A Combination of Feature Tracking and Pattern Matching with Optimal Parametrization for Sea Ice Drift Retrieval from SAR Data

Anton Andreevich Korosov, Pierre Rampal

A Combination of Feature Tracking and Pattern Matching with Optimal Parametrization for Sea Ice Drift Retrieval from SAR Data

Anton Andreevich Korosov * and Pierre Rampal

Nansen Environmental and Remote Sensing Center, Bergen, 5009, Norway; Pierre.Rampal@nersc.no
* Correspondence: anton.korosov@nersc.no; Tel.: +47-41334568

Abstract: Sea ice drift strongly influences sea ice thickness distribution and indirectly controls air-sea ice-ocean interactions. Estimating sea ice drift over a large range of spatial and temporal scales is therefore needed to characterize the properties of sea ice dynamics and to better understand the ongoing changes of the climate in the polar regions. An efficient algorithm is developed for processing SAR data based on the combination of feature tracking (FT) and pattern matching (PM) techniques. The main advantage of the combination is that the FT rapidly provides the first guess estimate of ice drift in a few unevenly distributed keypoints, and PM accurately provides drift vectors on a regular or irregular grid. Thorough sensitivity analysis of the algorithm is performed, and optimal sets of parameters are suggested for retrieval of sea ice drift on various spatial and temporal scales. The algorithm has rather high accuracy (error is below 300 m) and high speed (the time for one image pair is 1 min), which opens new opportunities for studying sea ice kinematic processes. The ice drift can now be efficiently observed in the Lagrangian coordinate system on an irregular grid and, therefore, used for pointwise evaluation of the models running on unstructured meshes or for assimilation into Lagrangian models.

Keywords: sea ice drift; feature tracking; pattern matching; Sentinel-1; SAR

1. Introduction

Sea ice drift strongly influences sea ice thickness distribution and, as a consequence, indirectly controls air-sea ice-ocean interactions, e.g., energy fluxes [1]. In both the Arctic and Antarctic, sea ice ends drifting towards lower latitudes, where it finally melts. Sea ice melt represents a significant source of surface fresh water and plays an important role in the dynamics of the World Ocean. Estimating sea ice drift over a large range of spatial and temporal scales is therefore needed to understand and characterize the properties of sea ice dynamics at large, but also to help better understand the ongoing changes of the climate in the polar regions and beyond [2].

The need for more accurate and high-resolution near-real-time sea ice drift estimates is becoming more crucial than ever before, in particular in the context of the ongoing shrinking of the Arctic sea ice cover [3]. Indeed, as a consequence of the sea ice decline, we observed a positive trend in the volume of marine transport [4] and offshore activities in the Arctic, which are likely to accelerate in light of the predicted expansion of open water [5]. We also note in this context the importance for current or future sea ice models used in forecasting platforms to simulate sea ice drift accurately in order to plan operations in ice-covered waters and reduce the risk of incidents. Fast-changing sea ice brings hazards that are still a challenge to predict with today's available observing and modeling systems [6]. As an example, an exploratory oil drilling in the Chukchi Sea was suddenly halted in summer 2012 when an unpredicted incursion of broken-up ice threatened the safety of personnel and vessels. Such events

will become more common as Arctic maritime activity grows and as the sea ice gets mechanically weaker [7] and more mobile [8]. The combination of an efficient algorithm to retrieve sea ice drift from Synthetic Aperture Radar (SAR) images with a state-of-the-art forecasting platform based on a skillful sea ice model, especially with respect to sea ice kinematics simulation, appears like one possible way to go in order to tackle the aforementioned challenge.

Assimilation of sea ice deformation derived from high-resolution (SAR) sea ice drift into sea ice models is becoming possible if the model uses, e.g., an elasto-brittle rheology and the local level of damage as a model variable [9,10]. The combination of observations with such models looks promising as it can significantly improve the sea ice drift and deformation forecasting skills. However, in order to be able to combine observations and models efficiently, both observed and simulated sea ice drift have to capture the known characteristics of sea ice drift at a given spatial and temporal scale. With the current satellite observations available and the last generation of sea ice models, one should target to provide accurate daily drift of sea ice at 1-km resolution in near-real-time over the largest possible area.

The motion of sea ice is mainly driven by winds and ocean currents, but there does not exist a simple linear relationship between them, especially in winter when the ice is compact and its mechanically-induced sea ice internal stresses are high. Indeed, as opposed to current and wind fields that are relatively smooth over scales larger than kilometers, the sea ice motion field is characterized by the presence of sharp and extremely localized spatial gradients. From a mechanical point of view, those gradients can be related to the brittle behavior of sea ice, which should be viewed as the intrinsic response of solids to strong-enough external mechanical forcing. When strong winds blow over sea ice, the sea ice cover breaks up along fractures, along which the two individual blocks of ice are moving independently and possibly in opposite directions. Ice motion gradients therefore correspond to areas of convergence, divergence or shear deformation associated with the formation of ridges, opening/closing of leads and shear along faults, respectively. The brittle behavior of sea ice is at the origin of the scaling properties of observed sea ice deformation fields, i.e., heterogeneity (in space) and intermittency (in time) [11]. Since sea ice motion in the high ice concentration areas is controlled by the aforementioned deformation fields, sea ice displacement estimates are scale dependent, as well, meaning that their correct interpretation requires their spatial and temporal scales of estimation to be known.

This poses one of the difficulties of estimating sea ice drift from satellite remote sensing data, as the employed images have their own inherent spatial resolution, and the geometry of the satellite orbit (or orbits of multiple satellites) defines the actual frequency of observations. In addition, the algorithms to derive sea ice drift introduce their own spatial and temporal scales of "averaging" throughout the drift estimation procedure.

That being said, sea ice drift has been retrieved from various sources of satellite remote sensing data over the last few decades. The maximum cross-correlation (MCC) approach has been applied for automatic processing of the optical data from the Advanced Very High Resolution Radiometer (AVHRR) since 1984 [12]. Nowadays, MCC is adapted for processing low resolution passive microwave [13], scatterometer [14] and high resolution SAR data [15–17]. Other approaches for motion detection on 2D data are also applied, including feature tracking [18], optical flow [19,20] and the combination of different methods [21,22]. These studies focus on methodological aspects of the algorithms, and the questions of temporal and spatial scales of ice drift are seldom addressed.

An additional challenge emerging recently is the enormous amount of data available from the SAR instrument onboard the two European Space Agency (ESA) satellites Sentinel-1A and B [23]. Observations from these two satellites cover the Arctic Ocean at least once or, in very high latitudes, twice and even thrice per day with a spatial resolution of 80 m providing 100 TB of data every day [24]. The unprecedented flow of data sets new demands on the efficiency of the ice drift retrieval algorithms.

The goal of our study is to develop an efficient algorithm for processing SAR data based on the combination of feature tracking (FT) and pattern matching (PM) techniques and to study its

applicability for retrieval of sea ice drift on various spatial and temporal scales. The main advantage of the combination is that the FT rapidly provides the first guess estimate of ice drift in a few unevenly-distributed keypoints, and PM accurately provides drift vectors on a regular grid. The algorithm is calibrated/validated for processing SAR data primarily from Sentinel-1 satellite. It is supplied with a methodology for sensitivity testing, which can also be applied to find optimal parametrization for processing data from other SAR instruments.

The paper is organized as follows: the section 'Data' briefly describes available satellite and in situ data; the section 'Methodology' provides a description of the ice drift retrieval algorithm, sensitivity tests and the validation approach; the section 'Results' shows the calibration/validation results; the section 'Discussion' highlights the main advantages of the algorithm and focuses on interpretation of the algorithm performance in three use cases.

