A Joint Coregistration of Rotated Multitemporal SAR Images based on the Cross-Cross-Correlation

Luca Pallotta[1], Carmine Clemente[2] , Teresa Borreca[1,3], Gaetano Giunta[1], John J. Soraghan[2]*

[1]*Department of Industrial, Electronic and Mechanical Engineering, University of Roma Tre, Rome, Italy*
[2]*Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK*
[3]*Graduated MS student*
**E-mail: carmine.clemente@strath.ac.uk*

1 Introduction

Coregistration of multitemporal synthetic aperture radar (SAR) images is one among the most important steps to be done during the generation of interferograms or other SAR-related products. As a matter of fact, an accurate coregistration procedure allows to remove differences in SAR images due to the acquisition process rather than variations in the observed scene. Therefore, coregistration has been been widely studied during the years [1, 2] essentially following two competing approaches referred to as a) feature-based [3–6], and b) area-based [7–11].

When dealing with a multitude of SAR images acquired at different times, the coregistration is performed by setting one image (typically the first acquisition) as master and the remainder as slave. Then, each slave is separately coregistered to the master without accounting for a respective displacement with the other slaves. In this paper, the idea is to perform a joint estimation of the required parameters considering a unique coregistration of all slaves to the master jointly accounting for the respective displacements among slaves during the estimation process. By doing so, the method is capable of providing a more robust estimate of the involved quantities. Specifically, the method extends the constrained least squares (CLS) algorithm designed in [10, 11] for rotated and translated images using the information extracted from the cross-cross-correlations (i.e., the cross-correlation with a couple of patches cross-correlations) of the same patches centered in the identifying extended targets/areas extracted from all images. Tests conducted using the multitemporal Gotcha Volumetric SAR Data Set V1.0 show the benefits that occur by using this jointly estimation method in comparison with a standard one-by-one registration procedure.

1.1 Notation

We use boldface lower case for vectors $\boldsymbol{a}$ and upper case for matrices $\boldsymbol{A}$. The symbols $\mathbb{R}$ and $\mathbb{C}$ denote 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 $(\cdot)^T$ and $(\cdot)^\dagger$ denote the transpose and conjugate transpose operators, respectively, while $|\cdot|$ and $\|\cdot\|$ are the modulus and Euclidean matrix norm, respectively. Finally, $j = \sqrt{-1}$ is the imaginary unit.

2 Multitemporal SAR image coregistration algorithm

Let us consider the availability of K images representative of the same observed scene, $\boldsymbol{I}_k(z) \in \mathbb{C}^{M\times N}$, $k = 0, \ldots, K-1$, with $z = x + jy$ the complex variable describing the Cartesian coordinates x and y. Indicating, without loss of generality, with $\boldsymbol{I}_m(z) \in \mathbb{C}^{M\times N}$ the image for $k = 0$ as master and with $\boldsymbol{I}_{s_k}(z) \in \mathbb{C}^{M\times N}$, $k = 1, \ldots, K-1$, the $K-1$ slaves, the effect of pixels translation and rotation of a slave w.r.t. the master can be defined as

$$\boldsymbol{I}_{s_k}(z) = \boldsymbol{I}_m((z-\delta_k)/\alpha_k) + \boldsymbol{E}_k(z), \quad k = 1, \ldots, K-1, \tag{1}$$

with $\boldsymbol{E}_k(z)$ the k-th error image accounting for noise and different scattering properties, $\delta_k = \delta_{x_k} + j\delta_{y_k}$ the complex displacement, and $\alpha_k = \gamma_k \exp[j\theta_k]$ a scaling factor accounting for rotation (θ_k) and zooming (γ_k). The aim of the proposed algorithm is to estimate the unknown parameters $\delta_k \in \mathbb{C}$ and $\theta_k \in \mathbb{R}$ (having set $\gamma_k = 1$ to enforce the absence of a zoom) for all the $K-1$ slave images accounting also for their respective misalignment in the problem formalization.

