SAR-SIFT: A SIFT-Like Algorithm for SAR Images

Flora Dellinger, Julie Delon, Yann Gousseau, Julien Michel, and Florence Tupin

Abstract—The scale-invariant feature transform (SIFT) algorithm and its many variants are widely used in computer vision and in remote sensing to match features between images or to localize and recognize objects. However, mostly because of speckle noise, it does not perform well on synthetic aperture radar (SAR) images. In this paper, we introduce a SIFT-like algorithm specifically dedicated to SAR imaging, which is named SAR-SIFT. The algorithm includes both the detection of keypoints and the computation of local descriptors. A new gradient definition, yielding an orientation and a magnitude that are robust to speckle noise, is first introduced. It is then used to adapt several steps of the SIFT algorithm to SAR images. We study the improvement brought by this new algorithm, as compared with existing approaches. We present an application of SAR-SIFT to the registration of SAR images in different configurations, particularly with different incidence angles.

Index Terms—Remote sensing, scale-invariant feature transform (SIFT), synthetic aperture radar (SAR), SAR image registration.

I. INTRODUCTION

THE last generations of Earth observation satellites are providing a large amount of high-resolution data, both in optical and synthetic aperture radar (SAR) domains, resulting in the multiplication of multisensor, multiresolution, and/or multiangle contexts. To jointly exploit these data for classification, 3-D reconstruction, rapid mapping, or change detection, feature-based approaches with some particular invariances may be more suitable than pixel-based ones. In this paper, we introduce an algorithm for the extraction of local descriptors from SAR images, enabling the application of feature-based approaches to such images.

The scale-invariant feature transform (SIFT) [1], [2] is a classical algorithm for interest points detection and local features description. Due to its efficiency [3], it is widely used in the field of computer vision to match images or to localize and recognize objects. Its invariances to scale changes, rotations, and translations and its robustness to illumination changes and affine distorsions make it suitable for different kinds of applications such as object retrieval, image indexing, stitch-

ing, registration, or video tracking. The outline of the SIFT algorithm and some of its variants will be presented later in Section II-A.

The SIFT algorithm and its variants are an interesting option for remote sensing images due to their efficiency for image matching and their invariances. Such algorithms have been applied mostly to satellite optical images since they have characteristics similar to images acquired with standard cameras. Several registration methods [4]–[6] use SIFT keypoints as control points to estimate deformation models. Li *et al.* [4] took into account the specificity of remote sensing images and introduce a new matching criterion with scale and orientation restrictions. A multilevel SIFT matching approach is proposed by Huo *et al.* [5] to register very high resolution images, with the help of RANdom Sample Consensus (RANSAC). Sedaghat *et al.* [6] adapted the algorithm to obtain keypoints that are uniformly distributed in space and filter the mismatches by applying a projective model. The SIFT algorithm also has some assets for the retrieval of remote sensing images or for classification applications. Yang and Newsam [7] used bag-of-word (BoW) representation of SIFT descriptors to perform image retrieval of land use/land cover classification in high-resolution imagery. Image classification is performed by Risojevic and Babic [8] using SIFT descriptors within a BoW representation. Object detection is another application field of the SIFT algorithm. Single buildings are detected on very high resolution optical images by Sirmacek and Unsalan [9] by using SIFT keypoints, multiple subgraph matching, and graph cut methods. Tao *et al.* [10] performed airport detection by considering both clustered SIFT keypoints and region segmentation.

While the SIFT algorithm has proven its efficiency for various kinds of applications in optical remote sensing, the situation is different for SAR images. SAR is an active system and has the advantage of acquiring images independently of weather conditions and solar illumination. SAR images are frequently used in disaster situations since they are often the first to be available. However, SAR images are corrupted by a strong multiplicative noise, which is called speckle, and data processing is thus made difficult. In particular, the SIFT algorithm does not perform well on this type of images. Considering the potential applications of local features in the context of SAR imagery, it is therefore of primary importance to adapt SIFT-like algorithms to this context. In particular, this would be potentially useful for all applications where robust and invariant image comparisons are needed (registration, change detection, template-based object detection, etc.).

Several improvements have been proposed to adapt the algorithm to this case. Some suggest to prefilter [11] or denoise [12] the images to reduce the influence of speckle noise. Others remove some invariances [13], [14] or modify some steps of

F. Dellinger, Y. Gousseau, and F. Tupin are with Institut Mines-Télécom, Télécom ParisTech, CNRS LTCI, 75013 Paris, France (e-mail: flora.dellinger@telecom-paristech.fr; yann.gousseau@telecom-paristech.fr; florence.tupin@telecom-paristech.fr).

J. Delon is with the Mathématiques Appliquées à Paris 5, Université Paris Descartes, 75006 Paris, France.

J. Michel is with the Centre National d'Etudes Spatiales, 31400 Toulouse, France.

the algorithm [15], [16] to improve the performances. Spatial relationships between keypoints are considered by Lv *et al.* [17] and Fan *et al.* [18] to suppress false correspondences. To limit the search space, Wessel *et al.* [19] used digital elevation model (DEM) and orbit information, and Xiaoping *et al.* [20] performed a manual preregistration. Details and limitations of these algorithms will be developed later in Section II-C. Performances of these newly developed algorithms are still relatively limited, and the number of correct matches is not sufficient to consider other applications than image registration. In particular, most of them do not consider the statistical specificities of speckle noise.

In this paper, we propose a new algorithm for the extraction of local descriptors, adapted to the statistical specificities of SAR images. This algorithm is largely inspired by the SIFT algorithm and will be referred to as SAR-SIFT. Section II presents the outline of the classical SIFT algorithm and some of its variants, shows some of their limitations on SAR images, and comments on previous attempts to overcome these limitations. Section III introduces a new gradient definition and the SAR-SIFT algorithm, both adapted to SAR images. Experimental validations and performances are presented in Section IV. Finally, Section V investigates the possibilities offered by this new algorithm in the case of multiangle and multiresolution SAR imaging. An application to SAR image registration and preliminary results for change detection are presented.

A preliminary version of this paper can be found in [21].

II. SIFT ALGORITHM, VARIANTS, AND SOME ADAPTATIONS TO SAR IMAGES

Here, the original SIFT algorithm and some of its variants are introduced. We also present its limitations when applied to SAR images and some of its adaptations to cope with such images.

A. Original SIFT Algorithm and Some Variants

The SIFT algorithm has been introduced by Lowe in 1999 [1] and later improved in 2004 [2] for the matching of local features in natural images. The algorithm combines two operators: a feature detector and a feature descriptor. The feature detection consists of selecting interest points. Descriptors are then computed to describe these features. The steps of the SIFT algorithm are described in the following paragraphs.

1) Keypoint Detection: First, keypoints are selected and characterized by their position (x, y), scale σ, and orientation θ, i.e.,

$$P(x, y, \sigma, \theta).$$

A difference of Gaussian (DoG) [2] scale space, as an approximation of the Laplacian of Gaussian (LoG) [22], is constructed with scales $\sigma_l = \sigma_0 \cdot r^l$ and $l \in [\![0, \ldots, l_{\max} - 1]\!]$. Local extrema in the three dimensions (x, y, σ) are then selected to obtain keypoints defined by their position and scale.

Candidates with low contrast or located on edges are filtered by a threshold on DoG values and a criterion based on the Hessian matrix, respectively [2]. Another possibility to filter points is to use a threshold on the multiscale Harris corner

detector, based on the Harris matrix [23]. The Harris matrix and corner detector are respectively defined as

$$C(x, y, \sigma) = \sigma^2 \cdot \mathcal{G}_{\sqrt{2} \cdot \sigma} * \begin{bmatrix} \left(\frac{\partial I_\sigma}{\partial x}\right)^2 & \left(\frac{\partial I_\sigma}{\partial x}\right) \cdot \left(\frac{\partial I_\sigma}{\partial y}\right) \\ \left(\frac{\partial I_\sigma}{\partial x}\right) \cdot \left(\frac{\partial I_\sigma}{\partial y}\right) & \left(\frac{\partial I_\sigma}{\partial y}\right)^2 \end{bmatrix}$$

$$R(x, y, \sigma) = \det\left(C(x, y, \sigma)\right) - t \cdot \operatorname{tr}\left(C(x, y, \sigma)\right)^2 \quad (1)$$

where $\mathcal{G}_{\sqrt{2} \cdot \sigma}$ is a Gaussian kernel with standard deviation $\sqrt{2} \cdot \sigma$, $*$ is the convolution operator, I_σ is the convolution of the original image by a Gaussian kernel with standard deviation σ, and t is an arbitrary parameter.

Among other interest point detectors, we can cite Harris-Laplace [24], which localizes points in space as extrema of the multiscale Harris function and in scale as maxima of the LoG. More accurate localization is achieved by Hessian-Laplace [25] by replacing the space selection with local maxima of the Hessian determinant. To achieve affine invariance, Harris-affine [26] and Hessian-affine [25] detectors refine the localization with an iterative adaptation based on the second-order derivative matrix. A review of local feature detectors can be found in [27].

In this paper, we will compare the proposed approach for keypoint detection (see Section III-B1) to keypoints that are detected as local extrema (in (x, y, σ)) in the LoG scale space. These points will be filtered by the multiscale Harris criterion to eliminate those lying on edges or low contrasted. We will refer to this approach as the LoG-Harris detector.

2) Orientation Assignment: To determine the orientations θ associated with each keypoint, Lowe [2] suggested to compute a local histogram of gradient orientations, weighted by the gradient magnitudes and a Gaussian window. The histogram is computed on a scale-dependent neighborhood. Every peak above 80% of the maximum is selected as an orientation. We can obtain several keypoints with the same position (x, y) and scale σ but with different orientations θ.