2. Data

2.1. Sentinel-1A Data

For calibration and validation of the algorithm, 322 Sentinel-1 SAR images were downloaded from the ESA Sentinel Scientific Hub [24]. In our study, we utilized Level-1 (L1) data in Extra Wide Swath mode (EW) from the ground product in medium resolution (GRDM, pixel spacing is 40 m) in cross-polarization (1SDH)[25]. Radiometric calibration was applied to the digital numbers (*DN*) provided in the TIFF file and calibration look-up-table (*A*) provided in accompanying metadata XML files in order to compute the normalized radar backscatter cross-section (σ^0) in dB [26]:

$$\sigma^0 = 10log_{10}(DN^2/A^2) \tag{1}$$

Values of σ^0 were scaled to the range 0–255 and converted from the 32-bit floating-point to the eight-bit integer data type compliant with the OpenCV library [27,28] using the following formula:

$$\sigma_8^0 = 255(\sigma^0 - \sigma_{MIN}^0)/(\sigma_{MAX}^0 - \sigma_{MIN}^0), \{\sigma_8^0 \in \mathbb{Z}|0 \leq \sigma_8^0 \leq 255\} \tag{2}$$

where σ_{MAX}^0, σ_{MIN}^0 are minimum and maximum values of σ^0 identified individually for HH and HV polarizations (see Sections 3.3 and 4.1).

The spatial resolution was reduced by down-sampling by two in both directions (pixels size equals 80 × 80 m) using averaging of the nearest pixels in order to reduce speckle noise. Images were used in swath projection and were not collocated on the same grid (reprojected).

The ground control points (GCPs) provided in the accompanying metadata files were used to transform the coordinates of the features identified on the image from the image coordinate system (rows, columns) to the geographic coordinate system (longitude, latitude) and back. Geolocation accuracy of the GCPs was estimated by manual comparison of the position of a reference point (a sharp tip of a small island in Svalbard archipelago) selected on one image to the position of this point on 30 other images acquired within one month. The average difference between positions of the reference point (a proxy for the geolocation accuracy) proved to be 4.9 ± 3.3 pixels (197 ± 135 m).

All preprocessing steps including radiometric calibration, downsampling and geolocation were performed using the Nansat package [29,30].

2.2. Sea Ice Drifters

During the N-ICE2015expedition, 42 buoys of various types were deployed from helicopter or the research vessel Lance in an ice-covered area north of Svalbard in January–February and April–May 2015 [31,32]. GPS receivers of the buoys were set to record position every hour (or every three hours on some buoys) with average accuracy of 50 m [33].

The buoys were deployed in ice pack, but during the drift, the ice concentration was constantly diminishing, and eventually, some of the buoys ended up in open water, where the FT or PM algorithms

are not applicable (see the map of buoy position in Figure 1). Eight hundred seventy seven buoy positions were selected as "valid" for our analysis only if a buoy was positioned on the ice (sea ice concentration from Ocean and Sea Ice Satellite Application Facility (OSI SAF) [34] must be above 85%), and the difference in time between satellite image acquisition and buoy GPS reading must be less that one hour. Valid match-ups were randomly split into two equal parts: one part was used for sensitivity tests (see Section 3.3), another part for independent validation (see Section 3.2). Buoy coordinates in the spherical coordinate system (longitude, latitude) were converted into the Cartesian coordinate system of a matching SAR image (rows, columns) utilizing the georeference information from the corresponding GCPs using Nansat [29].
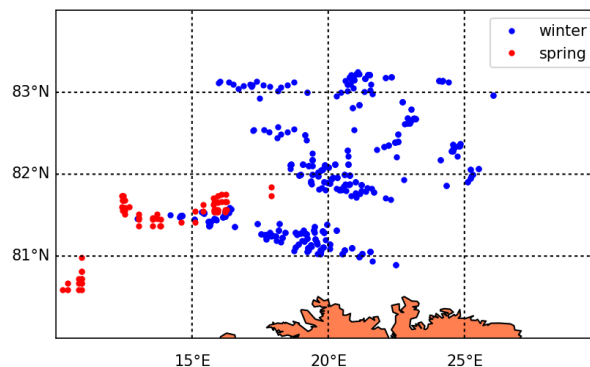
Figure 1. Position of buoys used for validation and deployed in winter (**blue dots**) or in spring (**red dots**).

3. Methodology

3.1. Combination of Feature Tracking and Pattern Matching

The proposed algorithm is based on subsequent application of the feature tracking [35] and pattern matching [36] techniques (flowchart shown in Figure 2). At the first step, FT is applied to a pair of Sentinel-1 SAR images covering the same area, but at a given time interval δT. The detailed description of the FT method is provided in [18], and here, we discuss only the steps important for understanding of the combined algorithm.

FT automatically identifies keypoints (distinct features) on the images and describes each keypoint with a vector of 256 binary descriptors [35]. The keypoints are matched using the brute-force matching (all keypoints on one image are compared to all keypoints on another image) and the Hamming distance (number of the descriptors in the vector with different values). For each keypoint on the first image, a ratio between the smallest and the second smallest Hamming distance to keypoints on the second image is computed. If the ratio is below a threshold (L_T), the keypoints with the smallest Hamming distance are considered as matched.

As a result, FT provides the location of matched features characterized by pixel/line coordinates on the first image (x_1, y_1) and corresponding coordinates on the second image (x_2, y_2). Two filters are applied to identify rogue vectors: First, vectors where ice drift exceeds a threshold F_1 m/s are omitted as incorrect. Second, in order to remove inconsistent drift vectors, the values of x_1 and y_1 are approximated using the second order polynomial functions of x_2 and y_2 computed by the least squares method (Equation (3)). The approximated coordinates are compared with the actual ones, and the vectors that have a start position further than threshold F_2 km away from the simulated point are discarded.

