114
Standard Deviation X	161.96	254.47	2.69
Standard Deviation Y	142.43	269.55	2.66
Filtered Matches	15	21	51
Standard Deviation X	0.61	0.68	0.96
Standard Deviation Y	0.23	0.26	0.51

It is also observed that the number of matches have increased considerably without any significant blunders highlighting the SIFT operator's capability to detect similar features in images with various acquisition differences. For this particular dataset, the SIFT operator variant with the standard matching procedure with a threshold of 0.6 could match 15 points with a subpixel consistency and more matches with similar consistency were found for the matching threshold of 1.0 (assisted by rough registration parameters). Further, these matches can be refined using MI locally around the matched point but as the match consistency is well within tolerance limits we have not performed the fine matching step for this dataset. The filtered matches for this scenario using the threshold of 1.0 can be visualized in Figure 5.

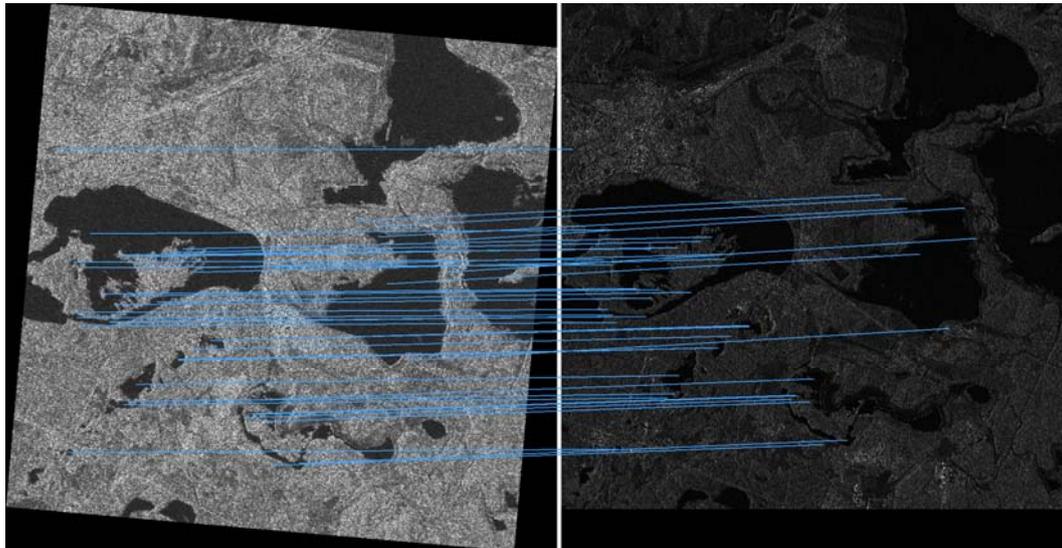


Figure 5: SIFT features matched between transformed ERS (left) and Radarsat-1 images of dataset 1a. 51 SIFT features were matched with a subpixel consistency level using a matching threshold of 1.0 with the rough registration parameters from intensity based techniques.

Results from Dataset 1b

The chief characteristics of this selected dataset can be summarized as:

- i. Different sensor images
- ii. Incidence angle difference of 20°
- iii. Different sensor orbiting directions

As for the previous dataset the multiresolution optimization framework initiated with the seed (-3°, 0, 0) resulted into a similar kind of performance by the intensity based techniques. The registration results obtained for different resolution images are tabulated in Table 4.

As done for the previous dataset the rough registration parameters can be utilized to assist the SIFT operator matching scheme and thus we repeat the same analysis again. The SIFT matching results for these different sensor, incidence angle, and sensor orbiting directions can be visualized in Table 6. Similar trends as for the previous datasets have been observed for matching ratio 0.6 and 0.8 but the number of matches have reduced significantly. This reduction in the number of matches might be attributed to the different sensor orbiting directions on top of already present different sensor and incidence angle for dataset 1a. Using the matching threshold of 1.0 with the rough registration parameters resulted in 72 matches with almost the similar consistency level as observed for the previous dataset. After using the iterative outlier elimination on the initial 72 matches, 34 of them remain with sub pixel match consistency in both x and y direction. The filtered matches for this dataset using the threshold of 1.0 can be visualized in Figure 6.