In the variant descriptor Speeded Up Robust Features (SURF) [28], the histogram is replaced by Haar wavelet responses in the x- and y-directions, and the sum of responses is computed within a sliding orientation window to estimate the principal orientation. In [29], it is proposed to compute principal orientations by applying an *a contrario* mode selection to the orientation histogram. This method allows improving the robustness of the orientation computation, with regard to the quantification process.

3) Descriptors Extraction: A SIFT descriptor is assigned to each keypoint $P(x, y, \sigma, \theta)$ to describe its local geometry. A square neighborhood is defined around each point with a size depending on σ to obtain translation and scale invariance. It is then rotated by an angle $-\theta$ to ensure rotation invariance. This normalized neighborhood is divided into 4×4 square sectors, upon which histograms of the gradient orientations, weighted by the gradient magnitudes, are computed. Before computing these histograms, gradient magnitudes are weighted by a Gaussian function defined over the local neighborhood. For each keypoint, the SIFT descriptor is obtained by concatenating and normalizing these histograms.
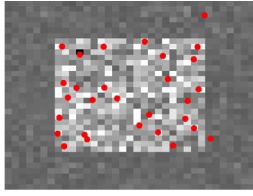
Fig. 1. Detection of keypoints on a rectangle corrupted by speckle noise with the LoG-Harris detector applied on the amplitude image (29 keypoints detected).

Different adaptations of the SIFT descriptor have been proposed in the literature. Principal component analysis (PCA)-SIFT [30] is obtained by applying PCA on normalized gradient neighborhood. Gradient location and orientation histogram [3] is computed on a log-polar grid and upon 17 sectors, the size of the resulting vector being reduced with PCA. SURF [28] is obtained by replacing gradient histograms by sums of Haar wavelet responses in vertical and horizontal directions. To achieve affine invariance, ASIFT (Affine-SIFT) [31] simulates different viewpoints to evaluate two camera axis orientation parameters. Recent reviews of local feature extraction and matching may be found in [27], [32], and [33].

In this paper, following [29], we choose to use a SIFT-like descriptor computed on a log-polar grid of nine sectors, which offers better performances than a square one [3]. In addition, we do not use a Gaussian weighting of the gradient magnitudes when computing histograms, since we observed similar performances with and without this weighting. This descriptor will be further called *circular-SIFT*.

B. Limitations of SIFT-Like Algorithms on SAR Images

The SIFT algorithm and its variants do not perform well on SAR images. This is no surprise, since the original SIFT algorithm was designed to detect structures that are specific to optical photography, such as blobs, in relatively low noise conditions. Many false keypoint detections and false matches occur. In particular, speckle noise leads to numerous false detections with classical keypoint detectors such as the LoG-Harris detector (see Fig. 1). On optical images, noise is usually relatively weak, and keypoint filtering steps such as the multi-scale Harris criterion [23] suppress most of the false detections due to their contrast dependence. However, SAR images present a large dynamic range, and the multiplicative noise leads to stronger gradient magnitude on homogeneous areas with high reflectivity than on areas with low reflectivity [see Fig. 2(b)]. False alarms on high-contrast areas are thus not suppressed, as shown on the example of a rectangle corrupted by speckle noise (see Fig. 1).

The orientations and, therefore, the local descriptors are also not robust to multiplicative noise, since their computation relies on a classical gradient by difference.

C. Previous Adaptations of the SIFT Algorithm to SAR Images

Modifications of the SIFT algorithm for SAR images have been already proposed in the literature. Some suggest to simplify the algorithm, by skipping the smallest scales for the keypoint detection [13] or by suppressing the orientation as-
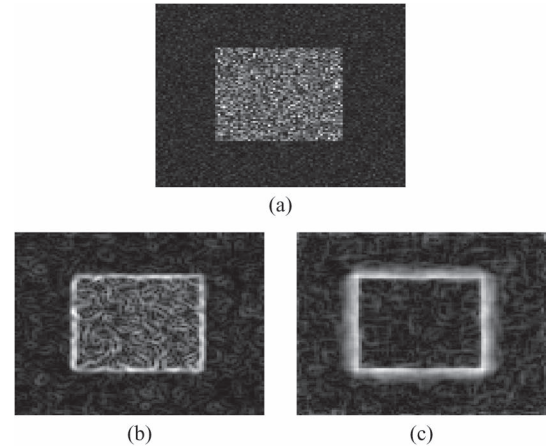

Fig. 2. Example of a rectangle corrupted by speckle noise and its gradient magnitude for two gradient computation methods. (a) Rectangle corrupted by speckle noise. (b) Gradient by difference. (c) Gradient by ratio [see (9) and (10)].

signment [14], [18]. While such a procedure does decrease the number of false detections since many occur at those scales, the remaining keypoints are still not precisely located. Suppressing the orientations limits the capability of the algorithm to match images with different viewing conditions.

To improve the algorithm, some steps can be adapted. In [16], intensity values are thresholded to obtain spatially uniformly distributed keypoints, and the size of the region descriptor is extended to increase matching performances. However, this limits the distinctiveness of descriptors and prevents the application of the algorithm on images with strong changes. In [15], a new pyramid with progressive downsampling is used for keypoint detection, and the SIFT descriptor is replaced by an improved version of Shape Context. While being faster, this new algorithm has lower performances than the original SIFT.

Some works propose to reduce the influence of speckle by replacing the Gaussian scale space with an anisotropic one [34] or by computing multilooks [19], [20]. However, this last process decreases image resolution and causes loss of information. Another solution is to denoise the images: curvelet transformation [11] or infinite symmetric exponential filter [12] can be used. Denoising is time consuming and can create artefacts that disturb the algorithm. While the performances of these algorithms are better than those of the original SIFT algorithm directly applied to SAR images, the number of correct matches is usually low (on the order of a few dozens).

Other studies suggest to improve performances by rejecting outliers. Lv *et al.* [17] divided the images into four subregions and consider the spatial relationships of the matched keypoints in every subregion. This, however, implies that the images represent the same scenes with nearly full overlap and no rotation. For a registration application, image transformation is estimated in [18] based on best correspondences, but a restrictive deformation is used. Wu and Yang [35] combined the SIFT algorithm and the cluster reward similarity measure to iteratively estimate an affine transformation. The process is time consuming and restricted to image registration.

The search space can be limited by performing a manual preregistration [20]. False correspondences can be removed by knowing orbit information, even if not precise [19], and a DEM.
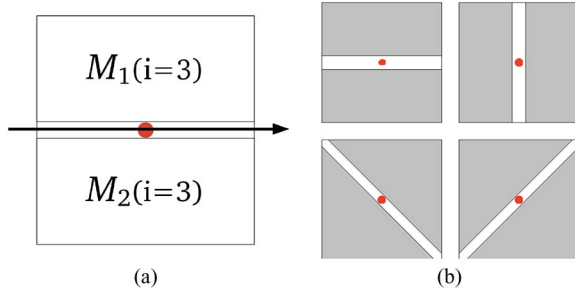
Fig. 3. Scheme of the ROA method [36], [37]. (a) Scheme of the ratio of local means for the direction $i = 3$, corresponding to the orientation $\pi/2$. (b) Four main directions, to compute T_3, T_1, T_2, and T_4, respectively, corresponding to the orientations $\pi/2$, 0, $\pi/4$, and $3\pi/4$.

Fig. 4. Exponential filter for computation of weighted means.

However, these kinds of information are not always known, and manual preregistration is time consuming and subject to interpretation errors.

To adapt the SIFT algorithm to SAR images, it is necessary to take into account the statistical specificities of SAR images. We suggest to develop first a new gradient computation due to which both the magnitude and the orientation are robust to speckle noise. Several steps of the algorithm can be then adapted to SAR images. A new keypoint detector is introduced, as well as a new orientation assignment and a SAR-adapted feature descriptor.

III. PROPOSED METHOD: NEW GRADIENT COMPUTATION AND SIFT-LIKE ALGORITHM FOR SAR IMAGES

A. Gradient Computation for SAR Images

1) State of the Art: Many works on edge detection have underlined the problem of using gradient by difference on SAR images. Indeed, variances of the gradient components depend on the underlying reflectivities [36]. Traditional approaches in edge detection consist in thresholding the gradient magnitude. For SAR images, this leads to higher false alarm rates in homogeneous areas of high reflectivity than in the ones of low reflectivity. The classical gradient by difference is, thus, not a constant false alarm rate operator. Statistical studies [36]–[38] have shown that the use of ratio is more suitable to multiplicative noise than the use of difference. Several edge detectors using ratio have been introduced in order to obtain a constant false alarm rate on SAR images:

- The ratio of average (ROA) [36], [37] consists in computing the ratio of local means on opposite sides of the studied pixel along one direction i [see Fig. 3(a)], i.e.,

$$R_i = \frac{M_1(i)}{M_2(i)}. \tag{2}$$

The ratio R_i is then normalized, i.e.,

$$T_i = \max\left(R_i, \frac{1}{R_i}\right). \tag{3}$$

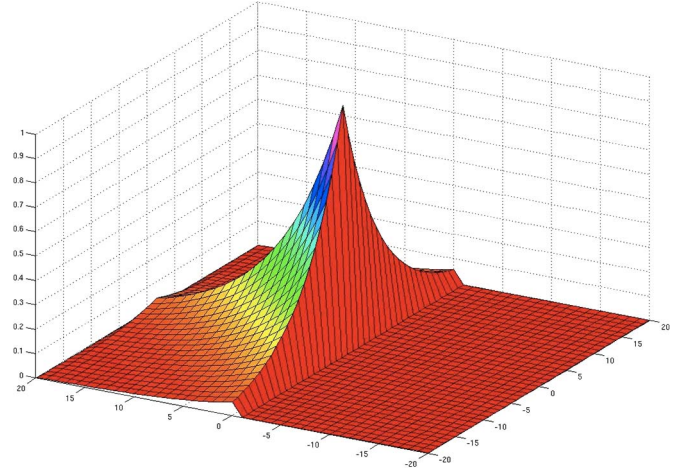These ratios are computed along the four main directions [see Fig. 3(b)]. The gradient magnitude D_n^1 and

orientation D_t^1 are defined as

$$D_n^1 = \max_i(T_i)$$
$$D_t^1 = \left(\arg\max_i(T_i) - 1\right) \times \frac{\pi}{4}. \tag{4}$$

Edges may be then obtained by thresholding the gradient magnitude D_n^1.

