Evaluating the applicability of BRISK for the geometric registration of remote sensing images

P. Schwind[a][*] and P. d'Angelo[a]

[a]*German Aerospace Center (DLR), EOC, Oberpfaffenhofen, Germany*

Binary Robust Invariant Scalable Keypoints (BRISK) is one of several relatively new matching algorithms aiming to improve well-established algorithms such as Scale-Invariant Feature Transform (SIFT) or Speeded-Up Robust Features. A detailed evaluation of the BRISK applicability for geometric registration of remote sensing images is performed. As the original algorithm was not developed with a focus on remote sensing image matching, a practical processing chain for the image registration of a newly acquired image with a reference image was developed. This chain also includes a modified Random Sample Consensus outlier removal based on the sensor-model of the to-be-registered image. The presented methodology is evaluated and compared to the SIFT operator in terms of repeatability, accuracy, recall and precision. Our results show that BRISK performs very well on remote sensing images and together with the sensor-model-based outlier removal offers a significant improvement over existing image registration methods such as SIFT.

1. Introduction

Geometric image registration is one of the most fundamental prerequisites when working with multiple remote sensing images. Using satellite orientation and position measurements alone is often not sufficient to achieve the high geometric accuracies required for many applications such as image fusion, change detection or digital elevation model (DEM) generation. Imagery acquired using airborne cameras (e.g. HySpex (Schwind, Schneider, and Müller 2014)) and even modern satellite platforms (e.g. Skybox (d'Angelo, Kuschk, and Reinartz 2014)) often has to be corrected using ground control points (GCP) to obtain the desired geometric accuracy.

The two main approaches to automatically obtain GCP from reference images are intensity (area)- and feature-based image matching (Zitová and Flusser 2003). While intensity-based algorithms can often achieve a higher accuracy, they are usually computationally more expensive than feature-based approaches. Another limitation of intensity-based approaches is that they do not perform ideally if precise information about the rotation and scale of the image is not available. In addition to that, intensity-based approaches require mostly flat surfaces to perform optimally. Matching two scenes from different viewpoints will usually result in co-registration inaccuracies if the scene features elevated objects, such as buildings or trees. Most state-of-the-art feature-based matching algorithms on the other hand can handle rotation, scale and viewpoint changes. Another important advantage is the possibility of storing the computed

[*]Corresponding author. Email: peter.schwind@dlr.de

descriptors, so they can be reused later to match newly acquired scenes. Therefore, it is often preferable to employ approaches based on point features to match remote sensing images.

While image registration and feature matching have been under research since the late 1980s in the remote sensing community (Förstner and Gülch 1987), in recent years, many new approaches have been developed in the field of computer vision. Arguably the most widely used methods of these are Scale-Invariant Feature Transform (SIFT, Lowe (2004)) and its derivative Speeded-Up Robust Features (SURF, Bay, Tuytelaars, and Van Gool (2006)). Subsequently, many authors developed further feature matching approaches, often promising similar or better accuracy than SIFT and SURF at reduced runtime. Some of the best known of them are:

- Features from Accelerated Segment Test (FAST, Rosten and Drummond (2006))
- Fast Retina Keypoint (FREAK, Alahi, Ortiz, and Vandergheynst (2012))
- Binary Robust Invariant Scalable Keypoints (BRISK, Leutenegger, Chli, and Siegwart (2011))

Many of these have been implemented as part of the OpenCV library (Bradski 2000), enabling an easy comparison for selected challenging test data sets.

Amongst these algorithms, especially SIFT and SURF have been evaluated and adapted for use in remote sensing applications. One recent application of SIFT was presented by (Ye and Shan 2014), who registered multi-spectral images using *Scale Restriction* SIFT for pre-registration followed by the feature extraction using a local self-similarity descriptor. A different example for SURF can be seen in (Wang et al. 2015), where SURF descriptors are calculated for interest points detected using a multi-scale Harris-Laplacian corner detector.