To do this, the proposed algorithm solves the constrained Least Squares (CLS) problem developed in [10], properly selecting the areas of interest in the images as in [11]. In particular, once the patches from the master and slaves are extracted, a cross-cross-correlation based method, devised in [12] for delay estimation for 1-dimentional signals, is applied to construct the overall displacement field of the stack of images with respect to their master.

The proposed algorithm is described by the functional scheme shown in Figure 1 whose main steps are detailed in the following.

The starting input of the algorithm is the image selected to be the master that is used to apply the procedure developed in [11] to properly select the patches needed for the displacement field evaluation. In particular, the method of [11] applies a constant false alarm rate (CFAR) [13] detection scheme to the entire image to detect strong reflective areas. The resulting binary detection map is then improved thanks to a clustering and false alarms cancellation procedure. In fact, the clustering based on the use of an order filter allows to improve the

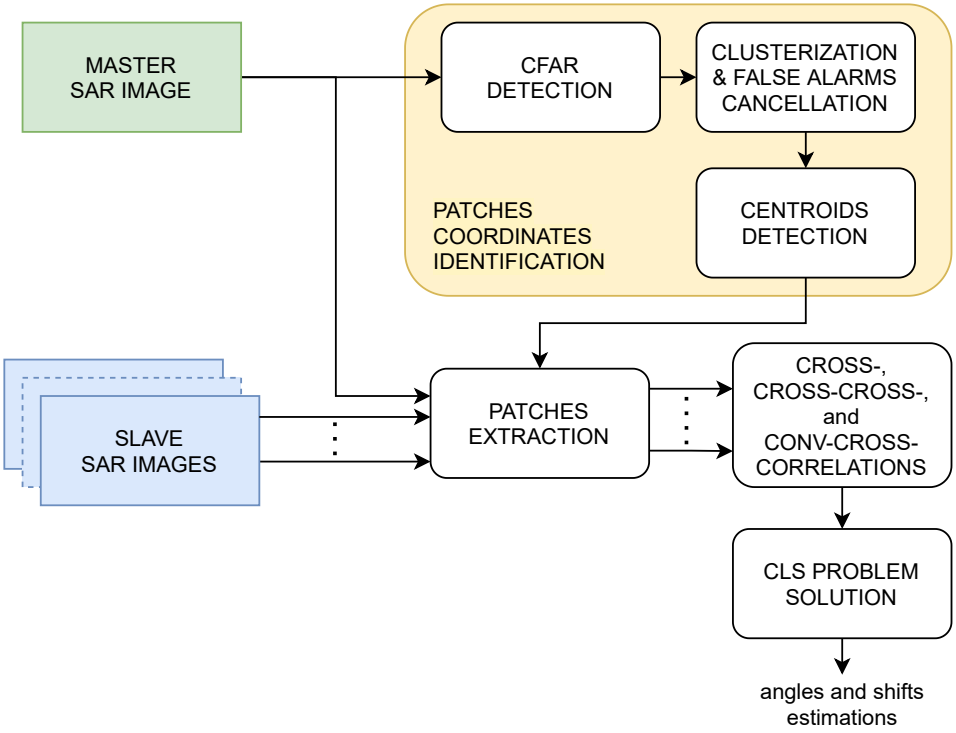


Fig. 1. Block scheme of the multitemporal SAR images coregistration algorithm based on the cross-cross-correlation.

shape of the detected object/area exploiting the behavior of the neighboring pixels of that under test. Meanwhile, the false alarm cancellation, based on a median filter, is performed to delete all noisy single detections belonging to the map. Once the extended objects are definitely obtained, their centroids are evaluated as their center of mass. Then, for each centroid in the master, a patch in its center is extracted from both the master and all slave images and all possible cross-correlations between corresponding patches in the K images are computed. Additionally, the cross-correlations between each couple of patch cross-correlations are evaluated before constructing the overall displacement field needed to solve the CLS problem detailed in Section 2.1.