Although FT is very fast, the generated vectors are heterogeneously distributed in space. Note that in this form, this set of drift vectors cannot be used further for studying, e.g., sea ice deformation, as it is a scale-dependent dynamical process. Therefore, the ice drift is to be further estimated on a regular grid with a given resolution.
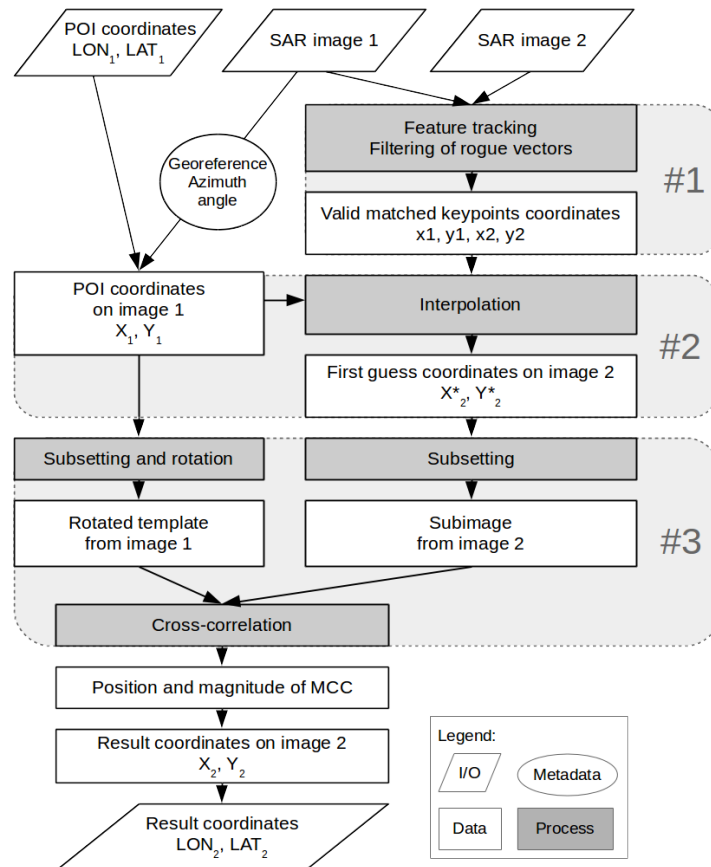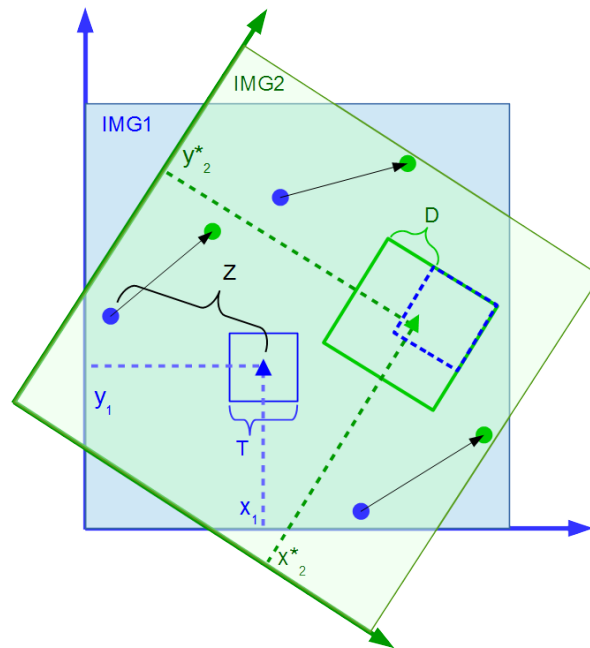
Figure 2. Flowchart of the combination of feature tracking and pattern matching algorithms.

At the second step (see Figure 2), the FT results are approximated at the regular grid or, generally speaking, in any point inside the overlap between the two SAR images. The approximation by a linear interpolation [37] is performed in the geographic space between the matched keypoints (Figure 3, blue dots from Image 1 and green dots from Image 2). In the surrounding area, the approximation is performed using *N*-th-order polynomial:

$$x_2^* = \sum_{i,j=0}^{n} A_{i,j} x_1^i y_1^j, i \leq N, j \leq N, i + j \leq N$$

$$y_2^* = \sum_{i,j=0}^{n} B_{i,j} x_1^i y_1^j, i \leq N, j \leq N, i + j \leq N$$

(3)

where x_2^*, y_2^* are the extrapolated coordinates of point from Image 2, $A_{i,j}$ and $B_{i,j}$ are polynomial coefficients derived using the least squares methods applied to all valid FT results and N is order of the polynomial discussed in Section 4.1. By definition, the approximation of the ice drift is inaccurate. The distance to the nearest keypoint (Figure 3, distance *Z*) may serve as a measure of the approximation uncertainty (Figure 3, distance *D*). The established approximation can now be used for estimating coordinates X_2^*, Y_2^* of the first guess point in Image 2 (Figure 3, green triangle) based on the coordinates X_1, Y_1 of any points of interest (POI) on Image 1 (Figure 3, blue triangle). Hereafter, block letters are used to denote sets of coordinates (e.g., x_1, y_1 denote coordinates of all output vectors from FT); capital letters are used to denote a single value (e.g., X_1, Y_1 denote coordinates of a point on the example image).

Figure 3. Explanation of the combination of feature tracking (FT) and pattern matching (PM) algorithms. **Blue (green)** semitransparent square: Footprint of Image 1 (2) in coordinates of Image 1. **Blue (green)** dots: Keypoints identified on Image 1 (2) by FT. Black arrows: Ice drift vectors identified by FT. **Blue** triangle: Point of interest on Image 1. **Green** triangle: The first guess point on Image 2. **Blue** square: Original template from Image 1. **Blue** square with dashed outline: Rotated template. **Green** square: sub-image from Image 2. Letters *Z*, *T* and *D* denote the distance to the nearest keypoint, template size and searching range.

At the third step (see scheme on Figure 2), the approximated ice drift is used as the first guess for the precise PM technique (Figures 3 and 4) in the following order. Azimuth angle (α_0) between the SAR scenes is calculated based on the georeference information. A template *t* (blue empty square on Figure 3) is subset from Image 1 around a starting point of interest with coordinates X_1, Y_1 (blue triangle on Figures 3 and 4A) and size $T \times T$ pixels. The corresponding approximate coordinates X_2^*, Y_2^* (green triangle on Figures 3 and 4C) are found using the aforementioned approximation. A larger sub-image *s* (green empty square on Figures 3, 4C) is subset from Image 2 with the center at X_2^*, Y_2^* and size $S \times S$ pixels.

The template *t* is rotated within the range of angles $[\alpha_0 - \beta, \alpha_0 + \beta]$ with step $\Delta\beta$ (Figure 4B). For each rotation, cross-correlation matrix *r* is calculated between the rotated template t_{ROT} and *s* (Figure 4D). Hereafter, the size of the cross-correlation matrix $D = S - T$ is referred to as a searching range. The rotation that provides maximum cross-correlation is recorded. The position (ΔX_2, ΔY_2) of the cross-correlation maximum r_{MAX} is found in the cross-correlation matrix *r* [36]. The result coordinates of ice drift on Image 2 are estimated as $X_2 = \Delta X_2 + X_2^*$ and $Y_2 = \Delta X_2 + Y_2^*$ (blue triangle in Figure 4C). Ice drift components U and V are found by converting X_1, Y_1 and X_2, Y_2 into geographic coordinates and calculating corresponding eastward and northward displacement. If the cross-correlation maximum r_{MAX} is below a predefined threshold r_{MIN}, the derived drift vector is discarded as unreliable.

The output of the combined algorithm consists of four values: *u*, ice drift eastward component; *v*, ice drift westward component; r_{MAX} maximum normalized cross-correlation; α, angle of ice rotation. We note that the following parameters of the algorithm are affecting its accuracy, effective resolution and speed: *T* size of the template; *D*, searching range; β, range of rotation; $\Delta\beta$, rotation step; *N*, order of the first guess polynomial; r_{MIN}, minimum required cross-correlation; F_1 and F_2, thresholds for rogue vector filters of the FT algorithm (see the beginning of the section).
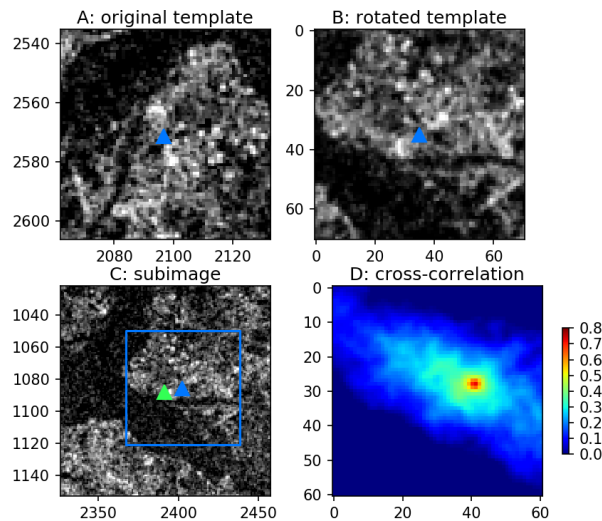
Figure 4. Explanation of the maximum cross-correlation (MCC) method with rotation. (**A**) Template *t* (71 × 71 pixels) is taken from Image 1 around the starting point shown as the blue triangle; (**B**) the template is rotated by 210° (t_{ROT}); (**C**) sub-image *s* (101 × 101 pixels) is taken from Image 2 around the first guess point (green triangle); the blue boundary and the blue dot indicate the position and the center of the template, which corresponds to the best match; (**D**) the cross-correlation matrix *r* is computed from the rotated template t_{ROT} and the subimage *s*.