In our preliminary tests using OpenCV, BRISK produced the most promising results out of all these algorithms. Consequently, we already used BRISK, followed by a sensor-model-based RANSAC (RANdom SAmple Consensus, (Fischler and Bolles 1981)) to improve the co-registration accuracy of two HySpex cameras (Schwind, Schneider, and Müller 2014). In that application, BRISK performed very well, often finding tens of thousands of matches in one scene pair. However, this was a relatively straightforward matching application, as the two matched scenes were acquired almost simultaneously from the same view point. Another important use case is the matching of a newly acquired *input image* with an older, precisely georeferenced *reference image* . The resulting matches are employed as GCP for the estimation of improved camera and sensor parameters of the input image. Finally, these parameters are processed together with a digital elevation model to create an orthorectified image. This is a much more challenging task, as the matched images in such a case are usually acquired at different times, from different viewing angles and with varying spectral ranges.

Therefore, the objective of this work was to develop an operational processing chain for the co-registration of an input image with a reference image based on BRISK. Subsequently, the applicability of the proposed methodology is evaluated in detail for selected test data sets, measuring the detector performance in terms of repeatability and accuracy and the descriptor performance in terms of precision and recall. The conducted evaluation shows that BRISK performs favourable in the context of remote sensing image registration and provides a valuable alternative to established matching algorithms such as SIFT.

2. Technical background and methodology

In the following subsections, the technical background, required to understand the proposed processing chain, is provided. Namely, the BRISK algorithm and descriptor are outlined and the working principle of the used RANSAC based outlier removal is introduced. Finally, in Section 2.3, the image registration processing chain, which makes use of these algorithms, is presented.

2.1. *BRISK*

Similar to other feature detectors such as SIFT, the BRISK (Leutenegger, Chli, and Siegwart 2011) algorithm can be subdivided into four main processing steps:

(1) Scale space Keypoint Detection
(2) Keypoint filtering and sub-pixel localization
(3) Orientation assignment
(4) Descriptor generation

For the keypoint detection, a pyramid scale-space is built by repeatedly down-sampling the input image into n octaves c_i and n intra-octaves d_i. The n octaves are created by repeatedly half-sampling the original image. The intra-octaves d_i are generated similarly, except that the first intra-octave d_0 is created by down-sampling the original image by a factor of 1.5. Using the FAST 9-16 detector (Rosten and Drummond 2006), keypoint candidates are selected in the image pyramid.

In the next processing step, these keypoint candidates are filtered by performing a 3D non-maxima suppression within the scale-space pyramid. The remaining keypoints are interpolated to a sub-pixel position. First 2D maxima are estimated by fitting a 2D quadratic function in a $3-\text{pixel} \times 3-\text{pixel}$ region in the pyramid layer of the keypoint as well as the bordering layers. Next, a 1D parabola is fitted to the three maxima to find the 3D maximum in the scale-space. For performance reasons this step is executed somewhat differently for octave c_0, where FAST 5-8 is used to create the intra-octave d_{-1} (see (Leutenegger, Chli, and Siegwart 2011) for details).

For the orientation assignment a sampling pattern, resembling that used by the DAISY descriptor (Tola, Lepetit, and Fua 2010), is applied. The points $\boldsymbol{p}_x$ of this pattern are paired with each other and the gradients of these pairs are calculated. In this work, a point pair is represented by the indices i and j. For the set of all long-distance pairs, L, meaning pairs with a distance bigger than $13.67t$ (t is the scale of the keypoint), the gradient $\boldsymbol{g}(\boldsymbol{p}_i, \boldsymbol{p}_j)$ between the point pair $\boldsymbol{p}_i, \boldsymbol{p}_j$ is calculated using the formula

$$\boldsymbol{g}(\boldsymbol{p}_i, \boldsymbol{p}_j) = (\boldsymbol{p}_j - \boldsymbol{p}_i) \frac{I(\boldsymbol{p}_j, \sigma_j) - I(\boldsymbol{p}_i, \sigma_i)}{\|\boldsymbol{p}_j - \boldsymbol{p}_i\|^2} \tag{1}$$

where $I(\boldsymbol{p}_x, \sigma_x)$ is the smoothed value around $\boldsymbol{p}_x$, computed using a Gaussian filter with standard deviation σ_x. Next the average direction $\boldsymbol{g}_{avg}$ of all possible long distance pairs

$$\boldsymbol{g}_{\text{avg}} = \begin{pmatrix} g_{\text{avg}_x} \\ g_{\text{avg}_y} \end{pmatrix} = \frac{1}{L} \cdot \sum_{(\boldsymbol{p}_i, \boldsymbol{p}_j) \in \text{L}} \boldsymbol{g}(\boldsymbol{p}_i, \boldsymbol{p}_j) \tag{2}$$

is calculated, with L being the number of long distance pairs. The orientation of the keypoint can then be calculated by $\alpha = \text{atan2}(g_{\text{avg}_y}, g_{\text{avg}_x})$, with g_{avg_x} and g_{avg_y} being the average gradient in horizontal and vertical direction, respectively.