- The ratio of exponentially weighted averages (ROEWA) [39] is an improvement of the ROA for a multiedge context, which is obtained by computing exponential weighted local means (see Fig. 4). For example, given a point (a, b), the means are defined for the direction $i = 3$ as

$$M_{1,\alpha}(i = 3) = \int\limits_{x=R}\int\limits_{y=R^+} I(a + x, b + y) \times e^{-\frac{|x|+|y|}{\alpha}}$$
$$M_{2,\alpha}(i = 3) = \int\limits_{x=R}\int\limits_{y=R^-} I(a + x, b + y) \times e^{-\frac{|x|+|y|}{\alpha}} \tag{5}$$

where α is the exponential weight parameter.

As in the ROA, the ratio and its normalization for a direction i are defined as

$$R_{i,\alpha} = \frac{M_{1,\alpha}(i)}{M_{2,\alpha}(i)}$$
$$T_{i,\alpha} = \max\left(R_{i,\alpha}, \frac{1}{R_{i,\alpha}}\right). \tag{6}$$

These ratios $T_{i,\alpha}$ are computed along the horizontal $(i = 3)$ and vertical $(i = 1)$ directions. By analogy to the edge detectors on optical images that are based on gradients, the edge image is obtained by

$$D_{n,\alpha}^2 = \sqrt{(T_{1,\alpha})^2 + (T_{3,\alpha})^2}. \tag{7}$$

The ROEWA is more precise in a multiscale edge context and more robust to noise than the ROA, since the weighting parameter α allows adaptive smoothing of the image.
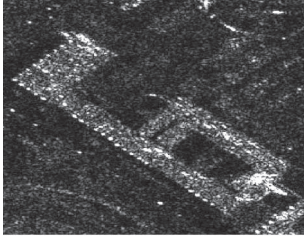
Fig. 5. Extract of Cosmo-SkyMed image with 34° of incidence angle and 1-m resolution.

Those operators have been designed for edge detection and provide a good estimate of the gradient magnitude. However, they do not give a precise measure of the gradient orientation since only a few directions are considered. This could be improved by increasing the number of directions, but it would be time consuming.

Suri *et al.* [14] proposed to define the horizontal and vertical gradient components as $T_{1,\alpha}$ and $T_{3,\alpha}$, respectively. By analogy to the gradient-based edge detector for optical images, they estimate the gradient magnitude and orientation as

$$D_{n,\alpha}^3 = \sqrt{(T_{1,\alpha})^2 + (T_{3,\alpha})^2}$$
$$D_{t,\alpha}^3 = \arctan\left(\frac{T_{3,\alpha}}{T_{1,\alpha}}\right). \quad (8)$$

This definition of orientation is highly questionable. Indeed, $T_{1,\alpha}$ and $T_{3,\alpha}$ always take positive values, and thus, $D_{t,\alpha}^3$ can only take values between 0 and $\pi/2$. Moreover, the gradient computation on a vertical edge with reflectivities m_a and $m_b(m_a < m_b)$ yields

$$T_{1,\alpha} = \frac{m_b}{m_a}$$
$$T_{3,\alpha} = 1$$
$$D_{t,\alpha}^3 = \arctan\left(\frac{m_a}{m_b}\right).$$

Therefore, the gradient orientation takes arbitrary values depending on the reflectivities of the areas, while it is expected to be equal to zero. Thus, unlike what Suri *et al.* suggested, the normalized ratios $T_{1,\alpha}$ and $T_{3,\alpha}$ should not be used to compute the gradient orientation.

2) Proposed Approach: We propose here to define the horizontal and vertical gradients as

$$G_{x,\alpha} = \log(R_{1,\alpha})$$
$$G_{y,\alpha} = \log(R_{3,\alpha}) \quad (9)$$

and to compute the gradient magnitude and orientation in the usual way as

$$G_{n,\alpha} = \sqrt{(G_{x,\alpha})^2 + (G_{y,\alpha})^2}$$
$$G_{t,\alpha} = \arctan\left(\frac{G_{y,\alpha}}{G_{x,\alpha}}\right) \quad (10)$$

where α is the parameter of the exponential weight used to compute the local means.
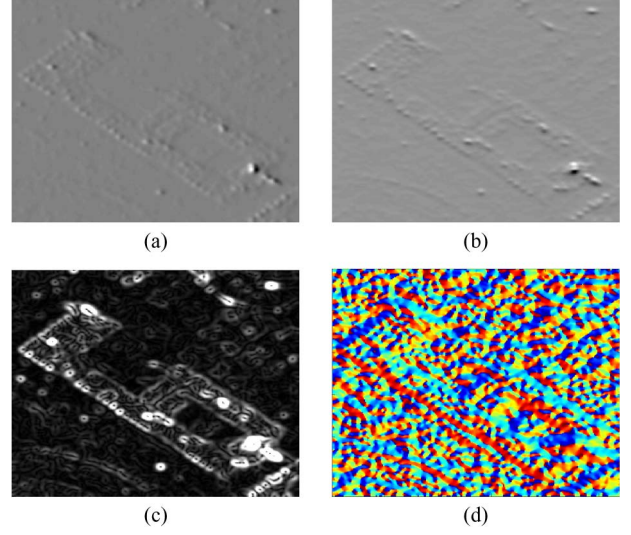

(a)   (b)   (c)   (d)

Fig. 6. Gradient by difference applied on the image in Fig. 5 with a Gaussian blur of $\sigma = 2$. (a) Horizontal component. (b) Vertical component. (c) Magnitude. (d) Orientation.
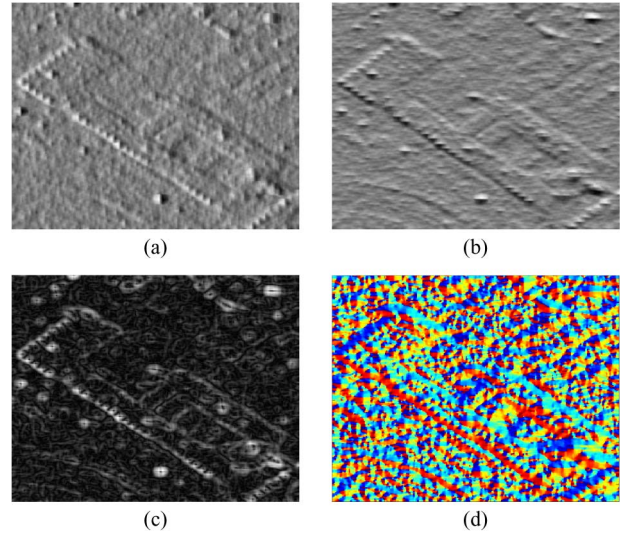

(a)   (b)   (c)   (d)

Fig. 7. New gradient computation (GR) applied on the image in Fig. 5 with $\alpha = 2$. (a) Horizontal component. (b) Vertical component. (c) Magnitude. (d) Orientation.

By using the logarithm, the aforementioned problem for the gradient orientation on a vertical edge is avoided, since the computation yields

$$G_{x,\alpha} = \log(m_a) - \log(m_b)$$
$$G_{y,\alpha} = 0$$
$$G_{t,\alpha} = 0.$$

There is no normalization with the minimum (or maximum) between the ratio and its inverse, in order to obtain negative and positive gradient values. With this approach, the whole possibilities of orientation values are taken into account. In addition, since the weighting parameter α allows smoothing the image at different scales, this gradient can be compared with the gradient by difference applied on an image with Gaussian blur. We call this new gradient computation method *gradient by ratio* (GR).

Fig. 2 presents the gradient magnitude on a rectangle corrupted by speckle noise for the two gradient computation methods. The GR method [see (9) and (10)] does not produce higher values on high-reflectivity areas than on the ones of low reflectivity, unlike the gradient by difference.

The gradient values of the image in Fig. 5 are presented in Fig. 6 for the gradient by difference and in Fig. 7 for the GR. It appears that the gradient by difference presents a better reduction to speckle noise, but gradient values (magnitude and vertical and horizontal components) are higher on high-reflectivity areas than on low-reflectivity ones. For the GR method, however, gradient responses are not higher on those areas.

This new gradient computation method will now help us to adapt the SIFT algorithm to SAR images.

B. SIFT-Like Algorithm Adapted to SAR Images

This new algorithm operates in a similar way as the original SIFT algorithm, with a feature detector followed by a feature descriptor. The outline of this algorithm, which we have called *SAR-SIFT*, is presented in Fig. 8.