3.2. Comparison with Sea Ice Drifters' Data

Calibration and validation of the combined algorithm was performed using the ice drifting buoys (Section 3.2). Image pairs overlapping in time and space with the position of the buoys were found using the Nansen-Cloud software [38]. A validation match-up consisted of two ice drift vectors: an in situ drift vector has been calculated as the difference between the starting and ending positions of the buoy closest to the image's acquisition time. A satellite drift vector has been calculated as the difference between the starting position of the corresponding buoy and ending position derived with the combined FT-PM algorithm. The retrieval error Ewas calculated as the distance between the derived end position on the second image and the buoy GPS position using the following equation:

$$E = \sqrt{(x_2^f - x_2^b)^2 + (y_2^f - y_2^b)^2} \tag{4}$$

where x_2^f and y_2^f are final coordinates retrieved by the combined algorithm and x_2^b and y_2^b are the buoys coordinates.

Under the assumption that the template centered at the initial buoy position should exactly match the subimage centered at the next buoy position, we first have tested the accuracy of collocation of buoys and SAR image pattern in the following way. The PM method was initiated from the buoys' position (blue triangle in Figure 4A, green triangle in Figure 4C) and the displacement between the actual position of the cross-correlation maximum (blue triangle in Figure 4C) and the buoy position was calculated using Equation(4). The dependence of the PM error *E* and number of match-ups *N* on the maximum correlation R was studied.

For the validation of the combined algorithm, the template was also centered at the initial buoy position. The subimage center and searching range *D* were set based on the approximated first guess from the FT algorithm. The *X* and *Y* components of the combined algorithm error ($E_X = x_2^f - x_2^b$, $E_Y = y_2^f - y_2^b$), as well as the total error *E* were studied.

3.3. Sensitivity Tests

The effect of various parameters on the accuracy and speed of the algorithm was analyzed in three sensitivity tests. The first test included assessment of the σ^0_{MIN} and σ^0_{MAX} (Section 2.1) impact on the number of identified keypoints and matched features. Values of σ^{0HH} were scaled within the following limits: $\sigma^{0HH}_{MIN} \in \{-24, -22, -20, -18, -16, -14\}$ dB, $\sigma^{0HH}_{MAX} \in \{-12, -10, -8, -6, -4, -2\}$ dB; and σ^{0HV} was scaled within the following limits: $\sigma^{0HV}_{MIN} \in \{-34, -32, -30, -28, -26, -24\}$ dB, $\sigma^{0HV}_{MAX} \in \{-20, -18, -16, -14, -12, -10\}$ dB. For each combination of σ^0_{MIN} and σ^0_{MAX} 10,000 keypoints were identified by the FT algorithm on each pair of the validation Sentinel-1A images, and the number of matching features per image pair was recorded. Keypoints were matched using four values of the Lowe ratio test L_T [35]: 0.6, 0.65, 0.7, 0.75.

The second test aiming at studying the approximation accuracy (using linear or polynomial interpolation) of the first guess was run for all pairs of Sentinel-1A images. For each pair, the FT algorithm was used to identify 100,000 keypoints, to match the keypoints and filter rogue vectors (Section 3.1). The remaining matched keypoints were randomly split into two equal subsets: the 'training' subset was used for gridding of the first guess, and the 'testing' subset for validation. Gridding was performed using either only 200, or 1000 or 2000 keypoints and using either the first order polynomial, or the second order polynomial, or the linear interpolation on a triangulated grid. The gridded x_2 and y_2 coordinates were compared with the coordinates of the keypoints from the testing subset. The gridding error was calculated using Equation (4), and the distance to the nearest keypoint Zwas recorded. The error E was averaged for each combinations of number of keypoints within realistic ranges of $Z \in [0, 20]$ km.

The third test aimed at studying how the accuracy of the PM method depends on the template size T, searching range D and the distance to the nearest keypoint Z. For each match-up, a template with size $T \in \{20, 40, 80\}$ pixels was centered at the buoy initial position. All available matched keypoints were used for gridding. If the initial buoy position was between the matched keypoints, the gridding was performed using linear interpolation, otherwise using the second order polynomial. The PM algorithm was applied with searching range $D \in \{5, 10, 20, 40, 80\}$. The combined algorithm error E was calculated using Equation (4), and the nearest keypoint distance Z was recorder. The error E was averaged for each combination of D and T and for four realistic ranges of Z: $0 \leq Z < 5$ km, $5 \leq Z < 10$ km, $10 \leq Z < 20$ km, $20 \leq Z < 40$ km.

The processing speed of the combined algorithm was tested on a personal computer with a 2.7-GHz processor clock frequency utilizing only a single CPU core with 4 GB RAM. FT was tested on one image pair with 80% overlap with detection and matching of 5000, 10,000, 50,000, 100.,000 and 200,000 initial keypoints. Generation of the first guess was tested with 1000, 2000, 5000 and 10,000 of matched keypoints. PM was tested with template size $T \in \{20, 40, 80\}$ and searching range $D \in \{10, 20, 40, 80\}$.

4. Results

4.1. Sensitivity of the Algorithm

The first sensitivity test (Section 3.3) showed that values of σ^0 follow the normal distribution (Figure 5) with a peak at -16 dB for HH and -26 dB for HV polarizations. Typical ranges (95% of all values) are $\sigma^{0HH} \in [-27, -2]$ dB, $\sigma^{0HV} \in [-34, -16]$ dB. The number of matched features is maximum (Figure 6) when σ^{0HH} is scaled within $[-18, -8]$ dB and σ^{0HV} is scaled within $[-28, -14]$ dB. The number of features matched on the HV channel is almost ten-times higher than on HH.

The value of the Lowe ratio test L_T significantly affects the number of matched and the ratio of rejected keypoints. With L_T increasing from 0.6–0.75, the average number of matched keypoints linearly increases from 1700 ± 2000–4000 ± 4000 points for one scene, while the average ratio of rejected keypoints decreases from 0.97 ± 0.12, to 0.96 ± 0.15, to 0.94 ± 0.2, to 0.85 ± 0.25.
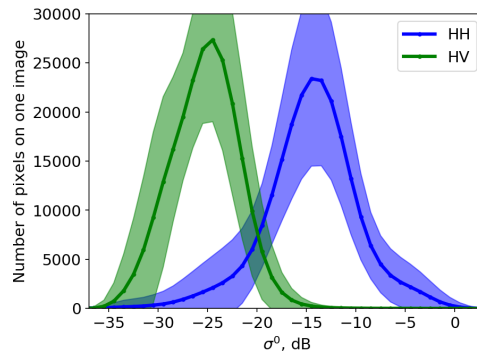
Figure 5. Average probability density function of σ^{0HH} (**blue**) and σ^{0HV} (**green**) on all validation images. Filled areas indicate the standard deviation of the average number of pixels.
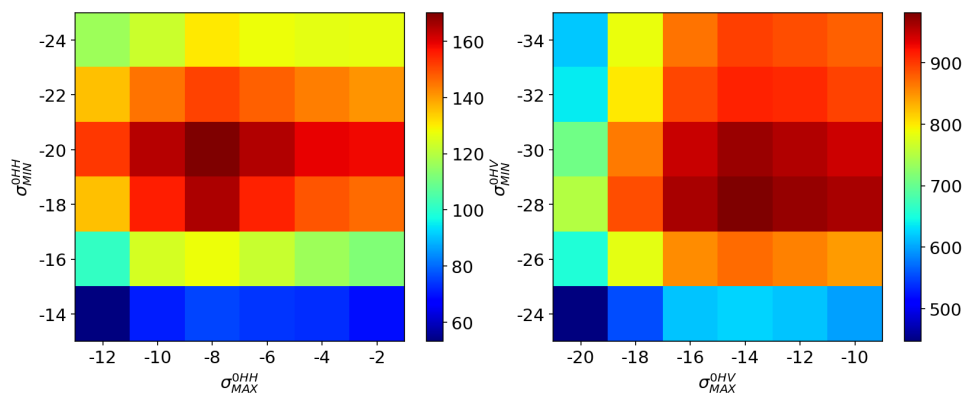

Figure 6. Dependence of the average number of matched keypoints per image pair (shown by color) on σ^0_{MIN} (*Y*-axis) and σ^0_{MAX} (*X*-axis) for HH (**left**) and HV (**right**) polarizations.