To finally generate the descriptor, the sampling pattern is rotated by α and all pair combinations with a distance smaller than $9.75t$ are compared to each other. This results in a descriptor vector with a length of 512 where each element b is defined as

$$b = \begin{cases} 1, & I(\boldsymbol{p}_j^\alpha, \sigma_j) > I(\boldsymbol{p}_i^\alpha, \sigma_i) \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

As this binary descriptor does not describe the absolute difference between two points, it is invariant to monotonic grey value changes between the scenes.

2.2. *Sensor-model RANSAC*

RANSAC is an algorithm very commonly used in computer vision and remote sensing to remove matching outliers and estimate the transformation model between two matched images. Usually an affine or polynomial model is used in RANSAC to transform matches from one image to the other. While this approach works well for relatively flat scenes, the affine or polynomial 2D image transformation is not always a good approximation for the sensor model, and does not capture effects due to undulating terrain and non-linear sensor perturbations. In such cases, RANSAC would require the use of a high error threshold, which can lead to rejection of good matches and acceptance of outliers. For the presented application, we chose to use a more accurate approach, by using the sensor model of the input image for the RANSAC transformation model.

For every RANSAC iteration, two random matches are selected in the input image. Using the sensor model as defined by Müller et al. (2012), the coordinates of these matches are transformed to the reference image. The corresponding matching coordinates in the reference images are used in a least squares fit to estimate the boresight angles of the input image sensor. In the next step, the sensor model together with the estimated boresight angles is applied to all the remaining matching coordinates in the input image and the distance of the calculated coordinates to the matching coordinates in the reference image is calculated. For each RANSAC iteration, the number of distances below a predetermined threshold is computed, and once the iterations end, the boresight angles which resulted in the highest such number are accepted as the final output of RANSAC. In a subsequent processing step, these angles can be used to orthorectify the input image and align it with the reference image. For a more detailed description of the sensor-model-based RANSAC, please refer to Schwind, Schneider, and Müller (2014).

2.3. *Image registration processing chain*

Using the methods described above, implementing a processing chain for multi-sensor matching and registration is straightforward. First, the input image is brought to the geometry of the reference ortho image by using the approximate corner coordinates of the image for an affine transformation. Next, the images are matched using BRISK, resulting in matching coordinates in both images. Then, the matching coordinates of the input image are transformed back to the coordinate of the raw input image. These matching coordinates are used as input for the sensor-model-based RANSAC, which removes outliers from the matches and computes the boresight angles needed to correct the geometry of the input image. The last step, which is not discussed in detail in this paper, is to orthorectify the input image using the sensor-model and taking the boresight angles into account.

The reference implementation of BRISK, available as part of the OpenCV library (Bradski 2000), unfortunately only supports images with 8 bits per channel. To allow the processing of images with 16 bits per channel and to achieve a better understanding of the algorithm, we implemented the BRISK algorithm ourselves in C++. For the matching of the BRISK descriptors, the locality-sensitive hashing (LSH) algorithm in the publicly available Fast Library for Approximate Nearest Neighbors (FLANN) was used, see (Muja and Lowe 2009).

As remote sensing images can usually be brought to the same scale and orientation before processing, options to turn off the rotation and scale invariance independently were added. If the rotation invariance is disabled, the orientation of all keypoints is assumed to be 0°. If the scale invariance is turned off, keypoints are still detected at different scales in the scale space

pyramid, but their position is not interpolated in 3D, only in 2D.

To avoid selecting the contrast threshold used by the FAST 9-16 detector for every new image type, a very low default threshold of 2 is employed for all images. An even interest point distribution is achieved by subdividing the image into a grid and selecting only the N points with the highest contrast threshold. In this work, $N = 500$ points were selected for each grid cell of 500 pixels $\times$ 500 pixels, avoiding manual selection of the contrast threshold. Of course, such an approach has a negative impact on the runtime of the algorithm, but eliminating the need for any kind of manual parameter selection is definitely worth this drawback, especially when using this method in an automated production chain. Also, since the N best points can already be selected before extracting the descriptor, the impact on the runtime is not very significant.