2.1 CLS formulation

The CLS problem designed in [10] allows the registration of a slave image $\boldsymbol{I}_s(z) \in \mathbb{C}^{M\times N}$, affected by rotation and shift, to a master $\boldsymbol{I}_m(z) \in \mathbb{C}^{M\times N}$ representative of the same observed scene. To do this, the method needs to identify some tie-points (described in Section 2.2) in both the master and slave, here indicated as $z_l = (x_{m,l} + jy_{m,l})$, and $\zeta_l = x_{s,l} + jy_{s,l}$, $l = 1, \ldots, L$, respectively. Then, the displacement field between them is the solution of a constrained over-determined linear system of L equations in 3 unknowns [10]

$$\begin{cases} \arg\min_{\boldsymbol{p}} \|\boldsymbol{Z}\boldsymbol{p} - \boldsymbol{\zeta}\|^2 \\ \text{s.t. } \boldsymbol{p}^\dagger \boldsymbol{D}\boldsymbol{p} - 1 = 0 \end{cases}, \tag{2}$$

with

$$\boldsymbol{p} = \begin{bmatrix} \alpha \\ \delta \end{bmatrix}, \ \boldsymbol{Z} = \begin{bmatrix} z_1 & 1 \\ \vdots & \vdots \\ z_L & 1 \end{bmatrix}, \ \boldsymbol{D} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \text{ and } \boldsymbol{\zeta} = \begin{bmatrix} \zeta_1 \\ \vdots \\ \zeta_L \end{bmatrix},$$

with the coordinates $\zeta_1, \ldots, \zeta_L$ expressed with respect to the reference system centered at the image center.

2.2 Cross-cross-correlation

A standard procedure to estimate the displacement field in (2), viz. ζ_l, $l = 1, \ldots, L$, consists in evaluating the position of the peak in the magnitude of cross-correlation between corresponding patches in the master and slave as

$$[\hat{y}, \hat{x}]_k = \arg\max_{y,x} \{|\boldsymbol{G}_k(y,x)|\}, \ k = 1, \ldots, K. \tag{3}$$

where

$$\begin{aligned} \boldsymbol{G}_k(y,x) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \boldsymbol{P}_m(m,n) \boldsymbol{P}_s^*(m-y, n-x), \\ -(N-1) \le y \le (M-1), \\ -(N-1) \le x \le (M-1) \end{aligned} \tag{4}$$

is the spatial cross-correlation of the quoted couple of patches $\boldsymbol{P}_m$ and $\boldsymbol{P}_s$ in the master and slave, respectively. Note that, in

the previous equations we have omitted the subscript s to indicate that x and y are the displacement in the x- and y-direction of the patch associated to a slave image.

Beyond the classic cross-correlation $\boldsymbol{G}$ evaluated with respect to the master, it is possible to consider all possible couples of images (discarding the auto-correlation). In this case, the total number of possible combinations of K images is $Q = (K^2 - K)/2$. Additionally, it would be useful to derive the cross-correlation and convolution (say conv-cross-correlation) between each couple of images cross-correlations, in order to obtain $T = K^4/4 - K^3/2 - K^2/4 + K/2$ combinations.

The cross-cross-correlation and the conv-cross-correlation can be, respectively, defined as

$$\begin{aligned}\boldsymbol{C}_{ijlp}(\rho_y, \rho_x) = &\sum_{y=0}^{M+N-1}\sum_{x=0}^{M+N-1} \boldsymbol{G}_{ij}(y,x)\boldsymbol{G}^*_{lp}(y-\rho_y, x-\rho_x),\\ &-(M+N-1) \le \rho_y \le (M+N-1),\\ &-(M+N-1) \le \rho_x \le (M+N-1)\end{aligned} \tag{5}$$

and

$$\begin{aligned}\boldsymbol{F}_{ijlp}(\rho_y, \rho_x) = &\sum_{y=0}^{M+N-1}\sum_{x=0}^{M+N-1} \boldsymbol{G}_{ij}(y,x)\boldsymbol{G}_{lp}(\rho_y-y, \rho_x-x),\\ &-(M+N-1) \le \rho_y \le (M+N-1),\\ &-(M+N-1) \le \rho_x \le (M+N-1)\end{aligned} \tag{6}$$