1) Keypoint Detection: A first simple approach to detect keypoints on SAR images would be to apply the LoG-Harris detector on the logarithm of the image. This allows dealing with an additive noise instead of a multiplicative one, as well as suppressing the false detections on the high-reflectivity areas. Although appealing because of its simplicity, this approach is not robust enough to noise and does not improve much the performances of the original LoG-Harris approach (see Fig. 1). The example of a rectangle corrupted by speckle noise in Fig. 9(a) shows that keypoints are indeed found near corners, as expected, but badly located. There are fewer spurious detections (detections on homogeneous areas and on edges) than when directly applied on the amplitude image (see Fig. 1), and they also occur equally on high- and low-reflectivity areas, but they are still numerous. By adapting the parameters on the multiscale Harris criterion [23], the number of spurious detections can be decreased, but so will the number of correct ones.

LoG and Hessian matrices do not seem convenient and easy to adapt to multiplicative noise since they rely on second derivatives. The multiscale Harris function [24], in contrast, is based on the first derivative. From the new gradient computation adapted to SAR images that we have developed in Section III-A, we propose a new approach based on this detector.

The multiscale Harris matrix and function are defined in (1). Observe that the weight σ^2 is needed here for full-scale normalization [24]. In the LoG-Harris detector, the multiscale Harris criterion allows suppressing low-contrast and edge points by applying a threshold d_H on $R(x, y, \sigma)$.

Considering this definition and the GR, we propose the new multiscale SAR-Harris matrix and the new multiscale SAR-Harris function, respectively, as

$$C_{\mathrm{SH}}(x, y, \alpha) = \mathcal{G}_{\sqrt{2} \cdot \alpha} * \begin{bmatrix} (G_{x,\alpha})^2 & (G_{x,\alpha}) \cdot (G_{y,\alpha}) \\ (G_{x,\alpha}) \cdot (G_{y,\alpha}) & (G_{y,\alpha})^2 \end{bmatrix}$$
$$R_{\mathrm{SH}}(x, y, \alpha) = \det\left(C_{\mathrm{SH}}(x, y, \alpha)\right) - d \cdot \mathrm{tr}\left(C_{\mathrm{SH}}(x, y, \alpha)\right)^2$$
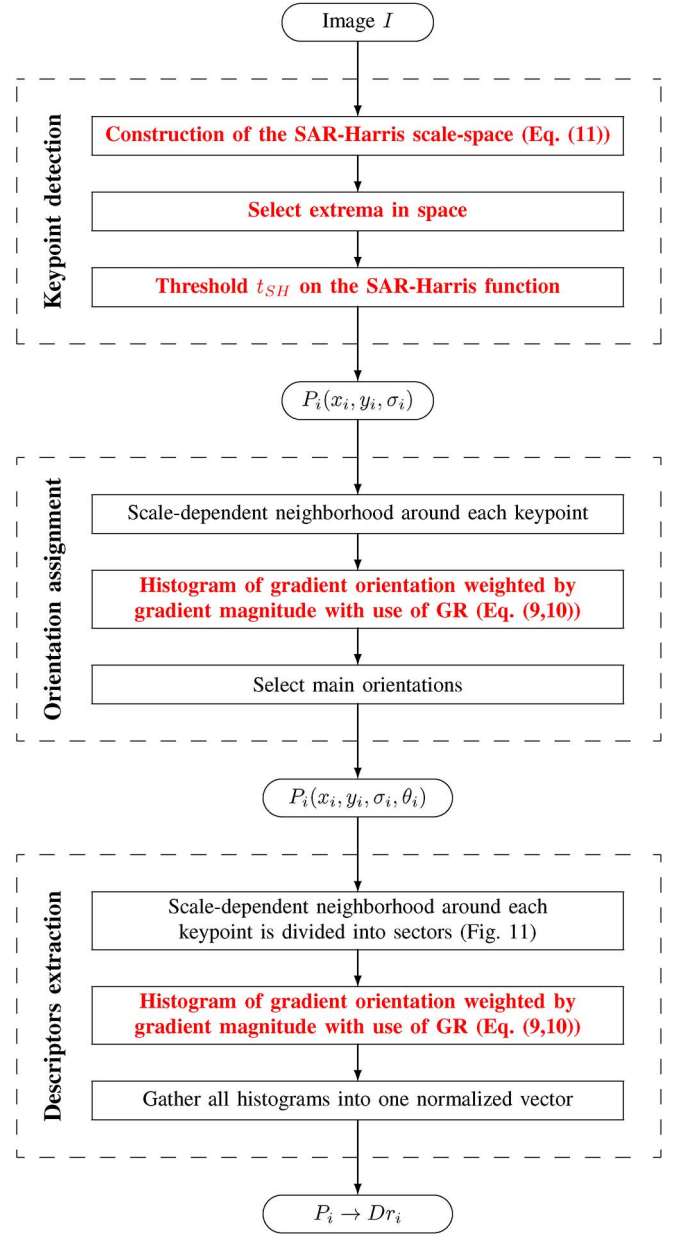$$\tag{11}$$

Fig. 8. Outline of the SAR-SIFT algorithm: keypoint detection (SAR-Harris), orientation assignment, and descriptor extraction. Contributions presented in this paper are in red.
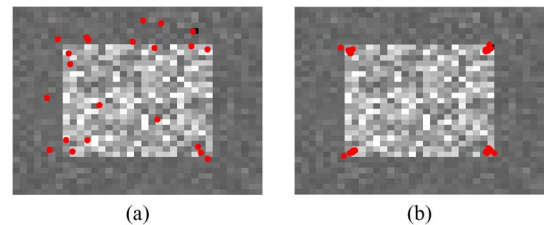

Fig. 9. Detection of keypoints on a rectangle corrupted by speckle noise with the LoG-Harris detector applied on the logarithm of the image and the SAR-Harris detector. (a) LoG-Harris detector applied on the logarithm image (22 keypoints). (b) SAR-Harris detector (32 keypoints).

where d is an arbitrary parameter and where the derivatives $G_{x,\alpha}$ and $G_{y,\alpha}$ are computed using (9).

In this case, it can be easily derived that the multiplication by σ^2 is not needed anymore to ensure scale invariance.
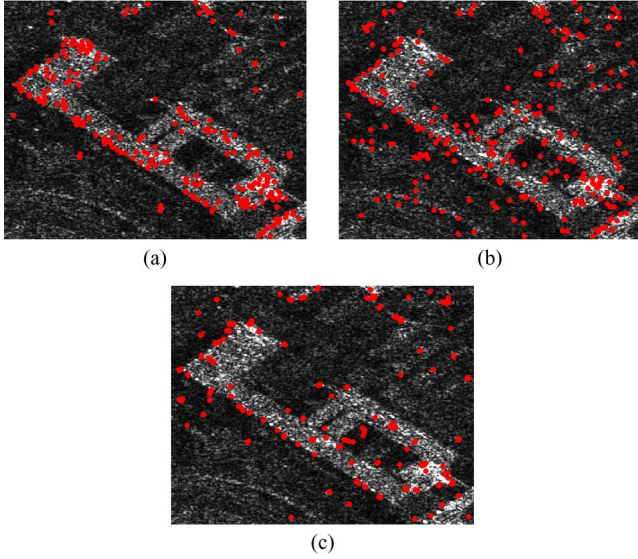
(a)

(b)

(c)

Fig. 10.   Detection of keypoints on the image in Fig. 5 with the LoG-Harris detector, applied on the amplitude image and the logarithm of the image, and the SAR-Harris detector. The thresholds d_H and d_{SH} were adjusted to obtain the same number of keypoints with all keypoint detectors. (a) LoG-Harris detector applied on the amplitude image (435 keypoints). (b) LoG-Harris detector applied on the logarithm of the image (435 keypoints). (c) SAR-Harris detector (433 keypoints).

For this keypoint detector, we replace the LoG scale space by a multiscale representation of the original image, which is obtained by computing the multiscale SAR-Harris function [see (11)] at different scales $\alpha_m = \alpha_0 \cdot c^m$, with $m \in [\![0, \dots, m_{\max} - 1]\!]$. Local extrema in space are then selected at each level to be keypoint candidates. Subpixel positions of the keypoints are refined by performing a bilinear interpolation of the SAR-Harris criterion around the local extrema. A threshold d_{SH} on the multiscale SAR-Harris function allows filtering edge and low-contrast points. We obtain keypoints characterized by their position (x, y) and their scale α.

This approach, called the SAR-Harris detector, merges the two steps of the LoG-Harris detector in order to avoid the use of second-order derivatives. As it is easily verified, it also has the advantage of being independent of the image dynamic.

We have noticed that this scale space rarely reaches extrema in three dimensions. This fact was also observed on optical images [24]. Only extrema in space are thus selected. Several detections can then occur at the same position but for different scales. However, some of them are suppressed by thresholding the multiscale SAR-Harris function.

The example of a rectangle corrupted by speckle noise in Fig. 9(b) shows the efficiency of this detector: keypoints are only found on the corners, as expected, and there are no spurious detections.

Fig. 10 presents an example of keypoint detection for different detectors. As expected, we observe that keypoints detected with the LoG-Harris detector applied on the amplitude image are mainly detected on high-reflectivity areas, and a lot of false detections occur on homogeneous areas with high radiometry. Concerning the keypoints detected with the LoG-Harris detector applied on the logarithm of the image, a lot of false detections are also found on homogeneous areas, but
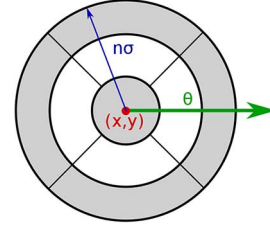

Fig. 11.   Scheme of the circular descriptor. The parameter n (set to 12 in this paper) relates the radius of the neighborhood to the scale of the keypoint. Ratio of inner to outer circle radius is 0.25 and 0.73, respectively.

they happen both on high- and low-reflectivity areas. However, the keypoints detected with the SAR-Harris detector are mostly located on corners and bright points, and the number of false detections on homogeneous areas is really low.