The RANSAC algorithm is implemented using previously developed sensor-model libraries (Müller et al. 2012). The error threshold of RANSAC is set to two pixels and the number of iterations to 200.

3. Evaluation

To test if BRISK followed by sensor-model-based RANSAC is applicable for the matching and registration of multi-sensor remote sensing images, the proposed methodology was evaluated rigorously. The examined use case is focused on the matching and registration of a newly acquired input image with an already orthorectified reference image.

3.1. *Evaluation methodology*

The evaluation is subdivided in two steps: First, the detector performance is evaluated, by computing the repeatability (Mikolajczyk et al. 2005) and accuracy of the found keypoints in the two images. Next, the efficiency of the descriptors is investigated by calculating the recall and precision (Mikolajczyk and Schmid 2005) of the matches.

The repeatability is defined as the ratio between number of match correspondences (repeats) and the minimum of the keypoints found in both images. We calculate the distance error ϵ_d for each keypoint pair, used to determine the repeats, according to the following formula:

$$\epsilon_{\mathrm{d}} = d_{\mathrm{k}} + d_{\mathrm{rs}} \tag{4}$$

where d_{k} is the Euclidean distance between the keypoints and d_{rs} is the Euclidean distance between their rotation/scale points. These rotation/scale points are located at a distance s from the keypoint, which corresponds to the scale and an angle r which corresponds to the rotation of the keypoint (see Figure 1). Only points with a distance error $\epsilon_{\mathrm{d}} < 2$ are included in the calculation of the repeatability statistics.

The geometric accuracy of the matched points is determined by fitting the sensor model against all matches and calculating the root mean squared error (RMSE) for these matches.

The precision and recall values computed to evaluate the performance of the descriptors are defined as the ratio between the true positives to all possible positives and the ratio between the true positives to the match correspondences, respectively.

The evaluation was performed for both BRISK and SIFT. The SIFT algorithm is widely used due to its good performance across a wide area of applications, and it provides a good baseline for the BRISK algorithm. The used SIFT program was implemented in C and has been used operationally for about 5 years within DLR, meaning it should by now have a similar or better degree of optimization than the used BRISK implementation. The detection strategy used for
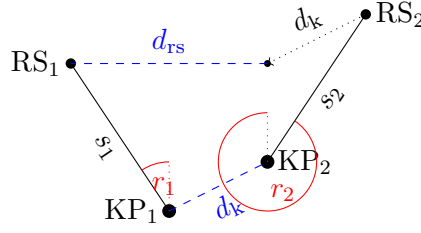
Figure 1. ϵ_d is computed by calculating the distance d_k of the keypoints (KP) and the distance d_rs of their rotation/scale points (RS). The location of each RS point is derived from the scale s and rotation r of the corresponding keypoint.

Table 1. Evaluation of the BRISK and SIFT detector

| | # Repeats | | Repeatability | | Runtime (s) | |
	BRISK	SIFT	BRISK	SIFT	BRISK	SIFT
Data set 1	18070	14403	0.25	0.24	132	427
Data set 2	2483	4258	0.07	0.04	36	634
Data set 3	592	873	0.01	0.02	90	354

SIFT is identical to the strategy used for BRISK (see Section 2.3), meaning a very low detection threshold and the same grid-based approach was used.

3.2. *Data sets*

Three data sets were selected for the evaluation. The data sets are challenging examples; a classical pyramidal correlation approach tested previously for these scenes has problems finding suitable matches. The input images were acquired using ALOS-AVNIR-2 and the reference images are based on a Landsat ETM+ reference mosaic. For the AVNIR-2 scenes an average of the four bands was used as input for the matching, while for the Landsat scenes the panchromatic channel was used. The dimensions of the input images are 7100 pixels × 8000 pixels but after bringing them to the reference image geometry for matching, their dimensions change to approximately 10,000 pixels × 10,000 pixels. The data sets were chosen with a focus on landcover diversity: Data set 1 features a desert area with almost no man-made structures and Data set 2 contains three islands with mostly agricultural areas. The final Data set 3 represents a very difficult test case, as it mostly consists of rain forest and the input scene features significant cloud cover. The used input and reference images can be seen in Figure 2.

