SAR-SIFT: A SIFT-LIKE ALGORITHM FOR APPLICATIONS ON SAR IMAGES

Flora Dellinger[1], *Julie Delon*[1], *Yann Gousseau*[1], *Julien Michel*[2], *Florence Tupin*[1]

[1]Telecom ParisTech, LTCI CNRS, Paris, France
[2]CNES, DCT/SI/AP, Toulouse, France

ABSTRACT

The scale invariant feature transform (SIFT) algorithm, commonly used in computer vision, does not perform well on synthetic aperture radar (SAR) images, in particular because of the strong intensity and the multiplicative nature of the noise. We present an improvement of this algorithm for SAR images. First, a robust yet simple way to compute gradient on radar images is introduced. This step is first used to develop a new keypoints extraction algorithm, based on the Harris criterion. Second, we rely on this gradient definition to adapt the computation of both the main orientation and the geometric descriptor to SAR image specificities. We validate this new algorithm with different experiments and present an application of our new SAR-SIFT algorithm.

Index Terms— SAR images, SIFT

1. INTRODUCTION

The last generation of earth observation satellites has led to an improvement of spatial resolution of satellite images allowing new applications to be developed. In a situation such as change detection after disasters, available data on the damaged area are often multi-sensor, multiresolution and multimodal images. A feature-based approach therefore seems more suitable than a pixel-based one in this context. The SIFT algorithm [1] is widely used in computer vision to detect and recognize an object present in different images and has shown great efficiency [2]. Its invariances to image translation, scaling, rotation and partially to illumination changes are appropriate to the challenge of change detection in heterogeneous images. Applications can include object recognition, image stitching or video tracking. This algorithm has been developed for natural images with low gaussian noise, and may therefore be easily adapted to optical satellite images. However, its performances are very poor when applied to synthetic aperture radar (SAR) images. Speckle noise creates a lot of false feature detections and only a few of them are usually matched correctly. Some developments have been made recently for the detection and extraction of local features on SAR images [3] [4] but they rely on questionable definitions of the gradient, as we will further discuss. In this paper, we propose a new SIFT-like algorithm adapted to the specificities of SAR images.

1.1. Presentation of the SIFT algorithm

The SIFT algorithm consists in the detection of keypoints and the association of a highly discriminative geometrical descriptor, called descriptor SIFT. Keypoints detected in two images can then be combined by comparing their respective descriptors.

This work is found under a CNES grant

The SIFT keypoints are selected as local interest points and characterized by their localisation, scale and orientation. These are detected as local extrema (both in space and scale) of the Laplacian of the Gaussian (LoG) scale-space of the image, constructed with the scales $\sigma_k = \sigma_0 \cdot r^k$ with $k = 0..k_{max} - 1$. In the remaining of this paper, we will refer to this keypoint detection method as LoG. To suppress candidates with low contrast or located on edges, the original SIFT algorithm relies on the Hessian matrix. A classical variation is to use the multi-scale Harris corner detector [5]. The Harris matrix is defined as:

$$C(x,y,\sigma) = \sigma^2 \cdot G_s \star \begin{bmatrix} (\frac{\partial I_\sigma}{\partial x})^2 & (\frac{\partial I_\sigma}{\partial x}) \cdot (\frac{\partial I_\sigma}{\partial y}) \\ (\frac{\partial I_\sigma}{\partial x}) \cdot (\frac{\partial I_\sigma}{\partial y}) & (\frac{\partial I_\sigma}{\partial y})^2 \end{bmatrix}$$

with G_s a gaussian kernel with standard deviation $s = \sqrt{2} \cdot \sigma$, $\star$ the convolution operator and I_σ the convolution of the original image by a gaussian kernel with standard deviation σ. The points are suppressed by applying a threshold t_H on the multi-scale Harris criterion, defined as $R(x,y,\sigma) = \det(C(x,y,\sigma)) - t \cdot \mathrm{tr}(C(x,y,\sigma))$, with t a parameter. In order to associate an orientation to each keypoint, a histogram of gradient orientations, weighted by the gradient magnitude, is then computed on a neighborhood around each keypoint and the main directions are selected. Eventually, to compute the descriptors, the neighborhood of each keypoint is divided into sectors and on each of them are computed histograms of the gradient orientations, again weighted by the gradient magnitude. For each keypoint, the SIFT descriptor is made of the normalized vector gathering these histograms.

Finally, keypoints extracted from two or more images are matched by thresholding distances between their respective descriptors.