Table 5: Intensity based registration of ERS and Radarsat image pairs of dataset 1b

Level	ERS to RS	RS to ERS	RC	TAT (sec)
2	(-5.06°, -1.76, 1.31)	(5.18°, 2.14, -1.40)	0.2	180
1	(-4.99°, -4.18, 3.26)	(5.01°, 4.56, -3.11)	0.16	604
0	(-4.99°, -8.88, 6.03)	(5.07, 10.20, -5.75)	0.50	1450

Table 6: SIFT matching scheme performance for Dataset 1b

Matching Ratio	0.6	0.8	1.0
Matches Found by SIFT	23	274	72
Standard Deviation X	112.33	255.09	2.75
Standard Deviation Y	31.52	262.02	2.64
Filtered Matches	16	26*	34
Standard Deviation X	0.85	1.27	0.88
Standard Deviation Y	1.09	1.32	0.83

*RANSAC

Results from Dataset 2

After encouraging processing chain performance from the first two datasets we test it for images having the following acquisition characteristics:

- i. Different sensor nature (airborne and spaceborne)
- ii. Time difference of 4 years
- iii. Different incidence angle
- iv. High resolution imagery acquired over semi urban land cover

The existing image acquisition differences present a complicated scenario for image matching applications. The influence of different sensor geometries and incidence angle is expected to bring strong aspect dependency on the appearance of urban settlements in the scene. As earlier, we start our analysis showing results from the intensity based registration process. In this case the images only differ by two translations in x and y directions. As far as the intensity based technique is concerned, a performance on similar lines as observed for the last two datasets has also been observed for TerraSAR-X and ESAR imagery. The intensity based registration results after an initialization of (-10, -45) resulted into the performance tabulated in Table 7.

Taking the initial guess obtained from level 2 images we continue with the SIFT operator performance analysis. The effect of complex acquisition differences mentioned above is observed on the SIFT matching results tabulated in Table 8. The standard matching procedures with matching ratios of 0.6 and 0.8 produced far less matches as compared to the matches produced by matching ratio of 1.0 helped by the initial estimate provided by the intensity based process. Considering the nature of the datasets the 57 matches found by the SIFT operator are encouraging (both on ground and within urban establishments). The matched features for this dataset can be visualized in Figure 7 roughly classified into plain area features (green lines) and features matched within the urban establishments (blue lines).