3.3. *Results*

The repeatability evaluation results, including the runtime for feature detection and descriptor extraction are given in Table 1. In all three test cases SIFT produces more repeats than BRISK, but the repeatability is higher when using BRISK for Data sets 1 and 2. Only for Data set 3, where both detectors have a very low repeatability, the repeatability of SIFT is higher than that of BRISK. The runtime of SIFT is at least three times as high as the time required by BRISK in all three test cases.

Table 2 shows the achieved accuracies of the BRISK and SIFT matches for the three test data sets. In all test cases, where matches remain after RANSAC, subpixel accuracy is achieved. For Data sets 1 and 2 SIFT achieved better accuracy than BRISK; however the differences are minimal.

Table 2. Accuracy of the BRISK and SIFT matches in pixels

| | BRISK | | SIFT | |
	RMSE(x)	RMSE(y)	RMSE(x)	RMSE(y)
Data set 1	0.77	0.80	0.72	0.71
Data set 2	0.87	0.67	0.66	0.57
Data set 3	0.97	0.64	-	-

Table 3. Descriptor Evaluation

| | # Inliers | | # Outliers | | Recall | | Precision | |
	BRISK	SIFT	BRISK	SIFT	BRISK	SIFT	BRISK	SIFT
Data set 1 w/o RANSAC	5525	771	33	3617	0.31	0.05	0.99	0.18
Data set 1 after RANSAC	4804	709	0	0	0.27	0.05	1.00	1.00
Data set 2 w/o RANSAC	475	173	22	825	0.19	0.04	0.96	0.17
Data set 2 after RANSAC	424	158	0	0	0.17	0.04	1.00	1.00
Data set 3 w/o RANSAC	15	6	32	593	0.03	0.01	0.32	0.01
Data set 3 after RANSAC	11	0	0	0	0.02	0.00	1.00	0.00

The results of the descriptor evaluation are shown in Table 3. While the number of matches decreased with an increase in the complexity of the data sets, it is notable that BRISK produced significantly more correct matches in all cases. In terms of recall, BRISK outperformed SIFT in all three cases by a factor higher than 4. Concerning the precision, the results also favour BRISK. For Data set 1 and 2, the precision is already close to 1 even without removing any outliers and for Data set 3 where it is much lower at 0.319 it is still a lot better than the precision of 0.01 achieved by SIFT.

Concerning the performance of the outlier removal, it can be seen in Table 3 that the sensor-model-based RANSAC successfully removes the outliers in all test cases. Only in one case, namely for the matches produced by SIFT for Data set 3, it fails to find any of the valid matches. In this case however there were only six possible inliers out of 599 matches, making the outlier removal very challenging.

For visual interpretation, the matches remaining after outlier removal are displayed in Figure 2. As can be seen, for Data set 1 (Figures 2(a), 2(d)) and Data set 2 (Figures 2(b), 2(e)) the number and distribution of matches over the entire scene is satisfying. For Data set 3 (Figures 2(c), 2(f)) on the other hand there are only 11 matches distributed over the lower half of the input image. While this is not optimal for the geometric correction, it could be expected as this data set represents a borderline case for matching, which is also why for SIFT no matches remain after outlier removal (see Table 3).