The apex of the magnitude of the cross-cross-correlation, $|\boldsymbol{C}_{ijlm}(\rho_y, \rho_x)|$, should be at the index

$$[y_i - y_j - y_l + y_p, x_i - x_j - x_l + x_p],$$

while that of $|\boldsymbol{F}_{ijlm}(\rho_y, \rho_x)|$ should be at the index

$$[y_i - y_j + y_l - y_p, x_i - x_j + x_l - x_p].$$

Hence, we can estimate the $K-1$ displacements in the MMSE sense solving the overdetermined system made by the T equations, consisting of the linear combination of the $K-1$ unknowns equal to the index of the maximum of the standard and flipped cross-cross-correlations considered in (5) and (6), that is

$$\begin{aligned}&[y_i - y_j - y_l + y_p, x_i - x_j - x_l + x_p] = [\bar{\rho}_y, \bar{\rho}_x]_{ijlp}\\ &i,j,l,p = 0,\ldots,K-1\,(j > i \text{ and } p > l),\end{aligned} \tag{7}$$

and

$$\begin{aligned}&[y_i - y_j + y_l - y_p, x_i - x_j + x_l - x_p] = [\breve{\rho}_y, \breve{\rho}_x]_{ijlp}\\ &i,j,l,p = 0,\ldots,K-1\,(j > i \text{ and } p > l),\end{aligned} \tag{8}$$

where

$$[\bar{\rho}_y, \bar{\rho}_x]_{ijlp} = \arg\max_{\rho_y,\rho_x}\{|\boldsymbol{C}_{ijlp}(\rho_y,\rho_x)|\}, \tag{9}$$

and

$$[\breve{\rho}_y, \breve{\rho}_x]_{ijlp} = \arg\max_{\rho_y,\rho_x}\{|\boldsymbol{F}_{ijlp}(\rho_y,\rho_x)|\}. \tag{10}$$

Resorting to a compact matrix form, (7)-(8) can be rewritten as

$$\boldsymbol{M}[\boldsymbol{y}, \boldsymbol{x}] = \left[\boldsymbol{\rho}_y, \boldsymbol{\rho}_x\right], \tag{11}$$

with

$$[\boldsymbol{y}, \boldsymbol{x}] = \begin{bmatrix} y_1 & x_1 \\ \vdots & \vdots \\ y_{K-1} & x_{K-1}\end{bmatrix},$$

$$\left[\boldsymbol{\rho}_y, \boldsymbol{\rho}_x\right] = \begin{bmatrix} \bar{\rho}_{y_{0102}} & \bar{\rho}_{x_{0102}} \\ \vdots & \\ \bar{\rho}_{y_{(K-3)(K-1)(K-2)(K-1)}} & \bar{\rho}_{x_{(K-3)(K-1)(K-2)(K-1)}} \\ \breve{\rho}_{y_{0102}} & \breve{\rho}_{x_{0102}} \\ \vdots & \\ \breve{\rho}_{y_{(K-3)(K-1)(K-2)(K-1)}} & \breve{\rho}_{x_{(K-3)(K-1)(K-2)(K-1)}}\end{bmatrix}$$

