SAR Coregistration by Robust Selection of Extended Targets and Iterative Outlier Cancellation

Luca Pallotta, Senior Member IEEE, Gaetano Giunta, Senior Member IEEE, Carmine Clemente, Senior Member IEEE, and John J. Soraghan, Senior Member IEEE

Abstract—This letter extends the constrained Least Squares (CLS) optimization method developed to coregister multitemporal synthetic aperture radar (SAR) images affected by a joint rotation effect and range/azimuth shifts enforcing the absence of zooming effects. To take advantage of the structural information extracted from the scene, the method starts with a detection stage that identifies extended targets/areas in the images. The selected tie-points allow the CLS problem to be reformulated to find its (initial) solution based on a robust subset of image blocks. Then, the mean square error (MSE) of each equation evaluated from the initial solution allows to implement an iterative cancellation procedure to further skim the CLS equation set. The effectiveness of the proposed procedure is validated on real SAR data in comparison with the standard CLS.

Index Terms—Synthetic Aperture Radar (SAR) coregistration, rotation and translation, CFAR Detection, outlier cancellation.

I. INTRODUCTION

COREGISTRATION is among the most important tasks of synthetic aperture radar (SAR) processing when dealing with multiple acquisitions of the same scene. When the acquisition process is repeated over time, it is expected that the acquired images differ each other as a consequence of the different trajectories followed by the platform. Even the minor mismatch in trajectories can have a significant effect on the position of the final scene represented, thus affecting final SAR-related products. A commonly employed solution to this problem consists in coregistering the images after the image formation process [1]. This methodology consists in aligning one image (slave) with respect to the other (master) in such a way that the pixels in the master-slave couple refer to the same scatterers of the physical scene. Two main philosophies have been generally followed in the open literature to achieve the image registration objective [2]. The first family of methods, called feature-based, relies on the identification of the tie-points in the two images that are then coupled to estimate the relative shift among them [3]–[6]. Another approach, referred to as area-based, exploits the misalignment information for instance embedded in the two-dimensional (2D) spatial cross-correlation between uniformly extracted patches in the master and slave [7]–[9].

In practical situations, a pure translation would not be representative of the misalignment effect between the two images and also a respective rotation should be accounted for at the design stage (affine transformations [10]). In [11], a method for SAR coregistration in the presence of rotation and translation has been proposed. The procedure divides the images into several patches and computes their misalignments from the peak of the 2D cross-correlation, formulating the resulting problem as a constrained least squares (CLS) enforcing the absence of a zooming effect between the two images. The main limitation of this approach is the fact that all patches are used in the CLS problem formulation even in the case of very low coherence values (for instance, in the presence of vegetation as well as geometric distortions like layover, shadowing, etc.), producing some performance degradations. To circumvent these limitations, in this letter we propose a new coregistration algorithm based on both extended target detection and outlier exclusion in the set of the CLS’ equations. More precisely, a pre-processing stage aimed at identifying extended targets/areas within both images is first applied. Then, their centroids are computed to be used as candidate tie-points for the subsequent formulation of the CLS problem. By doing so, differently from [11] where all patches in the scene are involved in the CLS, in this work only the more reliable areas in the two images are selected before the CLS formulation, excluding those of low interest. Additionally, after the problem is formalized, the set of equations is iteratively reduced performing an outlier cancellation procedure in agreement with their mean square error (MSE). Tests conducted on the Gotcha Volumetric dataset demonstrate the effectiveness of this method in correctly estimating the rotation angle.\(^\dagger\)

The letter is organized as follows. The coregistration problem is defined in Section II and the proposed procedure to solve it is detailed in Section III. Section IV is dedicated to the performance evaluation. Finally, Section V concludes the letter by providing some hints for future developments.