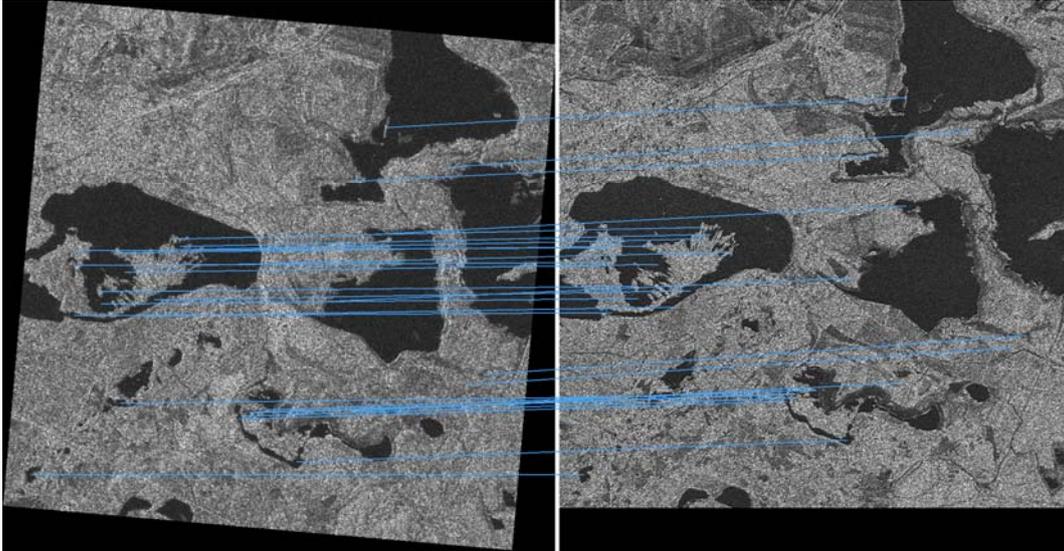


Figure 6: SIFT features matched between transformed ERS (left) and Radarsat-1 images of dataset 1b. 34 SIFT features were matched with a subpixel consistency level using a matching threshold of 1.0 with the rough registration parameters from intensity based techniques.

Table 7: Intensity based registration of TerraSAR-X and ESAR imagery (dataset 2)

Level	TSARX to ESAR	ESAR to TSARX	RC	TAT (sec)
2	(-3.44, -13.15)	(3.44, 13.14)	0.00	70
1	(-6.89, -26.29)	(6.88, 26.28)	0.12	150
0	(-14.80, -52.61)	(15.01, 52.90)	0.36	440

Table 8: SIFT matching scheme performance for Dataset 2

Matching Ratio	0.6	0.8	1.0
Matches Found by SIFT	7	188	57
Standard Deviation X	5.70	266.67	3.49
Standard Deviation Y	146.25	245.54	3.05
Filtered Matches	3*	10*	26*
Standard Deviation X	-	1.66	1.49
Standard Deviation Y	-	0.83	0.79

*RANSAC

All the 57 SIFT features detected here, might be useful for various SAR image matching scenarios but for co-registration only on ground features (38 green colored lines in Figure 8) should ideally be considered for registration parameter estimation. In isolation the on ground features reported a consistency of around 3 pixels in both x and y direction. The 3 pixel consistency achieved by the matched SIFT features might not be enough to co-register images with high accuracy. Henceforth, we run a further refinement of the on ground matches (Green color lines in Figure 7) using MI. After the refinement through local chip matching has been performed, possible control point pair outliers need to be checked for. These outliers might surface due to two main reasons. The first and the most common is the optimizer failure to detect the global maximum and the other possible scenario is lack of enough information in the windows marked around the control points for the metric to produce a sharp enough peak in the registration search space.

For this dataset we could refine 19 conjugate features out of the 38 on ground features as a window of size 300 x 300 (intensity based techniques require good amount of information to produce favorable results) needs to be demarcated centered on each feature. Thus, features lying near the image boundaries could not be considered for further refinement. Finally we obtained 12 conjugate features with a match consistency of 0.75 pixels in x direction and 0.54 pixels in y direction. The transformation parameters for the input TerraSAR-X imagery to the reference ESAR imagery achieved by these

finally matched 12 points are $(-14.10, -53.80)$. It can be observed that the obtained parameters are very much similar to the registration results obtained by MI based registration in Table 8.