The model matrix $\boldsymbol{M}$ of size $T \times (K-1)$ depends only on the number of multitemporal images K and comprises several null elements and some non-zero elements equal to ± 1 and ± 2 (the values ± 1 are related to measurements where an image is involved in a single operation, e.g., one cross-correlation. Whereas, the values ± 2 are related to measurements where an image is involved twice, e.g., in both the cross-correlations used in the cross-cross). For this reason it can be computed and a-priori stored reducing the computational complexity in real-time algorithms. In fact, the solution to (11) is obtained through the pseudo-inverse of $\boldsymbol{M}$, that is

$$[\hat{\boldsymbol{y}}, \hat{\boldsymbol{x}}] = \left(\boldsymbol{M}^T\boldsymbol{M}\right)^{-1}\boldsymbol{M}^T\left[\boldsymbol{\rho}_y, \boldsymbol{\rho}_x\right]. \tag{12}$$

Finally, these values are used to construct the displacement field for rotation angle and displacement estimation in (2).

3 Performance assessment

This section is aimed at assessing the performance of the proposed method for joint coregistration of multitemporal SAR images. Tests are conducted on the challenging full-polarimetric *Gotcha Volumetric SAR Data Set V1.0* [14], characterized by having a full azimuth coverage and eight different elevation angles with images acquired at different time instants. The sensor used for the acquisitions is located on a plane and operates at a carrier frequency of 9.6 GHz with a wide bandwidth of 640 MHz. The observed scene is a car parking containing several civilian vehicles (cars, forklift, tractor) and also calibration targets. For the conducted study, the

aperture was divided in azimuth sub-apertures of 4°, providing approximately equal range-azimuth resolution cells of 23 cm. By doing so, the resulting dataset comprises 90 images (looks) of 501×501 pixels for each of the 8 circular passes (different elevations) in the four polarisations (viz., HH, VV, HV, VH). To better understand the observed scene, Figure 2 depicts the span (expressed in dB) of the full-polarimetric Gotcha SAR image at $0 - 3$ degrees in azimuth.

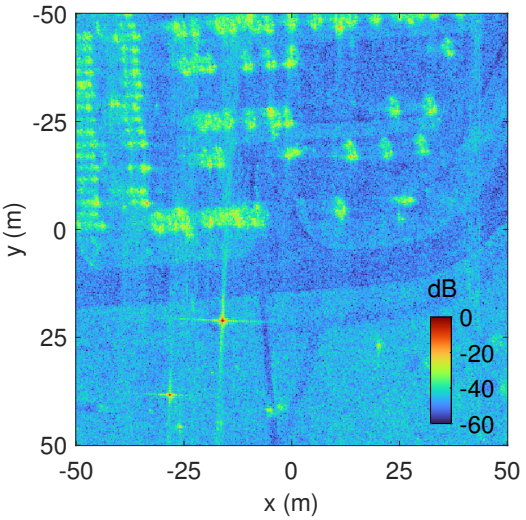


Fig. 2 Span (dB) of the full-polarimetric Gotcha SAR image at 0° azimuth.

In the next tests, without loss of generality, we focus on the HH polarization, considering all eight passes of acquisition once the azimuth angle has been fixed; then, the master is chosen to be the image at pass 0, whereas the slaves are those from pass 1 to 7. Since, the Gotcha images are provided already registered each other, in the devised tests, each slave is clockwise or counterclockwise rotated by an angle θ_k, $k = 1, \ldots, K-1$, followed by a nearest neighbor interpolation to compensate the non-integer translation of the pixels.

The analyses are conducted considering as figure of merit the root mean square error (RMSE) of the estimated angles

$$\text{RMSE} = \sqrt{\mathbb{E}\left[\left\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\right\|^2\right]}, \tag{13}$$

where $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_7]^T$ is the vector containing the seven angles to estimate, whereas $\hat{\boldsymbol{\theta}}$ is the vector containing their estimates. Now, because of the lack of a closed form expression for the RMSE, it is numerically evaluated resorting to the Monte Carlo simulation procedure. More precisely, at each Monte Carlo trial, $i = 1, \ldots, M_c$ (with M_c the number of runs), each slave image is rotated by an angle θ_k, $k = 1, \ldots, 7$, randomly chosen in the interval $[-2^\circ, 2^\circ]$.