II. PROBLEM FORMULATION

In this section we formalize the problem of SAR image registration like in [11] when the couple is affected by both a respective rotation and shift effect. Therefore, letting

\(^\dagger\)Notation: We use boldface lower case for vectors \(\mathbf{a}\) and upper case for matrices \(\mathbf{A}\), and \(\text{diag}(\mathbf{a})\) is the diagonal matrix whose diagonal entries are the values in vector \(\mathbf{a}\). Then, \(\mathbb{R}\) and \(\mathbb{C}\) are the set of real and complex numbers, respectively, and \(\mathbb{C}^{K \times N}\) is the Euclidean space of \((K \times N)\)-dimensional complex matrices (or vectors if \(N = 1\)). The symbols \(|\cdot|^2\) and \((\cdot)^\dagger\) denote the transpose and conjugate transpose operators, respectively, while \(|\cdot|\) and \(||\cdot||\) are the modulus and Euclidean norm, respectively. Finally, \(j = \sqrt{-1}\) is the imaginary unit.
$z = x + jy$ be the complex variable describing the Cartesian coordinates $x$ and $y$, each SAR image is indicated as $I(z)$, with $z \in \mathbb{C}$. Now, considering two images of the same scene, $I_m \in \mathbb{C}^{N \times N}$ (master) and $I_s \in \mathbb{C}^{N \times N}$ (slave), the effect of pixels translation and rotation between them can be formally defined in the complex domain (in spite of the corresponding affine transformation of the 2D spatial domain) as

$$I_s(z) = I_m((z - \delta)/\alpha) + E(z),$$

with $E(z)$ the error image comprising noise and the effects of different scattering properties of the considered pair. Moreover, the parameter $\delta = \delta_x + j\delta_y$ is the complex displacement accounting for horizontal $\delta_x$ and vertical $\delta_y$ displacements, whereas $\alpha = \gamma |\exp(j\theta)|$ is a complex scalar factor accounting for rotation, described by the rotation angle $\theta$ and the zooming factor $\gamma$ (assumed to be 1 because we enforce the absence of zooming effect). As a consequence, the problem to be solved consists in estimating the unknown complex parameters $\delta$ and $\alpha$. To do this, we search for some ground control points (or tie-points) within the two images in order to have a coarse estimation of the respective 2D displacement. Precisely, indicating with $\zeta_l = x_{s,l} + jy_{s,l}$, $l = 1, \ldots, L$, the complex valued coordinates position of the selected ground control points in the slave image and with $z_l = (x_{m,l} + jy_{m,l})$, $l = 1, \ldots, L$, those of the selected ground control points in the master, the overall displacement field between the quoted images is computed solving a constrained over-determined linear system of $L$ equations in $3$ unknowns [11]

$$\left\{ \begin{array}{l}
\arg \min_{p} |Ap - \text{diag}(w)\zeta_l|^2 \\
\text{s.t. } p^TDp - 1 = 0
\end{array} \right.,$$

with

$$D = \begin{bmatrix}
0 & 0 \\
0 & 1
\end{bmatrix}, \quad A = \text{diag}(w)Z, \quad \text{and } Z = \begin{bmatrix}
z_1 & 1 \\
\vdots & \vdots \\
z_L & 1
\end{bmatrix},$$

where $p = [\alpha, \delta]^T$, $\zeta_l = [\zeta_1, \ldots, \zeta_L]^T$ (with the coordinates $\zeta_1, \ldots, \zeta_L$ expressed with respect to the reference system centered at the image center), and $w = [w_1, \ldots, w_L]$ is the vector containing the weights (herein we assume $w = [1, \ldots, 1]^T$) applied to each considered couple of ground control points.

In [11], the optimal solution to the CLS problem described in (2) has been derived in closed-form. Moreover, the CLS problem is constructed dividing the entire images into several corresponding patches and taking as ground control points the positions of the maxima of the cross-correlation modulus computed between them. However, this strategy could show some performance losses due to low coherence between some patches, the presence of noisy patches, or poor estimates in some correlation peaks. To overcome these issues, the proposed method (detailed in Section III) comprises two important novelties:

- the ground control points are selected by following a target detection-based approach;
- among all the selected equations composing the CLS, those indicated to be outliers are excluded from the registration problem.