We will use the following parameters, commonly chosen for optical images: $t = 0.04$, $\sigma_0 = 0.63$, $r = 2^{1/3}$ and $k_{max} = 13$, being the number of scales. t_H is usually set to 2000 for 8-bits images, but will be adapted for each SAR image, due to their difference of dynamic.

1.2. Limitations of the SIFT algorithm on SAR images

When applied on SAR images, the first step of the keypoint detection, the LoG method, leads to many false detections. Contrary to what happens with natural images, these false detections are not suppressed by the second step (a threshold on the Harris or Hessian criterion). Indeed false candidates made by strong speckle noise on SAR images still remain after the multi-scale Harris criterion step because of their high contrast. Many of these false alarms occur at the smallest scales. Suppressing the detections at these scales as suggested in [3] removes many false alarms but also strongly decreases the number of keypoints. Moreover, because of the multiplicative nature of speckle noise, the regular gradient by difference creates

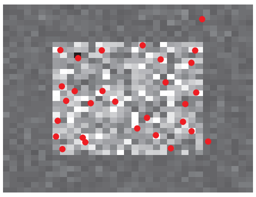


Fig. 1: Results of the LoG keypoints detection method applied on a speckle noised square amplitude image (29 keypoints detected).

stronger gradient magnitudes in homogeneous areas of high reflectivity than in the ones of low reflectivity. This leads to many false detections on homogeneous areas of high reflectivity, as seen on the example of a speckle noised square on Figure 1. This local interest point detection method is thus not suitable to SAR images.

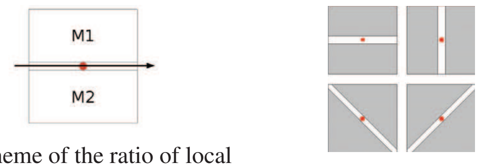
The other steps of the algorithm, that is the computation of orientations and descriptors, also rely on a classical gradient by difference. The resulting orientations and descriptors are therefore not robust to multiplicative noise.

To resolve these issues, a new gradient computation, adapted to SAR images, needs to be developed in order to improve the SIFT algorithm. We present in Section 2 a new computation of the gradient for SAR images. Then, in Section 3, a variant of the SIFT algorithm adapted to SAR images is introduced and numerically validated. In Section 4 an example of application of this new algorithm is given.

2. GRADIENT COMPUTATION FOR SAR IMAGES

2.1. State of the art

Many works have been done to overcome the problem of using gradient by difference on SAR images, especially for edge detection applications. The use of the ratio instead of the difference is more suitable to multiplicative noise. The Ratio of Average (ROA) [6] consists in computing the ratio of local means on opposite sides of the studied pixel along one direction $R_i = \frac{M1}{M2}$ (Figure 2a), then in computing $T_i = \min(R_i, \frac{1}{R_i})$. After computing T_x, T_y, T_{d1} and T_{d2} along the four main directions (Figure 2b), the gradient magnitude and orientation are defined as $G_n = \min(T_x, T_y, T_{d1}, T_{d2})$ and $G_t = \underset{i}{\operatorname{argmin}}(T_x, T_y, T_{d1}, T_{d2})$.

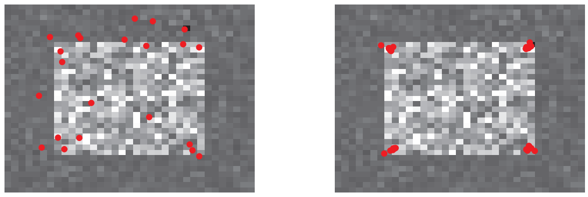


(a) Scheme of the ratio of local means

(b) Four main directions

Fig. 2: Scheme of the ROA method

The Ratio of Exponentially Weighted Averages (ROEWA) [7] is an improvement of the ROA for a multi-edge context, obtained by computing exponentially weighted local means. Those operators have been designed for edge detection and give a good estimate of the gradient magnitude. However, they do not directly allow to compute orientations. Also, computing ratios along several directions is time consuming.



(a) LoG method applied on the logarithm image (22 keypoints)

(b) SAR-Harris method (32 keypoints)

Fig. 3: Detection of keypoints on a speckle noised square with LoG method applied on the logarithm of the image and the SAR-Harris method.

By defining the vertical and horizontal gradient as $D_x = T_x$ and $D_y = T_y$, and by analogy to the gradient-based edge detector for optical image, estimates of gradient magnitude and orientation are proposed in [4] as $D_n = \sqrt{(D_x^2 + D_y^2)}$ and $D_t = \arctan(\frac{D_y}{D_x})$. However this definition of orientation is questionable. For example, if we compute the gradient on a vertical edge with reflectivities m_a and m_b ($m_a < m_b$), we get $D_x = \frac{m_a}{m_b}$, $D_y = 1$ and $D_t = \arctan(\frac{m_b}{m_a})$. Therefore the gradient orientation takes arbitrary values depending on the reflectivities of the areas, while it is expected to be zero. T_x and T_y should not be used directly to compute the gradient orientation.