The first test aims at empirically evaluating the optimum choice for the patch size. In fact, the size of the patches extracted from the imagery will directly impact on the final coregistration performance. This size can be a-priori set based on considerations about the overall image extent as well as the size of targets that are expected to be in it contained. Hence, Fig. 3 shows the RMSE (expressed in $^\circ$) versus the patch size, having considered, without loss of generality, square shaped patches. The curves are related to the sequence of 8 images for three different azimuth angles, viz. 0°, 176°, and 356°. Moreover, the RMSE is evaluated over a total of $M_c = 100$ Monte Carlo runs chosen the true rotation angles as above described.

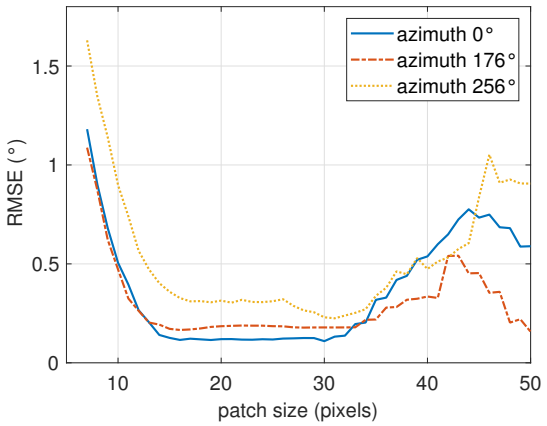
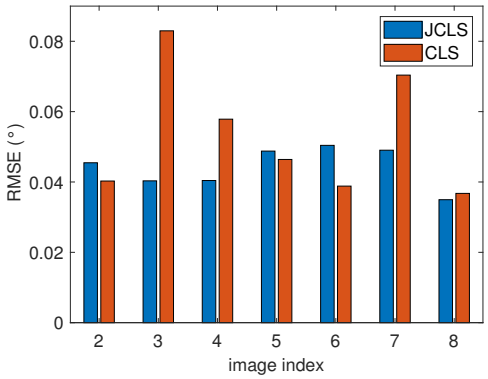


Fig. 3 RMSE ($^\circ$) versus patch size for three sequences of 8 SAR images acquired at different azimuth angles, viz. 0°, 176°, and 356°. A total of 100 Monte Carlo runs is performed randomly selecting the rotation angles in the interval $[-2^\circ, 2^\circ]$.

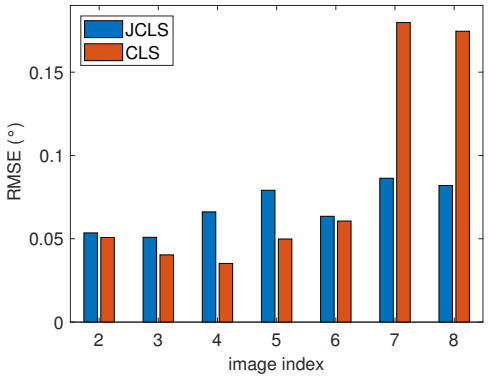
As expected the three curves show a coherent behavior each other. In fact, the RMSE is higher when the patch size is chosen to be very small (in this case, possible extended targets are spread over more patches) and also when it is chosen to be too much large (in this case, more than one target could be contained in the same patch). In particular, from the graph, it can be assured that the optimal patch size for these images is between 18×18 and 32×32 pixels.

For the above mentioned reason, in the successive tests, we set the patch size equal to 25×25 and 30×30 pixels. Therefore, Figs. 4 and 5 compare the proposed method, indicated as joint CLS (JCLS), with the CLS of [11] evaluating their achieved RMSE values for each slave image. The tests are again conducted for the same settings as in the previous analyses with the results achieved for the three different azimuth angles (viz. 0°, 176°, and 356°) shown in the respective subplots.