The second sensitivity tests (Section 3.3) showed that in the case of polynomial interpolation, the error *E* grows with an increasing distance from the nearest keypoint *Z* (Figure 7). The first order polynomial expectedly provides the worst results: the minimum achievable error is 500 m; it rapidly grows with distance, and when distance exceeds 5 km, the error exceeds 1 km. The second order polynomial provides more accurate estimates: the error reaches 1 km only when the distance is up to 10 km. Linear interpolation provides the highest accuracy: the error of 500 m is not reached even at the distance of 15 km from the nearest point. The number of matched keypoints does not affect the results significantly (not shown in the figure).

The third sensitivity test showed how significant was the role played by all of the three studied parameters taken together (template size *T*, searching range *D* and distance to the nearest keypoint *Z*) in the combined algorithm accuracy. For the HH channel (Figure 8), the template must be larger than 20 pixels, otherwise the error starts to grow rapidly if the searching range is too large (*D* > 40). With a large enough template (*T* > 20) and dense keypoints (*Z* < 10 km), high retrieval accuracy (*E* < 400 m) is obtained for any searching range above 10 pix. With a medium density of keypoints (10 < *Z* < 20 km), high accuracy is only obtained if the searching range is large (*D* > 40), while with sparse keypoints (*Z* > 20 km), the retrieval of sea ice drift from the HH channel is just not possible (*E* > 6 km).

The HV channel (Figure 9) seems more favorable for high resolution sea ice drift retrieval. With relatively high density of matched keypoints (*Z* < 10 km), the high accuracy (*E* < 300) is obtained even if the template is rather small (*T* = 20) and the searching range is small (*D* > 20). If the keypoints are sparser (10 < *Z* < 20 km), medium error (500 m) is obtained only if the searching range is large (*D* > 40). When the distance to the keypoint is too large (*Z* > 20 km), the algorithm fails (*E* > 7 km).
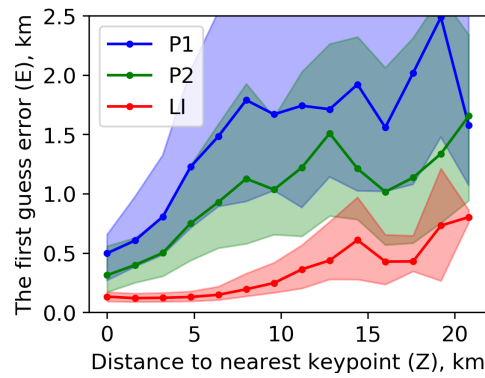
Figure 7. Dependence of the of the first guess error *E* (*Y*-axis) on the distance to the nearest keypoint *Z* (*X*-axis). The three lines show results for the first-order polynomial (**blue**), second-order polynomial (**green**) and linear interpolation (**red**). The filled area of the corresponding color indicates the standard deviation.
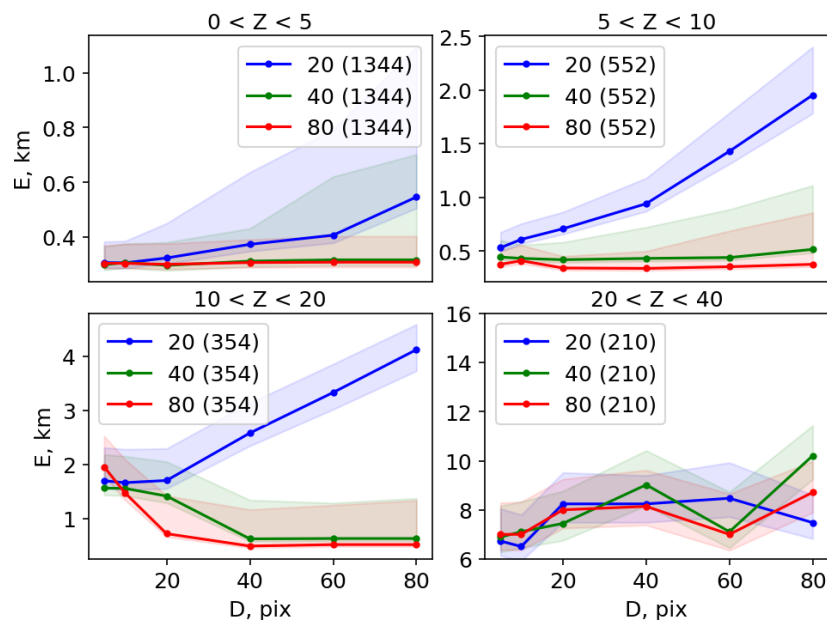
Figure 8. Dependence of the combined algorithm error *E* (*Y*-axis) on template size *T* (shown by color), on searching range *D* (*X*-axis) and distance to the nearest keypoint *Z* (four subplots) for HH polarization.
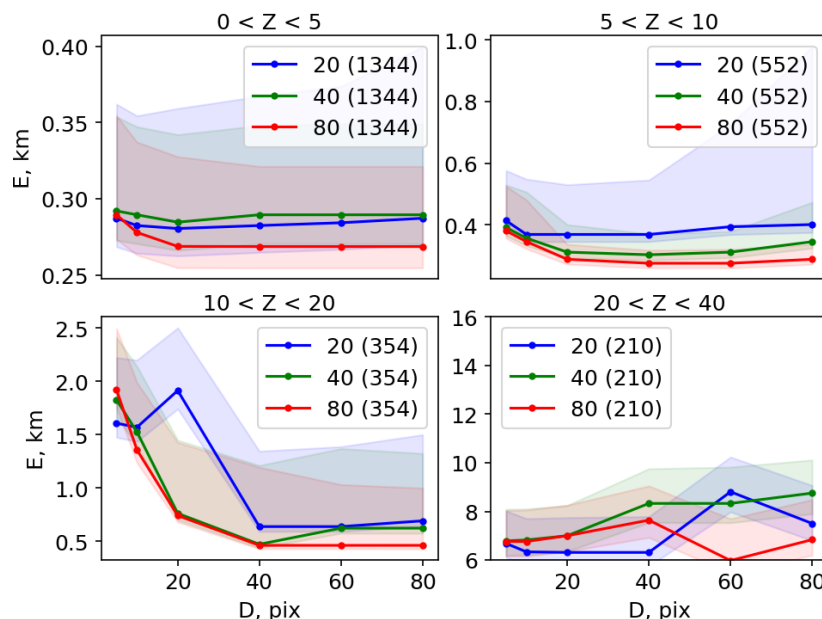
Figure 9. Dependence of the combined algorithm error *E* (*Y*-axis) on template size *T* (shown by color), on searching range *D* (*X*-axis) and distance to the nearest keypoint *Z* (four subplots) for HV polarization.