2) Orientation Assignment and Descriptor Extraction: In the original SIFT algorithm, both the steps of orientation assignment and descriptor extraction rely on histograms of gradient orientation. These histograms are computed on a neighborhood of each keypoint and weighted by the gradient magnitude and a Gaussian window. Here, we propose to use the GR method, introduced in Section III-A, to compute those histograms.

First, one or several principal orientations are computed for each keypoint. These are obtained from a local orientation histogram corresponding to a circular neighborhood (radius of 6σ). Following [29], we select the main modes of the local orientation histogram through to an a contrario mode selection [40]. Up to two orientations are selected at each point.

A descriptor is then computed for each orientation. Instead of using a square neighborhood and 4×4 square sectors as in the original SIFT descriptor, we rely on a circular neighborhood (radius of 12σ) and log-polar sectors as in [29] (see Fig. 11). Again, orientations are computed due to the gradient defined in Section III-A. The resulting descriptor is called *Ratio descriptor*. This descriptor is constructed the same way as the circular-SIFT descriptor, by concatenating the entries of the orientation histograms corresponding to the log-polar sectors. The only difference is the use of the GR instead of the gradient by difference. Let us observe that the GR method to compute the gradient could straightforwardly be adapted to other spatial configurations of sectors, such as the one of the original SIFT.

IV. EXPERIMENTAL VALIDATION OF THE SAR-SIFT ALGORITHM UNDER SPECKLE NOISE

Here, the proposed keypoint detectors and descriptors are compared with the LoG-Harris detector and to the circular-SIFT descriptor (see Section II-A3), respectively. We first investigate the stability and robustness of the keypoint detectors by measuring their repeatability rate. Then, performances of keypoint detectors and descriptors are evaluated with the help of receiver operating characteristic (ROC) curves. To only assess the ability of the algorithms to deal with speckle noise, the study is conducted on image pairs acquired under the same conditions.

A. Test Images and Parameters

For these experiments, we use 18 pairs of extracts of TerraSAR-X images with a subpixel registration, representing the city of Toulouse, France. All images have a size of 512×512 pixels and have been acquired under the same viewing conditions ($34°$ incidence angle, 2-m resolution, SpotLight mode). A visual check showed that no temporal changes occur between the two images of each pair, so that only the noise realization differs.

For the detection of keypoints with the LoG-Harris method (local extrema in LoG and threshold on the multiscale Harris criterion), we choose the following parameters to construct the scale space: $\sigma_0 = 0.63$, the first scale; $r = 2^{1/3}$, the ratio between the two scales; and $l_{\max} = 13$, the number of scales. The scales are then obtained with $\sigma_l = \sigma_0 \cdot r^l$ and $l \in [\![0, \ldots, l_{\max} - 1]\!]$. For the multiscale Harris criterion, the parameter t is set to 0.04, as suggested in [24]. The threshold on this criterion d_H is usually set to 2000 for 8-bit images, but will be adapted for each SAR image, since they have different dynamics.

For the SAR-Harris detector, the chosen parameters are as follows: $\alpha_0 = 2$, the first scale; $c = 2^{1/3}$, the ratio between the two scales; $m_{\max} = 8$, the number of scales; and $d = 0.04$, the arbitrary parameter of the SAR-Harris criterion. The scales are then obtained with $\alpha_m = \alpha_0 \cdot c^m$, with $m \in [\![0, \ldots, m_{\max} - 1]\!]$. We choose to select the same parameters as the multiscale Harris criterion, except for the first scale α_0 and the number of scales $m_{\max}$. Indeed, many false alarms occur at small scales. The threshold d_{SH} has been set to 0.8 after an experimental study of the probability distribution of the SAR-Harris criterion computed on corners, borders, and homogeneous areas.

For both the Ratio and circular-SIFT descriptors, histograms are computed on 12 bins, and log-polar grid is used with $k = 12$ (see Fig. 11).

B. Repeatability of Keypoint Detection

The repeatability criterion [41] gives a measure of the stability of keypoint detection, with regard to the image changes. Given a pair of registered images, we look, for each keypoint of the first image, at the closest one extracted on the other image with the same detector. Then, for different thresholds u, the percentage of keypoints repeated on the other image at a distance lower than u is observed.

The new SAR-Harris keypoint detector presented in Section III-B1 is compared with the LoG-Harris detector[1] applied on either the amplitude image or the logarithm image. Results are shown in Fig. 12. The thresholds d_H have been adapted to obtain on average the same number of keypoints with the SAR-Harris detector. We obtain for the entire set a total of 25 032 keypoints extracted with the original LoG-Harris detector, 24 729 keypoints when the logarithm of the image is used, and 21 253 keypoints with the SAR-Harris detector. The keypoint density is thus almost the same for each detector.

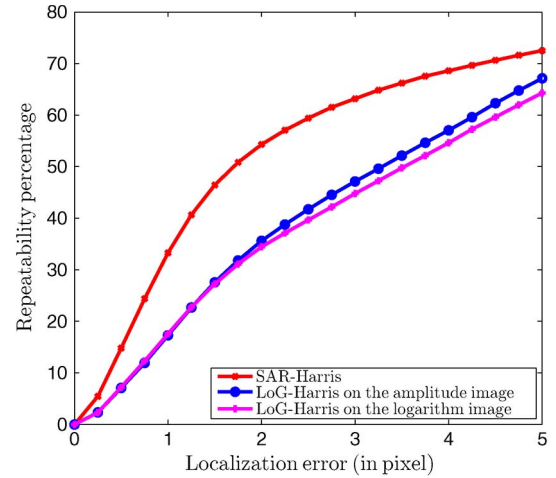[1]Comparisons with the corresponding DoG detectors yield very similar results.

Fig. 12. Repeatability rate of keypoints, computed on 18 image pairs, with respect to the localization error. Keypoints are extracted with three different detectors: the LoG-Harris detector applied on the amplitude image and on the logarithm of the image and the SAR-Harris detector.

It can be observed that the SAR-Harris detector performs better than the two LoG-Harris detectors. For example, at a localization error of $u = 1.5$ pixels, more than 50% of the keypoints extracted by the SAR-Harris detector are repeated. This rate is only 30% for the keypoints extracted with the two other detectors. We also observe that the performances of the LoG-Harris detector do not present an improvement when applied on the logarithm image rather than on the amplitude one.

C. Matching Performances

Global ROC curves are computed for different combinations of keypoint detectors and types of descriptor.

Keypoints of two different images are matched according to their respective descriptors. Different matching criteria exist in the literature, but the most commonly used is the nearest neighbor distance ratio (NNDR) method [2]. First, Euclidean distances are computed between one descriptor and the ones of the other image, and the nearest neighbor is chosen. To filter false matches, distances to the closest and second closest neighbor are compared. A threshold th is applied on the ratio of those respective distances. We will further call the first step as the nearest neighbor (NN) step and the second as the distance ratio (DR) step.

Another approach [29] consists in computing a probability of false alarm for all possible matches using an a contrario method. This allows different matches for one keypoint and permits to recognize multiple occurrences of one object.

Here, we will use the NNDR method.

Let $M_k(P_{1,i}, P_{2,j})$ be a match between a point $P_{1,i}(x_{1,i}, y_{1,i}, \sigma_{1,i}, \theta_{1,i})$ of an image I_1 and a point $P_{2,j}(x_{2,j}, y_{2,j}, \sigma_{2,j}, \theta_{2,j})$ of an image I_2. Considering the deformation T of the image I_1 in comparison with the image I_2, M_k is defined as correct if

$$\|T(x_{1,i}, y_{1,i}) - (x_{2,j}, y_{2,j})\|_2 < t_1 \cdot \min(\sigma_1, \sigma_2). \quad (12)$$

The parameter t_1 was empirically set up to 5. Here, the image pairs are registered with a subpixel precision; thus, there is no deformation between the images, and T is the identity function.
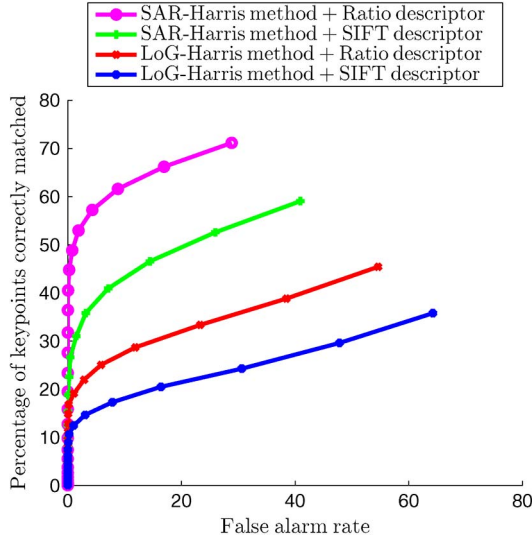
Fig. 13. Global ROC curves, computed on 18 image pairs, to evaluate the performances of the Ratio and circular-SIFT descriptors and the LoG-Harris and SAR-Harris detectors.

The quantities $\#CM_{all}$ and $\#FM_{all}$ are defined as the total number of correct and false matches for the entire set with the NN step, respectively. The numbers of correct and false matches $\#CM$ and $\#FM$, respectively, are evaluated for a certain value of the threshold th on the ratio of the distance to the closest and second closest matches (DR step). To obtain ROC curves, the percentage of correctly matched keypoints $\#CM/(\#CM_{all} + \#FM_{all})$ is plotted against the false alarm rate $\#FM/(\#CM + \#FM)$ by varying th.