### III. PROPOSED COREGISTRATION ALGORITHM

The block scheme synthesizing the proposed coregistration algorithm is depicted in Fig. 1. The first step of the procedure is the application of a constant false alarm rate (CFAR) detection scheme to both images to reveal all possible targets/strong reflective areas dislocated within the observed scene [12]. The CFAR detector utilizes a sliding window over the entire images to adaptively estimate the surrounding power level of the pixel under test that is then compared to a specific threshold set to ensure a predetermined false alarm rate. In the following tests, we focus on the classic cell-averaging CFAR (CA-CFAR), even more sophisticated CFAR detectors, such the greatest-of CA-CFAR (GOCA-CFAR), smallest-of CA-CFAR (SOCA-CFAR), trimmed mean or censored CFAR (TM-CFAR, CS-CFAR), ordered statistics CFAR (OS-CFAR) can be also applied [12]–[16].

The CFAR detection algorithm generates a binary detection map. After that, two additional steps need to be applied:

1) target clustering,
2) false alarms cancellation.

As to point 1), it is worth noting that, during the detection process, some targeted pixels may not exceed the detection threshold resulting in a non-complete target shape extraction. Hence, the clustering procedure allows the levels of the missed pixels to be determined. This operation is performed using a so-called order filter that replaces the element of the pixel under test by an element in the sorted set of neighbors. By doing so, the detection capabilities of the CFAR algorithm results in a clear enhancement. Regarding point 2) the reduction in false alarms is used to remove isolated detections, that are mainly due to noise spikes or strong reflections due to point-like targets. This operation is performed by means of a median filtering [17], i.e., a nonlinear operation typically used to reduce salt and pepper noise in optic images or speckle in SAR. This method allows to reduce spikes, isolated detections, and false alarms.

The binary maps at the output of this pre-processing stage are the input to the next step where, for each identified target of interest, its centroid is estimated as the center of mass of the detected region, namely computing the arithmetic mean of
the coordinates of the activated pixels. Once the centroids are estimated, their association is determined using a minimum distance criterion. More precisely, for each centroid in the master, the corresponding in the slave is found as the one sharing the lower Euclidean distance from that in the master. Using this approach allows the ambiguities in the association process to be resolved. Since the number of detected centroids could be different in the two images, those not associated are directly removed, possibly reducing the number of coupled centroids to \( L_1 \leq L \). At this stage, the tie-points considered at defining the CLS problem in (2) can be obtained as one of the following three alternative choices:

1. centroids’ positions;
2. peak of the complex 2D cross-correlation modulus;
3. peak of the real 2D cross-correlation;

For choice 1) denoting \( z_l, l = 1, \ldots, L_1 \), as the coordinate values of each of the \( L_1 \) detected and selected centroids in the master, whereas \( \zeta_l, l = 1, \ldots, L_1 \), are the coordinates of the counterparts in the slave. The displacement field, computed from (2), is formed with the displacement field directly composed by the targets’ centroids. The remaining two choices are motivated by the fact that some associated detections could not belong to the same target (e.g., a target is present only in one image). In such a case, it would be better to refer to the area surrounding the detected centroids and evaluates the cross-correlation of the patches centered at the centroid position. Choice 2) refers to case in which the complex SAR sub-images (patches) of size \( W_1 \times W_2 \) are used as is to evaluate the complex cross-correlation \( C_l \in \mathbb{C}^{(2W_1 - 1) \times (2W_2 - 1)} \) for each of the \( L \) couples. Conversely, for choice 3) the modulus of the complex patches are used to derive the real cross-correlation.