2.2. Proposed approach

We propose here to define the vertical and horizontal gradient as $G_{x,\alpha} = \log(R_{x,\alpha})$ and $G_{y,\alpha} = \log(R_{y,\alpha})$, and to compute the gradient magnitude and orientation in the usual way: $G_{n,\alpha} = \sqrt{(G_{x,\alpha}^2 + G_{y,\alpha}^2)}$ and $G_{t,\alpha} = \arctan(\frac{G_{y,\alpha}}{G_{x,\alpha}})$. α, being the parameter of the exponential weight used to compute the local means. With this method, on a vertical edge, we obtain $G_{x,\alpha} = \log(m_a) - \log(m_b)$, $G_{y,\alpha} = 0$ and $G_{t,\alpha} = 0$, as expected. Also, not taking the minimum between the ratio and its inverse gives the possibility to obtain negative and positive gradient values, and therefore allows to use the whole possibilities of orientation values. We propose to use this gradient computation method, that we call Gradient by Ratio (GR), to adapt the SIFT algorithm to SAR images.

3. SIFT ALGORITHM ADAPTED TO SAR IMAGES

3.1. Keypoints detection

3.1.1. Proposed approach

A first possible approach to the detection of keypoints on SAR images would be simply to apply the LoG method (extrema of the Laplacian in the Gaussian scale-space) on a suitably rescaled logarithm of the image. However, such an approach is not robust enough to noise and, as we will see in the next section, yields no improvements in comparison with the original LoG approach. The use of a Laplacian of Gaussian with second derivatives does not seem convenient and easy to adapt to multiplicative noise, which requires to compute ratios for better treatment. The approach we propose is based on the Harris detector [8] and the GR, presented in Section 2. We define the SAR-Harris matrix as:

$$C_{SH}(x, y, \beta) = G_{\sqrt{2}\cdot\beta} \star \begin{bmatrix} (G_{x,\frac{1}{\beta}})^2 & (G_{x,\frac{1}{\beta}}) \cdot (G_{y,\frac{1}{\beta}}) \\ (G_{x,\frac{1}{\beta}}) \cdot (G_{y,\frac{1}{\beta}}) & (G_{y,\frac{1}{\beta}})^2 \end{bmatrix}$$

and the multi-scale SAR-Harris criterion as $R_{SH}(x, y, \beta) = \det(C_{SH}(x, y, \beta)) - d \cdot \mathrm{tr}(C_{SH}(x, y, \beta))$. For this method, a multi-scale representation of the original image is built by computing the SAR-Harris criterion at different scales $\beta_k = \beta_0 \cdot c^l$ with $l = 0..l_{max} - 1$. Then, local extrema in space are selected at each level to be keypoints candidate. Edge and low contrast points are suppressed with a threshold t_{SH} on the SAR-Harris criterion. Extracted keypoints are characterized by their position (x, y) and their scale β. This approach therefore merges the two steps of the usual keypoint detection method in order to suppress the use of second order derivatives. We use the following parameters: $d = 0.04$, $\beta_0 = 2$, $c = 2^{1/3}$ and $l_{max} = 8$, being the number of scales. t_{SH} has been set to 0.8 after a study of the probability distribution of the SAR-Harris criterion computed on corners, borders and homogeneous areas.

The example of a speckle noised square on Figure 3b shows indeed the efficiency of this method: keypoints are only found on the corners, as expected, and there are no false detection.

3.1.2. Results

To evaluate the performance of these two methods compared to the original SIFT one, we use 18 pixel-registered pairs of TerraSAR-X images of size 512x512, supposing there is no temporal changes between the two images of each pair. On each image, keypoints are extracted with three different methods: LoG on the intensity image, LoG on the logarithm image and the introduced SAR-Harris criterion. In order to compare these methods, the repeatability of the corresponding detections is studied. For a given image of a pair, we look for the closest keypoint extracted with the same method on the other image of the pair. We observe, for different distances d, the percentage of keypoints that are repeated on the other image at a distance lower than d. Results are shown on Figure 4. The threshold t_H have been adapted to obtain on average the same number of keypoints detected with the SAR-Harris method. Retaining the minimum number of keypoints for each pair, we obtain a total of 25032 keypoints extracted with the original LoG method, 24729 keypoints when the logarithm of the image is used and 21253 keypoints with the SAR-Harris method.