We compare here two feature detectors, i.e., the LoG-Harris detector on the amplitude image and the new SAR-Harris detector, and two feature descriptors, i.e., the proposed Ratio descriptor and the circular-SIFT descriptor. The results of the four considered situations are displayed in Fig. 13. Since the images do not present any rotation, no orientations are assigned to the keypoints for this experiment, neither for the SIFT nor for the SAR-SIFT approach. The L^1 and L^2 distances are tested to compute similarities between descriptors at the matching step. We have observed that results with the L^1 distance are always better than with the Euclidean distance. To simplify the reading, only matches with the L^1 distance are displayed.

We observe that the best performance is achieved by the combination of the SAR-Harris detector and the Ratio descriptor. Indeed, for a false alarm rate of 1%, almost 50% of the possible correct matches are obtained, when for the other configurations this rate is less than 30%. The 1% false alarm rate is the percentage of false matches among the correspondences obtained with a certain value of th and not among all the possible matches. In addition, using the SAR-Harris detector with the circular-SIFT descriptor already significantly improves the performances of the algorithm. In contrast, the use of the LoG-Harris detector with the Ratio descriptor offers a limited enhancement.

In summary, the SAR-Harris detector is more stable and robust to noise than the LoG-Harris detector, and the Ratio descriptor outperforms the circular-SIFT Descriptor. The combination of SAR-Harris and Ratio descriptor is used in the SAR-SIFT algorithm (see Fig. 8). This algorithm thus achieves better results than the original SIFT algorithm on SAR images.

TABLE I
AVAILABLE IMAGES, THEIR CHARACTERISTICS, NUMBER FOR EACH SET, AND SIZE IN PIXELS. TSX IS THE ABBREVIATION FOR TERRASAR-X AND CSK FOR COSMO-SKYMED. D IS FOR DESCENDING PASSES, AND A IS FOR ASCENDING PASSES. ALL IMAGES WERE ACQUIRED ON SPOTLIGHT OR STRIPMAP AND RIGHT-LOOKING MODE

Set	Sensor	Angle	Resolution	Passes	Number	Size
a	CSK	48°	1m	D	4	2048×2048
b	CSK	34°	1m	D	2	2097×1914
c	CSK	43°	3m	D	6	646×550
d	TSX	34°	2m	A	2	1500×3000
e	TSX	34°	1×2m	A	1	750×1500

V. EXPERIMENTAL VALIDATION OF THE SAR-SIFT ALGORITHM IN MORE COMPLEX SITUATIONS AND APPLICATIONS

After having experimentally validated the efficiency of the SAR-SIFT algorithm in dealing with strong SAR noise, we analyze its behavior in more complex situations and, in particular, with different acquisition modes. The use of the RANSAC algorithm is then suggested to increase the number of correct matches while keeping a low false alarm rate for some specific situations where a global deformation exists. Finally, two applications of the SAR-SIFT algorithm are considered, i.e., registration and change detection.

A. Behavior of SAR-SIFT With Different Image Viewing Conditions

We focus now our study on image pairs with different resolutions and/or different incidence angles. In order to have a reference situation, we also consider image pairs having the same viewing conditions. We can expect some good results for the situations with only a difference in resolution. The situation is more complex when the incidence angle is changed, since the image geometry and the SAR signal are different. Therefore, it can be expected that the number of keypoints correctly associated will decrease in this case. This section investigates the variability of this number in different conditions.

The available sets of images are presented in Table I, with their characteristics (sensor, incidence angle, resolution, pass), number of images, and size in pixels.

For the images with the same viewing conditions, matches were computed between every image of their respective sets a, b, and c. To obtain a situation with a varying resolution, a multilook has been computed in the azimuth direction for the images of the set d to obtain a scale factor of 2 with respect to the images of the set e. For the case with difference in incidence angles, sets a and b are used to study a situation with a difference of 14°. Both are also compared with set c, but in those cases, there is also a difference in resolution. Five configurations are thus considered.

Ground truth deformation grids have been manually estimated[2] between all considered images, and matches are evaluated by using (12). The parameter t_1 is set to 5, except

[2]Starting from a registration using the sensor parameters provided by space agencies, a fine manual registration is computed.

TABLE II
AVERAGE PERCENTAGE OF KEYPOINTS THAT ARE CORRECTLY
MATCHED WITH THE NN METHOD FOR DIFFERENT
CONFIGURATIONS (SCALE FACTOR AND/OR
DIFFERENT INCIDENCE ANGLES)

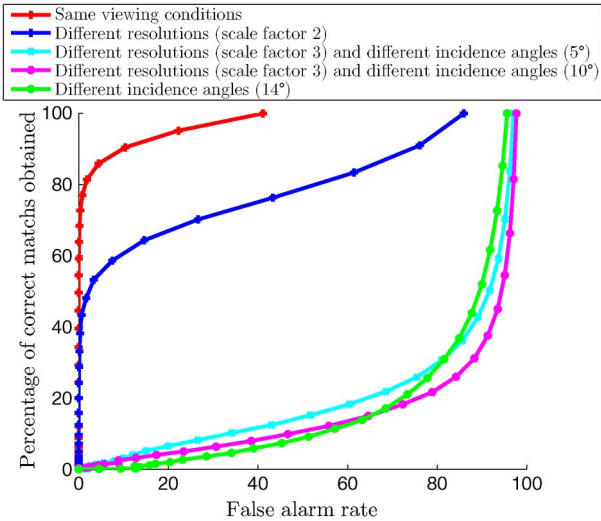| Scale factor	Difference of incidence angle	Percentage of keypoints correctly matched
-	-	61%
2	-	27%
3	5°	9.4%
3	9°	6.7%
-	14°	4.5%

Fig. 14. Global ROC curves, computed with the images in Table I, to assess the number of SAR-SIFT matches with different acquisition conditions.

for situations with incidence angle differences, where it is set to 7 because of layover effects. Indeed, due to slant-range distorsions, building sizes may vary in those cases, and a more loose threshold should be considered. Here, SAR-SIFT algorithm is performed with orientation assignment.

We first observe the ability of the algorithm to match corresponding keypoints using the NN step only (see Section IV-C). We want to verify that the algorithm can obtain a sufficient number of correct matches, without considering the DR step.

Table II presents the average percentage of keypoints correctly matched for each configuration.

Next, the DR threshold (allowing to reject false correspondences) is varied to produce ROC curves. Using the same notations as in Section IV-C, the percentage of good matches $\#CM/\#CM_{all}$ is plotted against the percentage of false matches $\#FM/(\#CM + \#FM)$ for different values of th (see Fig. 14).

As it can be expected, situations with either the same viewing conditions or difference in resolution present really good scores. The NN step allows a high percentage of keypoints to be correctly matched, i.e., 61% and 27%, respectively. Among these correct matches, roughly 80% and 45%, respectively, can be obtained with the DR step with only 1% of false alarms.

When considering an incidence angle difference, the scores are much lower, i.e., between 5% and 10%. Nevertheless, between roughly 500 and 2500 correctly matched keypoints can be obtained, depending on the image (differences are due

to the size of the images). These numbers would be perfectly usable, but the DR ratio fails at selecting correct matches, which yields high false alarm rates. It is interesting to note that the influence of the difference in resolution is very limited, unlike the incidence angle. Indeed, the matching between sets a and b, with an incidence angle difference of $14°$ and the same resolution, shows weaker scores than the matching between sets b and c and a and c, both with a scale factor of 3 and an incidence angle difference of $9°$ and $5°$, respectively.

On images with the same viewing conditions or with only a resolution difference, image geometry barely varies, and SAR-SIFT achieves a high number of correct keypoints matching (always more than a thousand in our experiments). However, with different incidence angles, the SAR signal varies, and the corresponding distorsions are not taken into account in the invariances of SAR-SIFT. The larger the difference is, the greater the image variations are and the more difficult it is for SAR-SIFT to find corresponding keypoints.

B. Filtering of False Matches to Search a Global Deformation: AC-RANSAC

1) Proposed Approach: As seen in the previous paragraph, the filtering part of the keypoints matching (DR step) fails in situations with differences of incidence angles.

However, for registration purpose and as a first approximation, satellite image pairs may be related by a global deformation. In multiview contexts and on roughly flat scenes, this deformation is often modeled by an affine transformation [42]. SAR images do present local displacements when acquired under different incidence angles. However, the global relation between two SAR images can be described by this model, provided that they are acquired by the same sensor and in the same pass.

This *a priori* information can be helpful to suppress false correspondences in registration applications. We propose to use the RANSAC algorithm [43] and an *a priori* of affine transformation in order to obtain a high number of correct matches with a low false alarm rate. This setting can be used to estimate a global affine deformation in presence of outliers (false correspondences) and, thus, to provide a set of coherent correspondences. We have chosen to use an a contrario version of RANSAC, which is called AC-RANSAC [44], that yields good results even with a very high percentage of outliers (up to 90%). It also has the advantage to require only one parameter, i.e., the number of iterations i_{max}.

2) Experimental Results: AC-RANSAC has been applied on the correspondences of the images of the fifth situation in Table II, presenting the largest incidence angle difference. In order to obtain the highest number of correct matches, the DR filtering part is not applied. We call simple matches all the resulting correspondences using the NN step only. However, to speed up the process, a random selection of matches used to estimate the deformation is done among the correspondences with a ratio of distances lower than the threshold value of th = 0.09. This value is significantly higher than classical choices [1]. However, the goal here is only to reduce the number of NN matches so that RANSAC can deal with outliers, while

Studied situation	Simple matches	AC-RANSAC
Correct matches	2251	1979
False matches	48747	104

keeping a sufficient number of matches to obtain an accurate registration. The value 0.09 is a good tradeoff for this task. We have chosen $i_{\max} = 10\,000$ as the number of iterations and affine transformation as the global deformation. Average numbers of false and true correspondences are presented in Table III, before and after applying AC-RANSAC.