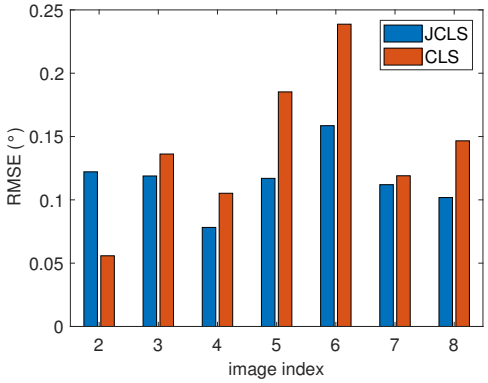
The graphical bars emphasize the superiority of the JCLS in jointly estimating the involved rotation angles. In fact, even if for some specific images the RMSE of the JCLS is slightly higher than that of the CLS, it gains much more for the others. In fact, the RMSE values shown by the JCLS are mostly homogeneous, whereas those of the CLS are strongly unbalanced. Therefore, we can conclude that the JCLS tends to mitigate the rotation angle estimation performance to provide a more balanced situation between the involved image stack.



(a) azimuth 0°

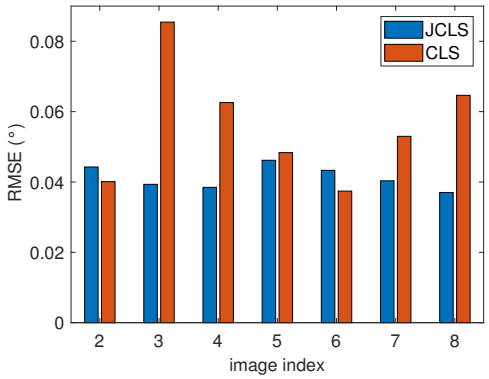


(b) azimuth 176°

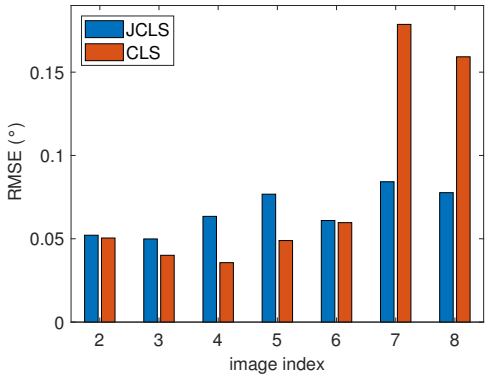


(c) azimuth 356°

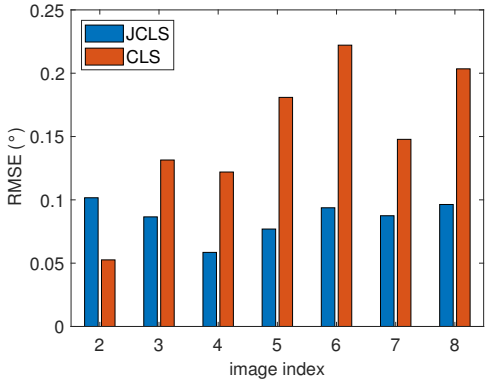
Fig. 4 RMSE (°) for each slave of the JCLS and CLS algorithms. The patch size is set to 25×25 pixels and subplots refer to 8 SAR images acquired at eight different elevations (marked by the image index). Moreover, three different tests are conducted fixing the azimuth angles to a) $0°$, b) $176°$, and c) $356°$. A total of 100 Monte Carlo runs is performed randomly selecting the rotation angles in the interval $[-2°, 2°]$.



(a) azimuth 0°



(b) azimuth 176°



(c) azimuth 356°

Fig. 5 RMSE (°) for each slave of the JCLS and CLS algorithms. The patch size is set to 30×30 pixels and subplots refer to 8 SAR images acquired at eight different elevations (marked by the image index). Moreover, three different tests are conducted fixing the azimuth angles to a) $0°$, b) $176°$, and c) $356°$. A total of 100 Monte Carlo runs is performed randomly selecting the rotation angles in the interval $[-2°, 2°]$.