The SAR-Harris method gives better performance than the two other ones. Indeed, more than 50% of the keypoints extracted with SAR-Harris are repeated on the respective pair at a distance less than 1.5 pixels, when the rate is only 30% for the keypoints extracted with the two other methods. The LoG keypoints detection method does not perform better when applied on the logarithm of the image rather than the amplitude.

3.2. Orientations Assignment and Descriptors Extraction

3.2.1. Proposed approach

We now consider the computation of the orientation and the descriptors associated with keypoints. Both are based on the histogram of the gradient orientations, weighted by the gradient magnitude, computed on a neighborhood of the studied keypoint. We propose to use the Gradient by Ratio (GR) method introduced in Section 2.2 to compute those histograms on SAR images. Except from the computation of the gradient, the remaining computations are left unchanged compared to the classical SIFT or its variations. In this paper, we use the variant of the descriptor introduced in [9], where a circular neighborhood is used. We call the resulting descriptor a Ratio Descriptor.

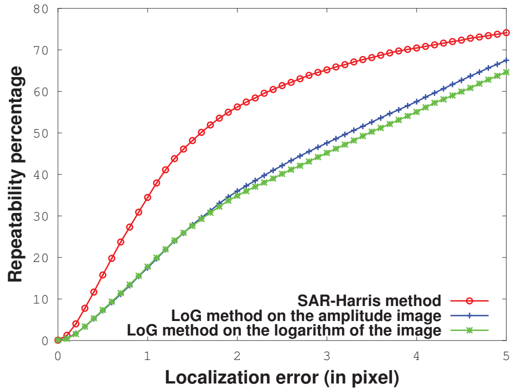


Fig. 4: Repeatability rate of keypoints, computed on 18 image pairs, with respect to the localisation error. Keypoints are extracted with three different methods : the LoG method applied on the amplitude image and on the logarithm of the image, and the SAR Harris method.

3.2.2. Results

The performance of both descriptors, the Ratio Descriptor and SIFT, are evaluated, along with two keypoints detection methods, the classical LoG and the newly introduced SAR-Harris. The four combinations are tested on the same pairs as in 3.1.2. For each situation, keypoints and descriptors are extracted on each image. The descriptors of the first image of each pair are then matched to the one of the other image by taking the closest match with the Euclidean distance. A match is defined as correct if the Euclidean distance between the matched keypoints, weighted by the minimum of the keypoints scales, is under an arbitrary threshold, which we set to 5 in our experiments. A varying threshold on the ratio between the Euclidean distance to the second and the first closest match [1] allows to suppress possible false matches and is used to compute Receiver Operating Characteristic (ROC) curves. Percentage of good matches among the entire set is computed with respect to the percentage of false matches on Figure 5.

We observe that the Ratio Descriptor along with the SAR-Harris keypoint detection method gives the best performance. We call the resulting algorithm *SAR-SIFT*. Almost 50% of keypoints are matched correctly without less than 1% of false alarms, when this rate is less than 30% for the other configurations. Also, using the SAR-Harris method with the SIFT descriptor already improves significantly the performance of the algorithm. In contrast, using the LoG method with the Ratio Descriptor offers a limited enhancement.

4. ILLUSTRATION OF THE SAR-SIFT ALGORITHM

As an application of our algorithm, we try to find a small extract of a TerraSAR-X image into another larger TerraSAR-X image with the same acquisition conditions. An example using the SAR-SIFT algorithm is shown on Figure 6a and the same experiment using the original SIFT algorithm is shown on Figure 6b. On Table 1, performances of the algorithms are summarized and illustrate the superiority of the SAR-SIFT algorithm.

This algorithm could also be used for precise co-registration. A set of matches between keypoints from two images can help to compute a polynomial or rational transform between images using mean-square regression and can be used to register images. The use

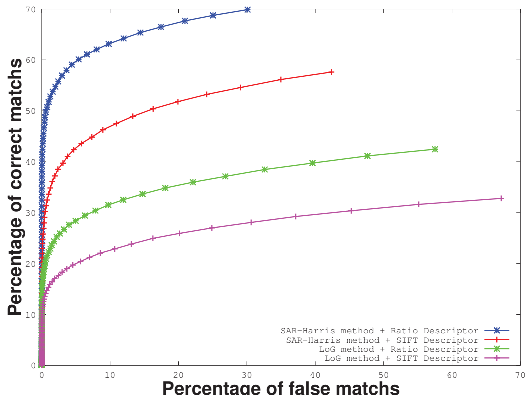


Fig. 5: Global ROC curves, computed on 18 image pairs, to evaluate the performance of the Ratio and SIFT descriptor, and the LoG and SAR-Harris methods. Histograms are computed on 12 bins.