The use of AC-RANSAC allows suppressing almost all false matches while keeping 88% of the correct matches. The percentage of outliers goes from about 95% to 5%. We can achieve a high number of correct matches with a reasonable percentage of false alarms.

C. Application of SAR-SIFT to Registration

As explained in Section V-B, the deformation between two images acquired by the same sensor and in the same pass can be approximated by an affine transformation. This approximation, however, is only valid for points on the ground. AC-RANSAC has proven its efficiency to filter a significant number of false matches for registration purpose and can be also used to estimate the affine transformation between two images. We propose here a registration application of the SAR-SIFT algorithm in situations where the incidence angle is varying.

1) Proposed Approach: To register such images, we need to estimate the coefficients of the following polynomial transformation:

$$x_2 = a_1 + a_2 \cdot x_1 + a_3 \cdot y_1$$
$$y_2 = b_1 + b_2 \cdot x_1 + b_3 \cdot y_1. \tag{13}$$

A least square estimation of the coefficients can be done using the keypoint matching between two images. However, as shown in Section V-A, the algorithm presents low efficiency in these conditions, and we risk to face a significant number of outliers. As presented in Section V-B, the AC-RANSAC algorithm can help to model the deformation in presence of outliers and is efficient to filter false correspondences.

We propose to apply this algorithm on the matches obtained by the SAR-SIFT algorithm between two SAR images in order to estimate the parameters of (13). The same process and parameter values presented in Section V-B are used. In order to facilitate the reading, this registration approach will be referred to as SAR-SIFT + AT + AC-RANSAC in the rest of this paper.

2) Evaluation: Table IV describes the pair of images used to assess the precision of the registration. They have the same resolution but a difference of incidence angle of 14°. Since there is no available ground truth, we have manually extracted 30 ground control points (GCPs) on the pair of images. These are then used to realize a manual registration and a measure of the registration accuracy.

Ten of the GCPs are randomly selected to evaluate the parameters of (13) and obtain a manual registration. The 20 other

Sensor	Angle	Resolution	Passes	Size
Cosmo-Skymed®	48°	1m	D	2048 × 2048
Cosmo-Skymed®	34°	1m	D	2048 × 2048

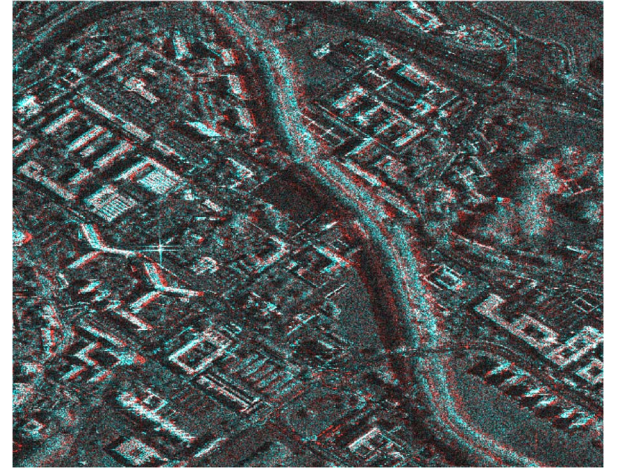| Registration method	Manual registration	SAR-SIFT + AT + AC-RANSAC
RMSE (in pixel)	2.46	2.03

Fig. 15. Superposition of the two registered images (angular difference of 14°) in Table IV with the SAR-SIFT + AT + AC-RANSAC method. The master image is in red, and the slave one is in blue.

points are used to measure the accuracy of the two registration methods, i.e., the one using SAR-SIFT + AT + AC-RANSAC and the previous manual one, by computing root-mean-square errors (RMSE). This process is repeated 10 000 times, and averaged RMSE is presented in Table V. The superposition of the two registered images by the SAR-SIFT + AT + AC-RANSAC method is presented in Fig. 15. The systematic shift around elevated objects (buildings, river borders) is due to layover effects.

The automatic method, i.e., SAR-SIFT + AT + AC-RANSAC, presents a good registration accuracy, comparable with the manual method. Considering that extracting GCP is time consuming and subject to errors, this new method is an interesting way to automatically register SAR images with different types of acquisition.

D. Change Detection Application: Preliminary Results

Another possible application of SAR-SIFT would be change detection. Such applications are often subject to misregistration errors, and the use of feature-based approaches allows avoiding a preregistration step.

As a preliminary result, we match an image pair presenting some changes (see Fig. 16). Both images were acquired by the same satellite and with the same incidence angle, but they do not have the same resolution. The image from 2007 has a
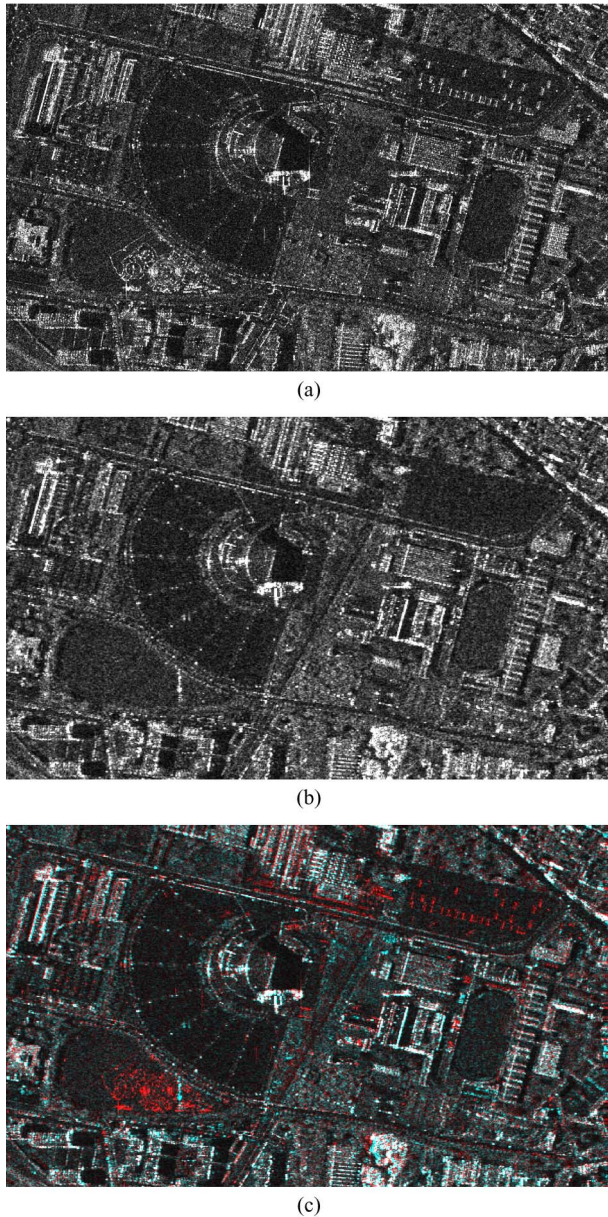
(a)

(b)

(c)

Fig. 16. Extracts of TerraSAR-X images with 34° of incidence angle, from the area of Toulouse, France. (a) Image from 2007 with 1-m resolution. (b) Image from 2008 with 2-m resolution. (c) Superposition of the two images, after registration.

resolution of 1 m, whereas the one from 2008 has a resolution of 1×2 m. For this last image, a multilook has been performed in the azimuth direction to obtain a 2-m resolution. To outline the changes, Fig. 16(c) represents a superposition of the images, after a preregistration step performed by the SAR-SIFT + AT + AC-RANSAC method. The obtained correspondences are then filtered with the AC-RANSAC algorithm. In areas with changes, we suppose that keypoints are detected but not matched. In order to test this hypothesis, the following experiment is conducted: every keypoint at a distance lower than 50 pixels of a match is thus discarded. Remaining keypoints from the 2007 image are displayed in Fig. 17.

We observe that remaining points are present on the three main sites with changes and that a few errors occur. Smaller changes are not detected, but the used algorithm is very simple,
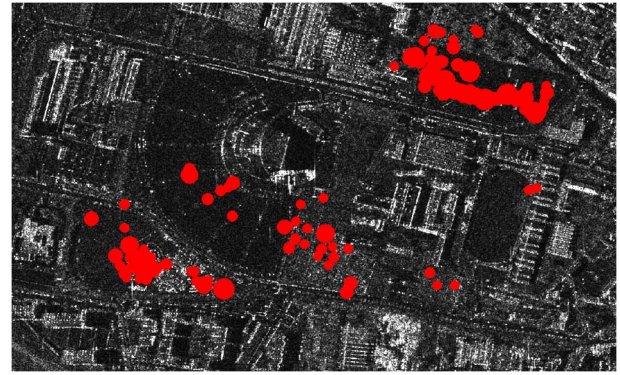

Fig. 17. Image from 2007 is displayed with all keypoints that are at a distance of at least 50 pixels from a match between the 2007 and 2008 images.

and the tolerant threshold of 50 pixels is high. Further tests and parameter settings are of course needed to fully develop the approach.

From this preliminary experiment, we think that the SAR-SIFT algorithm can be considered for change detection applications. A more sophisticated algorithm should be developed, using keypoint density, for example. Further tests are also required for the choice of the distance threshold, here set to 50.

VI. CONCLUSION

This paper has presented a new SIFT-like descriptor adapted to SAR images. It relies on a new gradient computation adapted to SAR images and robust to speckle noise. This new gradient computation method, i.e., GR, is then used to improve steps of the SIFT algorithm. A new keypoint detector based on the multiscale Harris detector offers stable keypoints. Robust gradient orientations provided by GR enable to obtain an efficient feature descriptor for SAR images.