The final step of the proposed procedure performs the exclusion of the outliers within the set of equations. More precisely, exploiting the absolute error of the estimated parameters vector \( \hat{p} \), those equations associated with its higher values are rejected and a new reduced dimension system of more accurate equations is solved. To do this the absolute error for each equation in (2) is computed as

\[
\epsilon = [\epsilon_1, \ldots, \epsilon_{L_1}]^T = |A\hat{p} - \text{diag}(w)\zeta|,
\]

and each absolute error value, \( \epsilon_l, l = 1, \ldots, L_1 \), is compared with a specific threshold, \( \eta \). Then the equations contained in (2) for which \( \epsilon_l > \eta \) are removed leading to a new reduced-size CLS problem. The threshold \( \eta \) is selected resorting to the median absolute deviation (MAD) criterion [18], [19], that is defined as

\[
\text{MAD} = \kappa \mu \text{median}(|\epsilon - \text{median}(\epsilon)|),
\]

where

\[
\mu = -1/(\sqrt{2}\text{erfcinv}(3/2)) = 1.4826\text{ is a constant motivated in [18], [19],}
\]

\[
\text{erfcinv}(\cdot)\text{ is the inverse of the complementary error function,}
\]

\[
\text{median}(\cdot)\text{ is the median of the data contained in its vector argument, and}
\]

\[
\kappa \text{ is a factor typically chosen between 2 and 3 [18], [19].}
\]

It is worth noting that a lower value of \( \kappa \) (e.g., 2) provides a stricter exclusion with the cancellation of a higher number of equations. Conversely, a higher value of \( \kappa \) allows the rejection of a lower number of incorrect equations. Therefore, if \( \kappa = 2 \) is directly used, the cancellation is based on a poor estimate of the equation error because of its dependence on the current approximation of unknowns’ vector. Conversely the use of a larger value of \( \kappa \) helps the algorithm to improve the numerical convergence to the asymptotic unknowns’ vector and hence to an enhanced estimate of equation errors. Consequently, the proposed method uses an iterative procedure aimed at gradually removing the most incorrect equations once they are detected as likely outliers. However, since the operations carried out at each iteration are obtained in closed-form (e.g., the solution to the CLS problem), the computational time with or without iterations is essentially the same. More precisely, the outlier cancellation process is iteratively repeated starting from \( \kappa = 3 \) and reducing it until reaching 2 with a step size chosen according to the total number of available equations. In our case, the step size has been set equal to 0.25 in the analyses.

IV. PERFORMANCE ANALYSES

In this section, the coregistration capabilities of the proposed algorithm are tested on the Gotcha Volumetric SAR Data Set V1.0 [20], acquired with a carrier frequency of 9.6 GHz and 640 MHz bandwidth, full azimuth coverage, eight different elevation angles and full polarization. The available data refer to a zone that contains several civilian vehicles as well as calibration targets. These data are pre-processed by dividing the aperture in sub-apertures of \( 4^\circ \) in azimuth, leading to approximately equal range-azimuth resolution cells of 23 cm. The resulting image has 501 × 501 pixels and is available for all polarizations (viz., HH, VV, HV, VH). In the work, we conduct our experiments with reference to the HH polarization and to the first pass of acquisition selecting each time one \( 4^\circ \) aperture image as master and the successive one as a slave after a counterclockwise rotation by an angle \( \theta = 4^\circ \) to return each image to its original acquisition. By doing so, the entire used dataset comprises 90 master-slave couples of SAR images that are used for the tests. Moreover, it is noted that the induced rotation generates a non-integer translation of the pixels, therefore the resulting image is re-sampled through a nearest neighbor interpolation. To better show the effect of a rotation of \( \theta = 4^\circ \), Fig. 2 shows a typical master-slave couple from the Gotcha dataset. Precisely, the master is representative of the first acquisition, whereas the slave is the second acquired image which is also rotated by \( 4^\circ \).