4 Conclusions

This paper has focused on the design of new coregistration algorithm for multitemporal SAR images. The focus of the devised method is the joint estimation of the registration parameters for all slave images, namely accounting at the design stage the respective displacements between slaves. Based on the exploitation of the cross-cross-correlations, the derived method is capable of ensuring a more robust behavior of the registration algorithm for all the involved slave images in the considered stack. Results conducted on real-recorded data have demonstrated the validity of the devised model.

References

[1] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, 'A Tutorial on Synthetic Aperture Radar', IEEE Geoscience and Remote Sensing Magazine, vol. 1, no. 1, pp. 6-43, 2013.

[2] Ms P. S. Tondewad and Ms M. P. Dale, 'Remote Sensing Image Registration Methodology: Review and Discussion', Procedia Computer Science, vol. 171, pp. 2390-2399, 2020.

[3] H. Goncalves, L. Corte-Real, and J. A. Goncalves, 'Automatic Image Registration Through Image Segmentation and SIFT', IEEE Transactions on Geoscience and Remote Sensing, vol. 49, no. 7, pp. 2589-2600, July 2011.

[4] Zhili Song, Shuigeng Zhou, and Jihong Guan, 'A Novel Image Registration Algorithm for Remote Sensing under Affine Transformation', IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 8, pp. 4895-4912, 2013.

[5] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, 'SARSIFT: A SIFT-Like Algorithm for SAR Images', IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 1, pp. 453-466, 2015.

[6] S. Paul and U. C. Pati, 'A Block-Based Multifeature Extraction Scheme for SAR Image Registration', IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 9, pp. 1387-1391, September 2018.

[7] E. Sansosti, P. Berardino, M. Manunta, F. Serafino, and G. Fornaro, 'Geometrical SAR Image Registration', IEEE Transactions on Geoscience and Remote Sensing, vol. 44, no. 10, pp. 2861-2870, October 2006.

[8] D. Li and Y. Zhang, 'A Fast Offset Estimation Approach for InSAR Image Subpixel Registration', IEEE Geoscience and Remote Sensing Letters, vol. 9, no. 2, pp. 267-271, 2011.

[9] L. Pallotta, G. Giunta, and C. Clemente, 'Subpixel SAR Image Registration Through Parabolic Interpolation of the 2-D Cross Correlation', IEEE Transactions on Geoscience and Remote Sensing, vol. 58, no. 6, pp. 4132-4144, 2020.

[10] L. Pallotta, G. Giunta, and C. Clemente, 'SAR Image Registration in the Presence of Rotation and Translation: A Constrained Least Squares Approach', IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 9, pp. 1595-1599, 2021.

[11] L. Pallotta, G. Giunta, C. Clemente, and J. J. Soraghan, 'SAR Coregistration by Robust Selection of Extended Targets and Iterative Outlier Cancellation', IEEE Geoscience and Remote Sensing Letters, vol. 19, no. 4501405, pp. 1-5, 2022.

[12] L. Pallotta and G. Giunta, 'Accurate Delay Estimation for Multi-Sensor Passive Locating Systems Exploiting the Cross-Correlation between Signals Cross-Correlations', IEEE Transactions on Aerospace and Electronic Systems, pp. 1-1, 2021.

[13] K. El-Darymli, P. McGuire, D. Power, and C. R. Moloney, 'Target Detection in Synthetic Aperture Radar Imagery: A State-Of-The-Art Survey', Journal of Applied Remote Sensing, vol. 7, no. 1, pp. 071598, 2013.

[14] E. Ertin, C. D. Austin, S. Sharma, R. L. Moses, and L. C. Potter, 'GOTCHA Experience Report: Three-Dimensional SAR Imaging with Complete Circular Apertures', in Algorithms for Synthetic Aperture Radar Imagery XIV. International Society for Optics and Photonics, 2007, vol. 6568, p. 656802.