By applying AC-RANSAC, a consistent number of correct matches can be achieved, allowing the use of this new SIFT-like algorithm for diverse applications. In this paper, an efficient registration application of SAR images is presented for difficult situations, such as incidence angle changes. Other applications such as change detection or object matching will be the subject of further work.

REFERENCES

[1] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proc. Int. Conf. Comput. Vis.*, 1999, pp. 1150–1157.
[2] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[3] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.

[4] Q. Li, G. Wang, J. Liu, and S. Chen, "Robust scale-invariant feature matching for remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 287–291, Apr. 2009.

[5] C. Huo, C. Pan, L. Huo, and Z. Zhou, "Multilevel SIFT matching for large-size VHR image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 171–175, Mar. 2012.

[6] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform robust scale-invariant feature matching for optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4516–4527, Nov. 2011.

[7] Y. Yang and C. Newsam, "Geographic image retrieval using local invariant features," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 818–832, Feb. 2013.

[8] V. Risojevic and Z. Babic, "Fusion of global and local descriptors for remote sensing image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 836–840, Jul. 2013.

[9] B. Sirmacek and C. Unsalan, "Urban-area and building detection using SIFT keypoints and graph theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1156–1167, Apr. 2009.

[10] C. Tao, Y. Tan, H. Cai, and J. Tian, "Airport detection from large IKONOS images using clustered SIFT keypoints and region information," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 1, pp. 128–132, Jan. 2011.

[11] J. Z. Liu and X. C. Yu, "Research on SAR image matching technology based on SIFT," in *Int. Archives Photogramm., Remote Sens. Spatial Inf. Sci.*, 2008, pp. 403–408.

[12] S. Suri, P. Schwind, P. Reinartz, and J. Uhl, "Combining mutual information and scale invariant feature transform for fast and robust multisensor SAR image registration," in *Proc. 75th Annu. Amer. Soc. Photogramm. Remote Sens. Conf.*, 2009, pp. 795–806.

[13] P. Schwind, S. Suri, P. Reinartz, and A. Siebert, "Applicability of the SIFT operator to geometric SAR image registration," *Int. J. Remote Sci.*, vol. 31, no. 8, pp. 1959–1980, Mar. 2010.

[14] S. Suri, P. Schwind, J. Uhl, and P. Reinartz, "Modifications in the SIFT operator for effective SAR image matching," *Int. J. Image Data Fus.*, vol. 1, no. 3, pp. 243–256, Sep. 2010.

[15] J. Lu, B. Wang, H. M. Gao, and Z. Q. Zhou, "SAR images matching based on local shape descriptors," in *IET Int. Radar Conf.*, 2009, pp. 1–4.

[16] L. Liu, Y. Wang, and Y. Wang, "SIFT based automatic tie-point extraction for multitemporal SAR images," in *Int. Workshop Edu. Technol. Training Int. Workshop Geosci. Remote Sens.*, 2008, vol. 1, pp. 499–503.

[17] W. Lv, W. Yu, J. Wang, and K. Wang, "SAR image matching based on SIFT keypoints and multi-subregions information," in *Int. Asia-Pac. Conf. Synthetic Aperture Radar*, 2011, pp. 1–4.

[18] B. Fan, C. Huo, C. Pan, and Q. Kong, "Registration of optical and SAR satellite images by exploring the spatial relationship of the improved SIFT," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 657–661, Jul. 2013.

[19] B. Wessel, M. Huber, and A. Roth, "Registration of near real-time SAR images by image-to-image matching," in *Proc. Photogramm. Image Anal.*, 2007, p. 179.

[20] Y. Xiaoping, L. Tong, L. Pingxiang, and H. Guoman, "The application of improved SIFT algorithm in high resolution SAR image matching in mountain areas," in *Proc. Int. Symp. Image Data Fus.*, 2011, pp. 1–4.

[21] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: A SIFT-like algorithm for applications on SAR images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2012, pp. 3478–3481.

[22] T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, Nov. 1998.

[23] Y. Dufournaud, C. Schmid, and R. P. Horaud, "Matching images with different resolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2000, pp. 612–618.

[24] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," in *Int. Conf. Comput. Vis.*, 2001, vol. 1, pp. 525–531.

[25] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *Int. J. Comput. Vis.*, vol. 60, no. 1, pp. 63–86, Oct. 2004.

[26] K. Mikolajczyk and C. Schmid, "An affine invariant interest point detector," in *Proc. 7th Eur. Conf. Comput. Vis.*, 2002, pp. 128–142.

[27] T. Tuytelaars and K. Mikolajczyk, "Local invariant feature detectors: A survey," *Found. Trends Comput. Graph. Vis.*, vol. 3, no. 3, pp. 177–280, Jan. 2008.

[28] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 404–417.

[29] J. Rabin, J. Delon, and Y. Gousseau, "A statistical approach to the matching of local features," *SIAM J. Imag. Sci.*, vol. 2, no. 3, pp. 931–958, Jul. 2009.

[30] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2004, vol. 2, pp. II-506–II-513.

[31] J.-M. Morel and G. Yu, "ASIFT: A new framework for fully affine invariant image comparison," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 438–469, Apr. 2009.

[32] M. Weinmann, "Visual features: From early concepts to modern computer vision," in *Advanced Topics in Computer Vision*, *Advances in Computer Vision and Pattern Recognition*, G. M. Farinella, S. Battiato, and R. Cipolla, Eds. London, U.K.: Springer-Verlag, 2013, pp. 1–34.

[33] R. Szeliski, *Computer Vision: Algorithms and Applications*, 1st ed. New York, USA: Springer-Verlag, 2010.

[34] S. Wang, H. You, and K. Fu, "BFSIFT: A novel method to find feature matches for SAR image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 649–653, Jul. 2012.

[35] Y. Wu and M. Yang, "A multi-sensor remote sensing image matching method based on SIFT operator and CRA similarity measure," in *Int. Conf. Intell. Sci. Inf. Eng.*, 2011, pp. 115–118.

[36] R. Touzi, A. Lopes, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 6, pp. 764–773, Nov. 1988.

[37] A. C. Bovik, "On detecting edges in speckle imagery," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 36, no. 10, pp. 1618–1627, Oct. 1988.

[38] C. J. Oliver, D. Blacknell, and R. G. White, "Optimum edge detection in SAR," *Proc. Inst. Elect. Eng.—Radar, Sonar Navigat.*, vol. 143, no. 1, pp. 31–40, Feb. 1996.

[39] R. Fjortoft, A. Lopes, P. Marthon, and E. Cubero-Castan, "An optimal multiedge detector for SAR image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 3, pp. 793–802, May 1998.

[40] A. Desolneux, L. Moisan, and J.-M. Morel, *From Gestalt Theory to Image Analysis: A Probabilistic Approach*, vol. 34. New York, NY, USA: Springer-Verlag, 2008.

[41] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *Int. J. Comput. Vis.*, vol. 37, no. 2, pp. 151–172, Jun. 2000.

[42] B. Zitova and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, Oct. 2003.

[43] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

[44] J. Rabin, J. Delon, Y. Gousseau, and L. Moisan, "MAC-RANSAC: A robust algorithm for the recognition of multiple objects," in *Proc. 3D Data Process., Vis. Transmiss.*, 2010.

Flora Dellinger received the Engineer degree from the École Nationale Supérieure de l'électronique et de ses Applications, Cergy, France, and the M.Sc. degree in electrical engineering from the Illinois Institute of Technology, Chicago, IL, USA. She is currently working toward the Ph.D. degree at Télécom ParisTech, Paris, France.

Her research interests are remote sensing application, such as feature detection and description, multisensor synthetic aperture radar image registration, and multisensor change detection.

Julie Delon received the Bachelor's degree in mathematics, the Ph.D. degree in applied mathematics, and the HDR degree from the École Normale Supérieure, Paris, France, in 2001, 2004, and 2011.

From 2005 to 2013, she was a CNRS Researcher with Télécom ParisTech, Paris. She is currently a Professor of applied mathematics with Paris Descartes University, Paris. She works on optimal transport algorithms and their application in image processing and computer vision, on image and movie restoration, and on invariant descriptions and stochastic approaches for computer vision.

Yann Gousseau received the Engineer degree from the École Centrale de Paris, France, in 1995 and the Ph.D. degree in applied mathematics from the University of Paris Dauphine, Paris, France, in 2000.

He is currently a Professor with Télécom Paris-Tech, Paris. His research interests include the mathematical modeling of natural images and textures, stochastic geometry, image analysis, computer vision, and image processing.

Julien Michel received the Telecommunications Engineer degree from the École Nationale Supérieure des Télécommunications de Bretagne, Brest, France, in 2006.

From 2006 to 2010, he was with Communications et Systèmes, Toulouse, France, where he worked on studies and developments in the field of remote sensing image processing. He is currently with Centre National d'études Spatiales (French Space Agency), Toulouse, where he is in charge of the development of image processing algorithms for the exploitation of Earth observation images, mainly in the field of very high resolution optical image analysis.

Florence Tupin received the Engineer degree and the Ph.D. degree in signal and image processing from the École Nationale Supérieure des Télécommunications, Paris, France, in 1994 and 1997, respectively.

From 1997 to 1998, she was with Sagem in fingerprint recognition. Since November 1998, she has been with Télécom ParisTech, Paris, where she was an Associate Professor and currently a Professor of signal processing and computer vision in the Image and Signal Processing Department. Her research interests are image processing and interpretation and 3-D reconstruction, particularly for remote sensing applications and Earth observation.