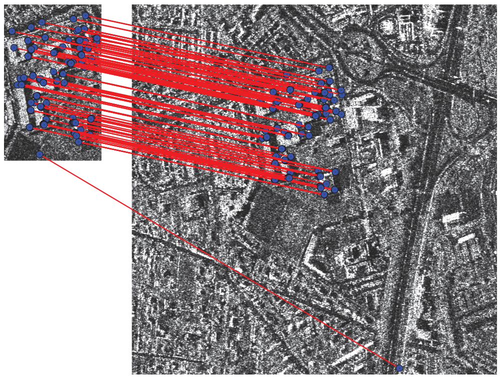
<table>
<tr><td colspan="2"></td><td colspan="2">Keypoints detection method</td></tr>
<tr><td colspan="2"></td><td>LoG</td><td>SAR-Harris</td></tr>
<tr><td rowspan="2">Descriptor</td><td>SIFT</td><td>16</td><td>146</td></tr>
<tr><td>Ratio</td><td>41</td><td>223</td></tr>
</table>

Table 1: Number of correct matches obtained for just one false for the example on Figure 6a and 6b with different keypoints detection methods and type of descriptor

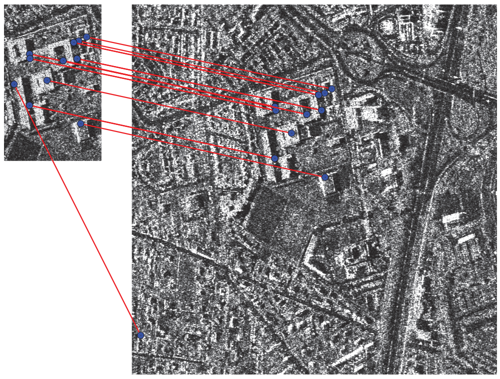
of the RANSAC algorithm can then help to be more tolerant to false matches and improve the registration.

5. CONCLUSION

An improved SIFT algorithm for SAR images, called SAR-SIFT, has been proposed, using a new keypoints detection method, an orientation computation more adapted to speckle noise, and a new descriptor that gives better performance than the usual SIFT approach on SAR images. The efficiency of the SAR-SIFT algorithm has been validated using several matching experiments. With this adaptation of the SIFT algorithm to SAR images, we hope to be able to use the full range of SIFT based methods, such as object recognition and tracking, stiching or registration, on SAR images. In a future work, we will also consider the use of keypoints for change detection in SAR high spatial resolution time series.



(a) Matching with the SAR-SIFT algorithm. 223 correct matches are found for a single false match.



(b) Matching with the usual SIFT algorithm. 16 correct matches are found for a single false match.

Fig. 6: Matching of a TerraSAR-X patch into a larger TerraSAR-X image with the same acquisition conditions with the SAR-SIFT algorithm and the usual SIFT algorithm.

6. REFERENCES

[1] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, pp. 91–110, 2004.

[2] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 10, pp. 1615 –1630, oct. 2005.

[3] Sahil Suri, Peter Schwind, Peter Reinartz, and Johannes Uhl, "Combining mutual information and scale invariant feature transform for fast and robust multisensor SAR image registration," in *75th Annual ASPRS Conference*, 2009.

[4] Sahil Suri, Peter Schwind, Johannes Uhl, and Peter Reinartz, "Modifications in the SIFT operator for effective SAR image matching," *International Journal of Image and Data Fusion*, vol. 1, no. 3, pp. 243–256, 2010.

[5] Yves Dufournaud, Cordelia Schmid, and Radu P. Horaud, "Matching images with different resolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA*. 2000, pp. 612–618, IEEE Computer Society Press.

[6] R. Touzi, A. Lopes, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 26, no. 6, pp. 764 –773, nov 1988.

[7] R. Fjortoft, A. Lopes, P. Marthon, and E. Cubero-Castan, "An optimal multiedge detector for SAR image segmentation," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 36, no. 3, pp. 793 –802, may 1998.

[8] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, 2001, vol. 1, pp. 525 –531 vol.1.

[9] J. Rabin, J. Delon, and Y. Gousseau, "A statistical approach to the matching of local features," *SIAM J. Img. Sci.*, vol. 2, pp. 931–958, September 2009.