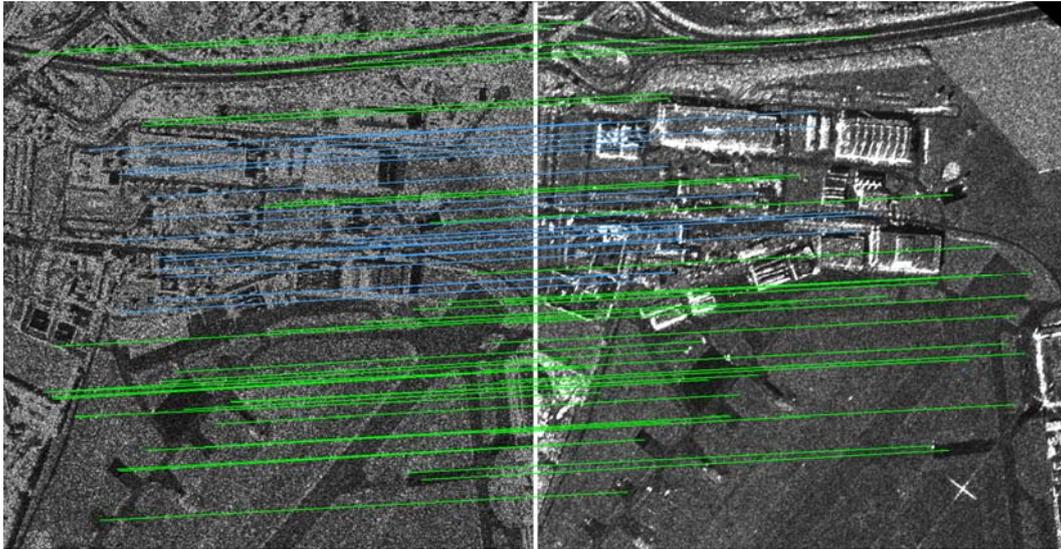


Figure 7: SIFT features matched between transformed TerraSAR-X (left) and ESAR images of dataset 2. In total 57 SIFT features were found using a matching threshold of 1.0 with the rough registration parameters from intensity based techniques. Green lines represent features matched on plain ground and blue lines represent features matched within the urban establishments.

DISCUSSION AND CONCLUSIONS

In this paper we have presented a combination of mutual information and scale invariant feature transform for SAR image matching and registration applications.

MI as a similarity metric for multimodal images has this capability to estimate rough registration parameters from down sampled imagery. It is observed that the magnitude of changes observed in the registration parameters from lower down the image pyramid to the original resolution images may not be of much significance considering the exponential increase in registration turn around times incurred with an increase in image sizes. Therefore, a feature detector capable of utilizing the initially extracted registration parameters can be cascaded with the intensity based technique to achieve registration results with much improved turn around times.

In the presented analysis, to counter the SAR speckle influence, images have been subjected to ISEF filtering and features detected at the first octave were not considered for matching. Previous tests showed that the resulting reduction in feature detection does not necessarily mean a reduction of matches (Schwind et al., in press). These steps provide a significant speed up to the entire processing chain as far lesser number of features are created, resulting in a faster matching process. ISEF filtering or any smoothing for that matter might have some influence on sub pixel feature localization and this might not be tolerant to certain critical applications therefore we recommend fine matching of the conjugate match pairs using MI on the original images to reach desired accuracy levels and to remove any bias introduced by image smoothing.

Originally developed and tested for optical camera images in the field of computer vision, the operator has shown promising results for it to be considered as an alternative for fast and robust feature matching in SAR images. It needs a special mention that the core components of the operator processing chain are not quite optimized for SAR imagery which is statistically very different from their optical counterparts. The technique used for feature detection in the scale space created using the DoG images followed by the local gradient estimation for descriptor formation might not be ideal for SAR imagery. Nevertheless, even with the mentioned shortcomings, the proposed SIFT processing chain produced promising results for SAR image pairs with various acquisition differences (different times, sensors, incidence angles and sensor heading directions)