4.2. Efficiency of the Combined Algorithm

The speed benchmarks showed that the time required for FT is proportional to the square of the initial number of keypoints and may vary from 3–800 s (Table 1). Very little time (<3 s) is required for building the first guess. The time required for PM of one template is negligible and is proportional to the square of template size *T* and to the searching range *D* (Table 2).

Table 1. Time required for identification, matching and filtering of keypoints.

| Number of Initial Keypoints	5000	10,000	20,000	50,000	100,000	200,000
Time, s	3.5	5	11	50	190	815

Validation of the algorithm was performed on an independent subsample of match-ups with the set of the parameters designed for the highest possible accuracy (Table 3), but without considering any efficiency constraints. As shown in Figure 10, the combined algorithm is capable of deriving sea ice drift with high accuracy (*r* = 0.99, *E* = 286 ± 3 m) for a wide range of displacement magnitude (up to 300 km). The retrieved drift is negligibly biased (−39 and 71 m in X- and Y-directions), and the error is log-normal distributed.

Table 2. Dependence of PM speed (expressed as the number of ice drift vectors per 1 s) on template size *T* and searching distance *D*.

D	Template Size (*T*)		
	20	**40**	**80**
10	4660	2699	960
20	4463	2072	839
40	2413	1677	703
80	974	818	494

Table 3. Values of the algorithm parameters used for validation.

Parameter	Value
σ_{MIN}^{0HV}, dB	-28
σ_{MAX}^{0HV}, dB	-14
Initial number of keypoints	100,000
F_1	0.5 m/s
F_2	8 km
T	40
D	80
r_{MIN}	0.3
L_T	0.7
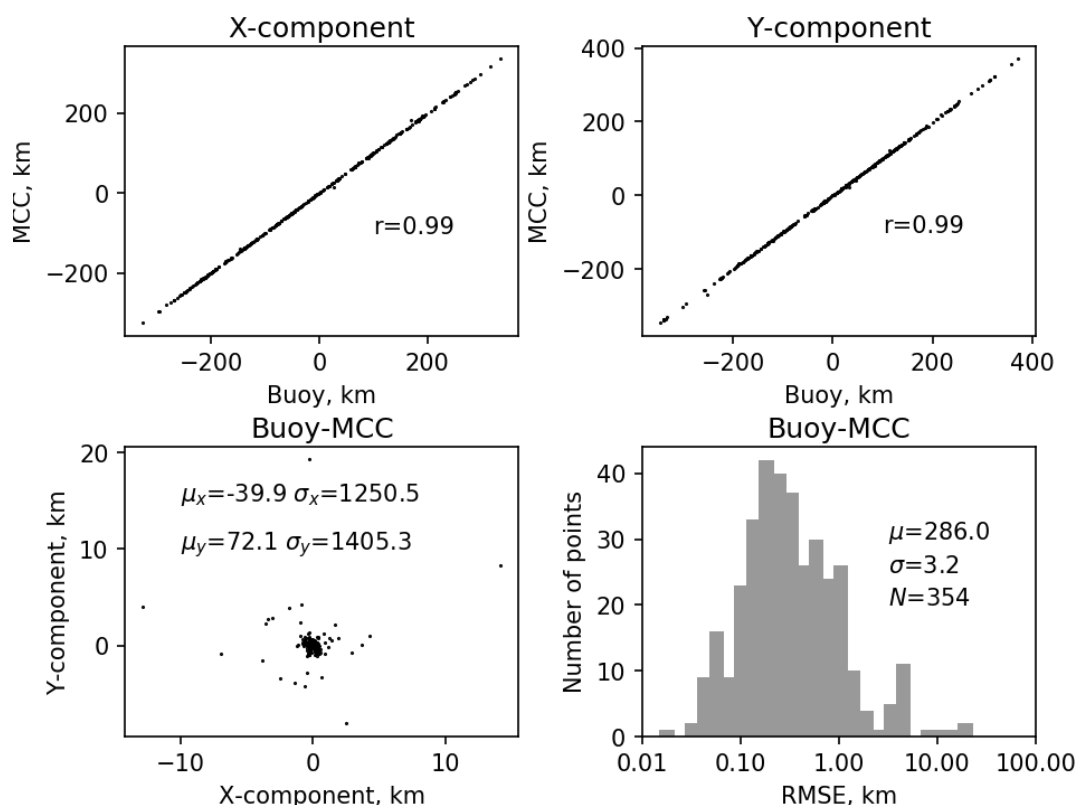
Figure 10. Results of validation on independent match-ups (not used in sensitivity tests). Upper scatterplots compare the X-component (left) and Y-component (right) of ice drift detected by in situ buoys (X) with drift derived by the combined algorithm (Y); *r* stands for correlation. The lower left scatterplot compares the X- and Y-components of difference between in situ and the combined algorithm drift; μ indicates the mean and σ the standard deviation. The histogram on the lower right shows the probability distribution of the retrieval error *E* (log-scale); *N* indicates the number of match-ups; μ the mean error (meters); σ the standard deviation (meters).

5. Discussion

One of the major advantages of FT and PM techniques' combination is the ability to retrieve the sea ice drift vector in any point of interest at very high speed. The algorithm's agility is due to several reasons: time-consuming spatial collocation (reprojection) of two SAR images is not needed; FT is very fast by nature; the first guess estimate from the FT results is computed only once for the entire image; the first guess helps to restrict time-consuming PM. The quickness of the algorithm opens new opportunities for studying sea ice kinematic processes. The ice drift can now be derived not only on

the traditional regular grid, but also on an irregular grid used by the models with the elasto-brittle rheology [9,10]. The classical Eulerian observations of ice drift can now be replaced by Lagrangian observations, when the location of a virtual ice drifter (or a set of drifters) evolves in time, so that the starting position of an ice drift vector changes at every step.

Several parameters affect the efficiency of the algorithm. Increasing the initial number of keypoints will decrease the FT speed, but will also create a denser field of matched keypoints, providing higher accuracy of the first guess, and allow higher resolution sea ice drift retrieval. Increasing the size of template T will decrease PM speed, but increase the robustness of the algorithm. Increasing the searching range D compensates for the low accuracy of the first guess (e.g., due to low number of initial keypoints), but also decreases speed. The actual requirements for accuracy and spatial resolution of the eventual sea ice drift product define the optimal parametrization of the algorithm. An optimum is achieved when the requirements are met within the least time possible. The optimal values of these parameters are summarized in the Table 4 for three example use cases presented below.

In the first use case, a pan-Arctic daily sea ice drift product is retrieved at a fixed grid with a spatial resolution of 6.5 km. Independent ice drift vectors at the given resolution can be generated if the template size is 80 × 80 pixels. At least 250 matched keypoint have to be identified on a 5000 × 5000 pixels image, so that the average distance between keypoints does not exceed 100 pixels (8 km). Given that on average, only 10% of found keypoints are successfully matched and points may be distributed unevenly, the FT should be initialized with at least 10,000 points. A searching range of 40 pixels will provide the required accuracy of 500 m. In such a scenario for an average sea ice area of 15×10^6 km^2, we will need around 350,000 vectors from approximately 100 Sentinel-1 pairs. Excluding reading of data, just the computation of ice drift will require ca. 40 min.

In the second use case, a regional daily sea ice drift product covering 700,000 km^2 (e.g., the Kara Sea) is generated at a 2-km spatial resolution with the required accuracy of 300 m. These requirements are met if at least 50,000 keypoints are initialized, the template size is 25 pix and the searching range is as large as 80 pixels. Retrieval of 175,000 vectors from four scenes will require ca. 7 min.