4. Discussion and conclusions

Concerning the results of the detector evaluation, it is not easy to say which of the two tested detectors performed better. For the first data set both methods have very similar repeatability scores. For Data set 2 BRISK performed somewhat better, but for Data set 3 SIFT achieved a better score. The very low absolute repeatability can be explained by the fact that a very low detection threshold was used (see Section 2.3). It would be possible to significantly improve the repeatability by fine-tuning the detection threshold based on the contrast of the input scenes.
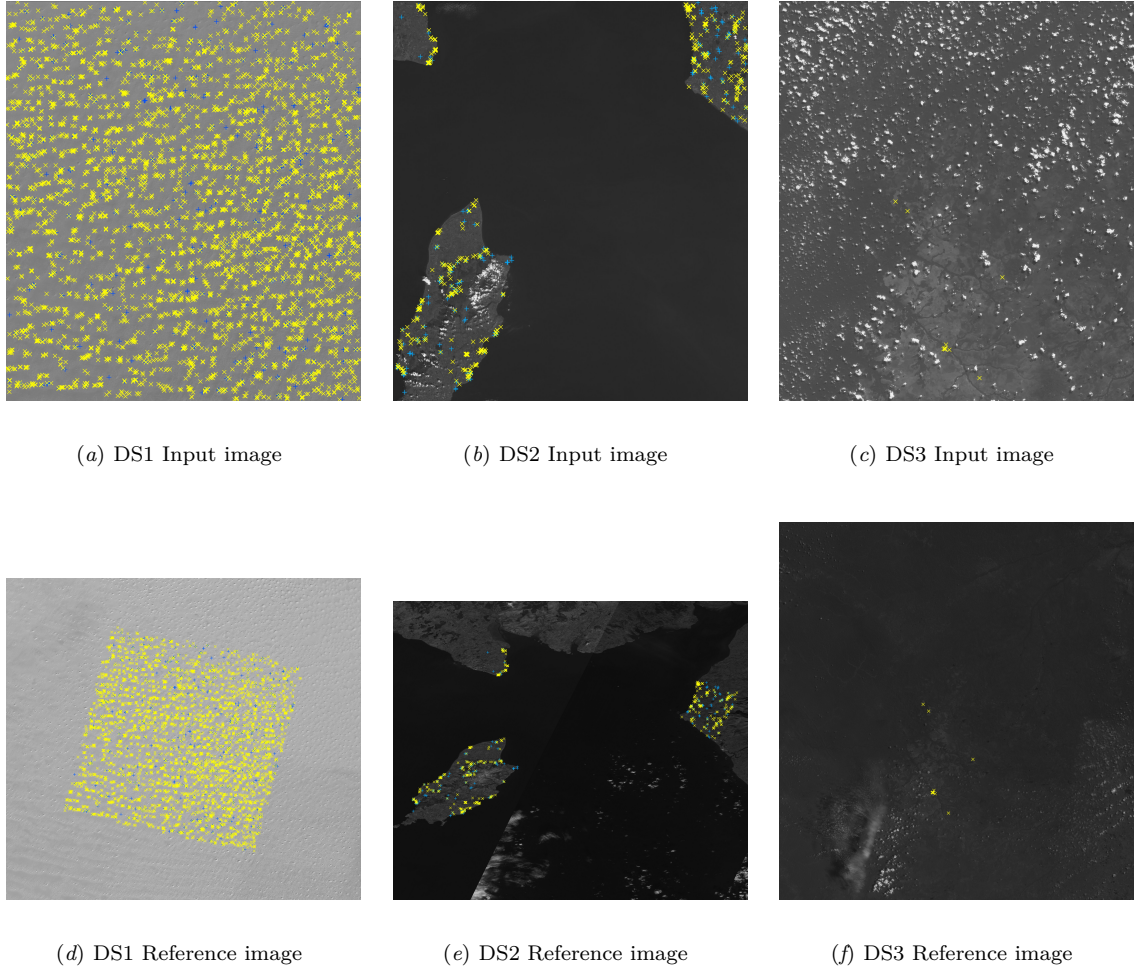
(*a*) DS1 Input image   (*b*) DS2 Input image   (*c*) DS3 Input image

(*d*) DS1 Reference image   (*e*) DS2 Reference image   (*f*) DS3 Reference image

Figure 2. Test Data sets containing BRISK ($\times$) and SIFT ($+$) matches remaining after outlier removal

As the contrast of remote sensing data varies strongly, reliable, automatic processing is not possible when using a fixed threshold.

For the descriptor tests, the results are less ambiguous. Even though SIFT produced a higher number of repeats for two of the Data sets (Data sets 2 and 3) and even a higher repeatability in one case (Data set 3), it generated much less correct matches than BRISK. The ratio of inliers versus outliers is also much better for the BRISK matches. Together with the computed recall and precision values, which clearly favour BRISK, the results show that the BRISK descriptor performs much better than the SIFT descriptor for the matching of the tested remote sensing images. This is all the more true if the runtime is also considered, as BRISK is at least three times faster in all test cases. Only the results achieved in the accuracy evaluation favour SIFT somewhat, if only marginally. In any, case both SIFT and BRISK achieve subpixel accuracy for the tested data sets.