REFERENCES

- Brown, L.G., 1992. A Survey of Image Registration Techniques, *ACM Computing Surveys*, 24:325-376
- Bovolo, F. and Bruzzone, L., 2005. A Detail-Preserving Scale-Driven Approach to Change Detection in Multitemporal SAR Images. *GeoRS*, 43: 2963-2972.
- Borghys, D., Perneel, C. and Acheroy, M., 2001. A hierarchical approach for registration of high-resolution polarimetric sar images. *Proceedings from the SPIE Conference on Image and Signal Processing for Remote Sensing*, September 17-21, 2001
- Chen, H., Varshney, P.K., and Arora, M.K., 2003 Mutual information based image registration for remote sensing data. *International Journal of Remote Sensing*, 24(18):3701-3706.
- Cole Rhodes, A.A., Johnson, K.L., LeMoigne, J., and Zavorin, I., 2003. Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient. *IEEE Trans. on Image Processing*, 12(12):1495-1511.
- Eineder, M., Breit, H., Fritz, T., Schättler, B. and Roth, A., 2005, TerraSAR-X SAR Products and Processing Algorithms. *Proceedings from the IGARSS*, 4870-4873, July 25-29, 2005.
- Fischler, M.A. and Bolles, R.C., 1981, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 2:381-395.
- Hegarar-Masce, L., S., Zribi, M., Alem, F., Weisse, A. and Loumagne, C., 2002, Soil moisture estimation from ERS/SAR data: toward an operational methodology. *IEEE Trans. on Geoscience and Remote Sensing*, 40:2647-2658.
- Hua X., Pierce, L.E., and Ulaby, F.T., 2003 Mutual information based registration of SAR images. . *Proceedings from the IGARSS*, 6:4028- 4031, 28June-02 July, 2003
- Holden, M., Hill, D.L.G., Denton, E.R.E., Jarosz, J.M., Cox, T.C.S., Rohlfing, T., Goodey J., and Hawkes, D.J., 2000. Voxel Similarity measures for 3-D serial MR brain image registration. *IEEE Transactions on Medical Imaging*, 19(2):94-102.
- Hough, P.V.C., Methods and means for recognizing complex patterns. *U.S.Patent*, 3069654, 1962.
- Krieger, G., Fiedler, H., Zink, M., Hajnsek, I., Younis, M., Huber, S., Bachmann, M., Hueso Gonzalez, J., Werner, M. and Moreira, A., 2007, The TanDEM-X mission: A satellite formation for high resolution SAR interferometry. *Proceedings from the International Astronautical Congress*, September 24-28, 2007.
- Li, H., Manjunath, B. and Mitra, S., 1995, A Contour-Based Approach to Multisensor Image Registration. *IEEE Trans. on Imaging Processing*, 4, 320-334.
- Lowe, D., 2004, Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 20:91-110.
- Lowe, D.G., 1999, Object Recognition from Local Scale-Invariant Features. *Proceedings from the International Conference on Computer Vision*, 1150-1157, September 20-25, 2007
- Moghaddam, M., Dungan, J.L. and Acker, S., 2002, Forest variable estimation from fusion of SAR and multispectral optical data. *IEEE Trans. on Geoscience and Remote Sensing*, 40:2176-2187
- Palubinskas, G., Meyer, F.J., Runge, H., Reinartz, P., Scheiber, R. and Bamler, R., 2005, Estimation of along-track velocity of road vehicles in SAR data. *Proceedings from the SPIE conference on Image and signal processing for remote sensing*, 5982:1-9, September 20-22, 2005.
- Roth, A., 2003, TerraSAR-X: A new perspective for scientific use of high resolution spaceborne SAR data. *Proceedings from the GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas URBA*, May 22-23, 2003
- Schreiber, R., Reigber, A., Ulbricht, A., Papathanassiou, K., Horn, R., Buckreuss, S. and Moreira, A., 1999, Overview of interferometric data acquisition and processing modes of the experimental airborne SAR system of DLR. *Proceedings from the IGARSS* 1:35-37, 28 June-02 July, 1999
- Schwind, P., Suri, S., Reinartz, P., Siebert A., (In Press) Applicability of the SIFT operator to geometric SAR image registration. To be published in *International Journal of Remote Sensing*
- Suri, S., and Reinartz, P., (2008). Application of generalized partial estimation for mutual information based registration of high resolution sar and optical imagery. *Proceedings from the International Conference on Information Fusion*, June 30- July 3, 2008.
- Touzi, R., Lopes, A. and Bousquet, P., 1988, A statistical and geometrical edge detector for SAR images. *IEEE Trans. Geoscience and Remote Sensing*, 26:764-773.
- Wachowiak, M.P., Smolikova, R., and Peters, T.M., 2003. Multiresolution Biomedical Image Registration Using Generalized Information Measures, *Lecture Notes in Computer Science* 2879 (MICCAI 2003):846-853.