Before analyzing the results of the conducted tests, some intermediate results obtained during the proposed process are given. Precisely, Fig. 3(a) shows the output of the CFAR detector with respect to the master image given in Fig. 2(a). The detection map clearly emphasize two aspects: on one side some extended targets of interest are revealed, on the other side several false alarms are present within the map. Then, the output of the clustering and false alarms removing process is shown in Fig. 3(b) to emphasize the detected targets. From this binary map, the cleaning process is evident as well as the shape of the revealed extended targets is much more visible.
Following the procedure of the proposed method, the last step in the image pre-processing consists in the estimation and coupling of the centroids of detected targets. This is graphically shown in Fig. 4, where the master centroids are indicated with • marker, whereas those of the slave are depicted with + marker. These represent the displacement field of the quoted couple, where the rotation effect affecting the slave is quite visible. It would be expected that some estimated couples may result to not be coherent with the other motion vectors. As an example, after a visual inspection of Fig. 4, it clearly appears that some couples (identified by the dotted circles) are definitely outliers if compared to the great majority of the motion field. However, since the CLS solves a problem with only three unknowns, it does not require the availability of a high number of correctly connected tie-points to obtain its solution. Moreover, in practical scenarios, it is expected that it not so rare to found a sufficient number of valid tie-points (e.g., due to tree trunks, buildings, rocks, cars, and so on). Therefore, even in presence of a poor detection performance, the proposed method should be able of ensuring satisfactorily coregistration performance.

The conducted analysis studied the capabilities of the proposed method in realigning the quoted rotated couples showing the estimated rotation angle as figure of merit. In particular, what we expect is that the rotation angle is estimated to be equal to 4° for each of the 90 couples. Additionally, to understand which of the three proposed methodologies, in terms of displacement fields generation works better than the others, we compare the three variants of our method with the direct use of centroid positions in the CLS (indicated as C-OBCLS), the use of the peaks of the modulus of the complex cross-correlation (referred to as CC-OBCLS), and that based on the peaks of the real cross-correlation between the images modulus (called RC-OBCLS), respectively. As to the other involved parameters, we consider a detection threshold in the CFAR detection so as to ensure a false alarm rate equal to $10^{-2}$ (see [15], [16] for details). Moreover, the order filter is set with the order equal to 17 over a domain of 25 pixels, and the median filter is applied with a neighborhood of $7 \times 7$ pixels. Fig. 5 shows the rotation angle for each couple and the different methods; in this case the evidence is that all of three variants of the method can ensure satisfactory performance for almost all the analyzed couples. However, the CC-OBCLS seems to be the worst among them especially for the noisier images in which the phase does not contribute in correctly estimating the displacement field. On the other hand, it is also clear that the RC-OBCLS represents the best choice for this dataset.

In addition, to demonstrate the improvements provided by the outlier cancellation procedure applied on the systems of equations derived from the displacement field, in Fig. 6, the estimated rotation angle is displayed for all the involved 90 images obtained using the proposed algorithm without the application of the final step of outlier cancellation. From the graphs it is evident that even if most of the rotation angles are correctly estimated, the dispersion around the true value (i.e.,
(°) is wider than that in Fig. 5.

Analogously, in Fig. 7 the rotation angles estimated with the standard CLS method of [11] are also plotted. From the results the evidence is that the CLS completely fails in this specific situation with estimated angles below 2° also when the outlier cancellation is applied to it.

Figure 6. Estimated rotation angle for the considered 90 couples of SAR images of the proposed method without outlier cancellation.

Figure 7. Estimated rotation angle for the considered 90 couples of SAR images using the CLS method of [11].

V. CONCLUSIONS

In this letter, a coregistration procedure accounting for rotation and shifts between the involved SAR images has been proposed. It improves the CLS optimization problem which also enforces a constraint on the absence of a zooming effect. The novelty of this method is twofold. First, a more robust subset of image blocks is selected before solving the CLS. Second, the most erroneous equations are iteratively removed from those composing the CLS. The performance of the proposed algorithm has been assessed in terms of rotation angle estimate over measured SAR data. The effects of the two algorithmic novelties have been both separately and jointly investigated. The results have demonstrated the effectiveness of the proposed approach in estimating the correct misalignment. Possible future works might consider the application of this method in the presence of moving targets in the images.

REFERENCES