It was also shown that the robust outlier removal using a sensor-model-based RANSAC complements the presented BRISK-based processing chain very well. In all test cases most of the inliers were found while all of the outliers were removed.

After evaluating the presented processing chain, it can be concluded that the detector stage of BRISK performs similar to that of SIFT if the repeatability is compared. Concerning the accuracy, SIFT performs marginally better than BRISK. If the runtime of the detection is taken

into consideration however, the BRISK detector is clearly in advantage. This also applies to the descriptor comparison, where BRISK performed better than SIFT in every evaluated metric. By using it in the proposed processing chain which makes use of all the geometric knowledge available, BRISK offers a significant improvement compared to other well-established matching algorithms such as SIFT.

Considering the somewhat better accuracy of the SIFT detector, it might be worth investigating a combination of the SIFT detector with the BRISK descriptor in the future. Ideally, when using such a combination, the BRISK descriptor should be adapted to be able to compute it in the SIFT scale space to avoid the overhead of computing two different scale spaces. Another possible improvement could be a different strategy for the descriptor matching. The currently used LSH algorithm does not make use of the fact that the approximate relative distribution of the matches is already known. If, for example, the matching algorithm would only compare descriptors in a selected area instead of all descriptors, the number of correct matches would likely be increased.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Alahi, A., R. Ortiz, and P. Vandergheynst. 2012. "FREAK: Fast retina keypoint." In *Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, June 16âĂŞ21,* 510–517. IEEE.

Bay, H., T. Tuytelaars, and L. Van Gool. 2006. "SURF: Speeded Up Robust Features." In *Computer Vision–ECCV 2006,* 404–417. Springer.

Bradski, G. 2000. "The OpenCV Library." *Dr. Dobb's Journal of Software Tools* .

d'Angelo, P., G. Kuschk, and P. Reinartz. 2014. "Evaluation of Skybox Video and Still Image products." *ISPRS Archives* XL-1: 95–99.

Fischler, M. A., and R. C. Bolles. 1981. "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography." *Communications of the ACM* 24 (6): 381–395.

Förstner, W., and E. Gülch. 1987. "A fast operator for detection and precise location of distinct points, corners and centres of circular features." In *Proc. ISPRS intercommission conference on fast processing of photogrammetric data,* 281–305.

Leutenegger, S., M. Chli, and R. Y. Siegwart. 2011. "BRISK: Binary robust invariant scalable keypoints." In *Computer Vision (ICCV), 2011 IEEE International Conference on,* 2548–2555. IEEE.

Lowe, D. G. 2004. "Distinctive image features from scale-invariant keypoints." *International journal of computer vision* 60 (2): 91–110.

Mikolajczyk, K., and C. Schmid. 2005. "A performance evaluation of local descriptors." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 27 (10): 1615–1630.

Mikolajczyk, K., T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. 2005. "A comparison of affine region detectors." *International journal of computer vision* 65 (1-2): 43–72.

Muja, M., and D. G. Lowe. 2009. "Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration." In *International Conference on Computer Vision Theory and Application VISSAPP'09,* 331–340. INSTICC Press.

Müller, R., T. Krauß, M. Schneider, and P. Reinartz. 2012. "Automated Georeferencing of Optical Satellite Data with Integrated Sensor Model Improvement." *Photogrammetric Engineering & Remote Sensing* 78 (1): 61 –74.

Rosten, E., and T. Drummond. 2006. "Machine learning for high-speed corner detection." In *Computer Vision–ECCV 2006,* 430–443. Springer.

Schwind, P., M. Schneider, and R. Müller. 2014. "Improving HySpex Sensor Co-registration Accuracy using BRISK and Sensor-model based RANSAC." *ISPRS Archives* XL-1: 371–376.

Tola, E., V. Lepetit, and P. Fua. 2010. "DAISY: An Efficient Dense Descriptor Applied to Wide-Baseline Stereo." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 32 (5): 815–830.

Wang, W., T. Cao, S. Liu, and E. Tu. 2015. "Remote Sensing Image Automatic Registration on Multi-scale Harris-Laplacian." *Journal of the Indian Society of Remote Sensing* 1–11.

Ye, Y., and J. Shan. 2014. "A local descriptor based registration method for multispectral remote sensing images with non-linear intensity differences." {*ISPRS*} *Journal of Photogrammetry and Remote Sensing* 90 (0): 83 – 95.

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