In the third use case, a virtual drifter is released in the Arctic, and its position is tracked from one SAR image to another one for 10 days in a row. To assure the best possible accuracy of the first guess, the FT is initialized with 100,000 keypoints, and a template of 40 pixels is matched within a 100-pix distance. That may provide accuracy of 300 m and will require 30 min.

Table 4. Summary of the combined algorithm application use cases.

| Use Case	1. Pan Arctic	2. Regional	3. Virtual Drifters
Grid type	Fixed, Eulerian	Fixed, Eulerian	Lagrangian
Resolution, km	6.4	2	Not applicable
ROI area, km	15,000,000	700,000	Not applicable
Required accuracy, m	500	300	300
Image pairs	400	4	10
Initial keypoints	10,000	50,000	100,000
Template size	80	25	40
Searching distance	40	80	100
Number of vectors	350,000	175,000	10
Time spent, s	2500	400	1900

It was previously reported [39] that the maximum value of cross-correlation r_{MAX} alone cannot be effectively used for assessing the uncertainty of PM algorithm. However, our comparison clearly shows (Figure 11) that the retrieval error E is decreasing from 1750 m–250 m with the correlation r_{MAX} increasing from 0.15–0.95 both for HH and HV channels. That allowed us to introduce a parameter r_{MIN} for the removal of rogue vectors with too low MCC. This parameter affects the number of retrieved drift vectors, and value of 0.3 seems to be optimal for keeping most of the vectors (85% for HH and 95% for HV) with the highest accuracy ($E < 500$ m). If r_{MIN} is reduced, an additional

rogue-vector detection procedure needs to be applied [40]. These results also confirm that the HV channel is more appropriate to use for sea ice drift estimation.

Calculation of MCC with several rotations of the template clearly reduces speed, but allows, first, to improve the robustness of retrieval (the correlation r_{MAX} is increased) and, second, identify consistent rotation of solid ice parcels (single ice floes or aggregates of many floes). An example map of combined sea ice drift and rotation (Figure 12) demonstrates how vectors with similar rotation are grouped together. In this example, rotation was applied in the range from $+20°--20°$ with a $3°$ step. That decreased speed by seven-times, but on average increased r_{MAX} by 15%. Specific values of these parameters depend strongly on the geographic region and ice dynamics.
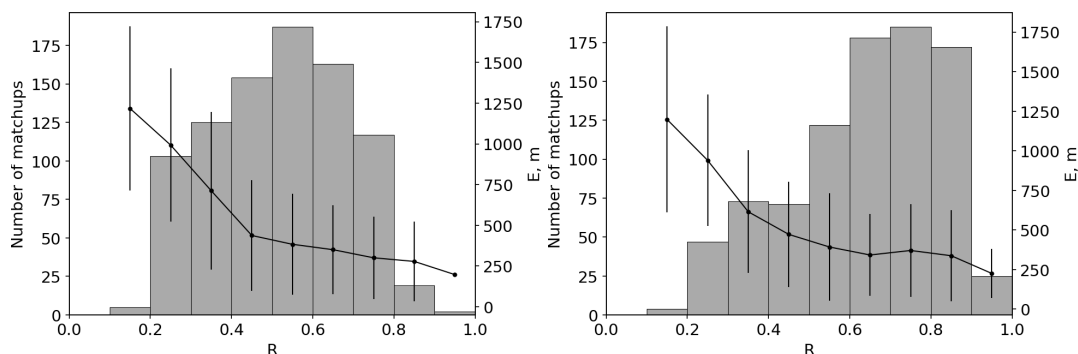
Figure 11. Dependence of the number of match-ups (left axis, gray bars) and retrieval error *E* (right axis, black solid line with error bars) on maximum cross-correlation (*X*-axis) for HH (**left**) and HV (**right**) polarizations.
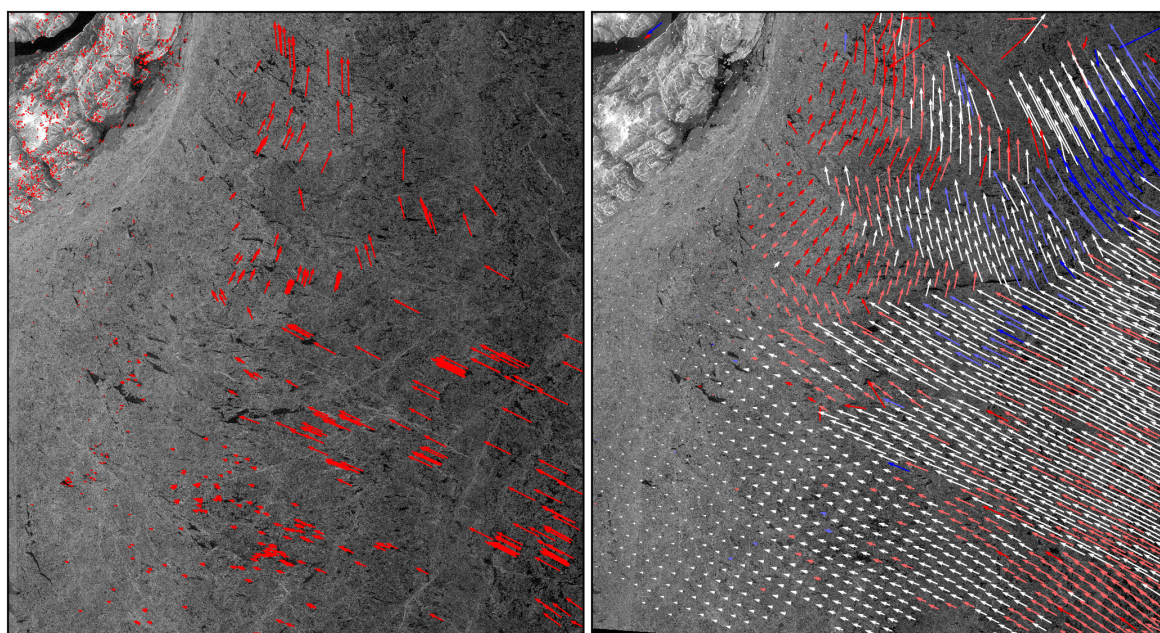
Figure 12. Pair of Sentinel-1A images of the Wandel Sea taken on 2 (**left**) and 4 (**right**) March 2016. Both images are in swath projection of the first image. Vectors from FT are displayed as red arrows on the left frame. Vectors from PM are displayed on the right frame with colors indicating rotation and ranging from blue ($-5°$) via white ($0°$) to red ($+5°$). The length of the vectors shows actual displacement of sea ice.

As can be seen in Figure 12, some rogue vectors still remain, and the next step of algorithm improvement could be to implement a better uncertainty estimate than using the *r*-value alone. Peakedness/flatness of the cross-correlation function, estimated as the second order spatial derivative

at the position of the identified maximum, may serve as a good candidate for that purpose [39]. Other improvements required in the future may include the following: rogue vector detection based on deformation [41]; proper filtering of small-scale drift anomalies that can be seen as spurious noise when computing the first order spatial derivative of the sea ice velocity field, i.e., sea ice deformation; different approximation of the first guess based on FT results that takes into account small-scale variations; improvement of the search for the highest MCC in 3D space (X,Y,rotation) using multivariate optimization technique.

6. Conclusions

A new algorithm based on the combination of feature tracking and pattern matching techniques has been developed for efficient retrieval of sea ice drift from SAR data. The algorithm has rather high accuracy ($E < 300$ m) and high speed (the time for one image is ca. 1 min). It outperforms the FT algorithm applied alone (the FT error is 563 m [18]) or together with interpolation (the FT interpolation error is up to 1500 m) and provides ice drift vectors on a predefined grid. Sources of the observed error include satellite image geolocation uncertainty, error in buoy GPS position and error of the PM algorithm *per se*.

The algorithm is calibrated/validated using 877 match-ups of satellite drift observations with ice drifter buoys in the Fram Strait where various ice types are present, including first-year ice, multi-year ice and fast ice. Optimal values of the algorithm parameters for retrieval of sea ice drift on various spatial and temporal scales are based on thorough sensitivity analysis. The algorithm application is illustrated on retrieval of sea ice drift and rotation from a pair of Sentinel-1A SAR images in March 2016. The provided methodology for testing the sensitivity of the algorithm parameters and choosing the optimal values can be applied to SAR data with other frequencies, polarizations and resolutions.

The algorithm is an open source Python code SeaIceDrift-0.3 available at GitHub [42]. The code is accompanied by an easy-to-understand example and unit tests and is integrated with the Travis-CI [43] and Coveralls [44] platforms.

Author Contributions: Anton Korosov conceived of and designed and performed the experiments. Anton Korosov and Pierre Rampal analyzed the experiment results. Anton Korosov and Pierre Rampal wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Vihma, T. Effects of Arctic Sea Ice Decline on Weather and Climate: A Review. *Surv. Geophys.* **2014**, *35*, 1175–1214.

2. Vaughan, D.G.; Comiso, J.C.; Allison, I.; Carrasco, J.; Kaser, G.; Kwok, R.; Mote, P.; Murray, T.; Paul, F.; Ren, J.; et al. Observations: Cryosphere. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y, Bex, V., Midgley, P.M., Eds.; Cambridge University Press: Cambridge, UK, 2013; pp. 317–382.

3. Parkinson, C.L. Global Sea Ice Coverage from Satellite Data: Annual Cycle and 35-Yr Trends. *J. Clim.* **2014**, *27*, 9377–9382.

4. Humpert, M.; Raspotnik, A. The future of Arctic shipping. *Port Technol. Int.* **2012**, *55*, 10–11.

5. Barnhart, K.R.; Miller, C.R.; Overeem, I.; Kay, J.E. Mapping the future expansion of Arctic open water. *Nat. Clim. Chang.* **2016**, *6*, 280–285.

6. Eicken, H. Ocean science: Arctic sea ice needs better forecasts. *Nature* **2013**, *497*, 431–433.

7. Gimbert, F.; Jourdain, N.C.; Marsan, D.; Weiss, J.; Barnier, B. Recent mechanical weakening of the Arctic sea ice cover as revealed from larger inertial oscillations. *J. Geophys. Res. Oceans* **2012**, doi:10.1029/2011JC007633.

8. Rampal, P.; Weiss, J.; Marsan, D. Positive trend in the mean speed and deformation rate of Arctic sea ice, 1979-2007. *J. Geophys. Res. Oceans* **2009**, *114*. C05013.

9. Bouillon, S.; Rampal, P. Presentation of the dynamical core of neXtSIM, a new sea ice model. *Ocean Model.* **2015**, *91*, 23–37.

10. Rampal, P.; Bouillon, S.; Ólason, E.; Morlighem, M. neXtSIM: A new Lagrangian sea ice model. *Cryosphere* **2016**, *10*, 1055–1073.

11. Rampal, P.; Weiss, J.; Marsan, D.; Lindsay, R.; Stern, H. Scaling properties of sea ice deformation from buoy dispersion analysis. *J. Geophys. Res. Oceans* **2008**, *113*, C03002.

12. Ninnis, R.M.; Emery, W.J.; Collins, M.J. Automated extraction of pack ice motion from advanced very high resolution radiometer imagery. *J. Geophys. Res. Oceans* **1986**, *91*, 10725–10734.

13. Lavergne, T.; Eastwood, S.; Teffah, Z.; Schyberg, H.; Breivik, L.A. Sea ice motion from low-resolution satellite sensors: An alternative method and its validation in the Arctic. *J. Geophys. Res. Oceans* **2010**, doi:10.1029/2009JC005958.

14. Girard-Ardhuin, F.; Ezraty, R. Enhanced Arctic Sea Ice Drift Estimation Merging Radiometer and Scatterometer Data. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 2639–2648.

15. Thomas, M.; Geiger, C.; Kambhamettu, C. High resolution (400m) motion characterization of sea ice using ERS-1 {SAR} imagery. *Cold Reg. Sci. Technol.* **2008**, *52*, 207–223.

16. Kwok, R.; Sulsky, D. Arctic Ocean Sea Ice Thickness and Kinematics: Satellite Retrievals and Modeling. *Oceanography* **2010**, *23*, 134–143.

17. Karvonen, J. Operational SAR-based sea ice drift monitoring over the Baltic Sea. *Ocean Science* **2012**, *8*, 473–483.

18. Muckenhuber, S.; Korosov, A.A.; Sandven, S. Open-source feature-tracking algorithm for sea ice drift retrieval from Sentinel-1 SAR imagery. *Cryosphere* **2016**, *10*, 913–925.

19. Petrou, Z.I.; Tian, Y. High-Resolution Sea Ice Motion Estimation With Optical Flow Using Satellite Spectroradiometer Data. *IEEE Trans. Geosci. Remote Sens.* **2016**.

20. SUN, Y. Automatic ice motion retrieval from ERS-1 SAR images using the optical flow method. *Int. J. Remote Sens.* **1996**, *17*, 2059–2087.

21. Berg, A.; Eriksson, L.E.B. Investigation of a Hybrid Algorithm for Sea Ice Drift Measurements Using Synthetic Aperture Radar Images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5023–5033.

22. Komarov, A.S.; Barber, D.G. Sea Ice Motion Tracking From Sequential Dual-Polarization RADARSAT-2 Images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 121–136.

23. Fletcher, K. *Sentinel-1: ESA's Radar Observatory Mission for GMES Operational Services*; ESA Communications: Noordwijk, The Netherlands, 2012.

24. ESA (European Space Agency). Copernicus Sentinels Scientific Data Hub. 2014. Available online: https://scihub.copernicus.eu/ (accessed on 11 January 2017).

25. Sentinel-1 SAR Technical Guide, 2000–2017. Available online: https://earth.esa.int/web/sentinel/technical-guides/sentinel-1-sar] (accessed on 11 January 2017).

26. Miranda, N.; Meadows, P. Radiometric Calibration of S-1 Level-1 Products Generated by the S-1 IPF. Available online: https://sentinel.esa.int/documents/247904/685163/S1-Radiometric-Calibration-V1.0.pdf (accessed on 11 January 2017).

27. Pulli, K.; Baksheev, A.; Kornyakov, K.; Eruhimov, V. Realtime Computer Vision with OpenCV. *Queue* **2012**, *1*.

28. Bradski, G. Open Computer Vision Library. Available online: http://opencv.org/ (accessed on 11 January 2017).

29. Korosov, A.; Hansen, M.; Dagestad, K.F.; Yamakawa, A.; Riechert, A.V.M. Nansat: A Scientist-Orientated Python Package for Geospatial Data Processing. *J. Open Res. Softw.* **2016**, *4*, doi:10.5334/jors.120.

30. Korosov, A.; Hansen, M.W.; Yamakawa, A.; Dagestad, K.F.; Vines, A.; Riechert, M.; Myasoedov, A.; Morozov, E.; Piotrovskaya, N.; Williams, T. Nansat-0.6.14, 2016. Available online: https://doi.org/10.5281/zenodo.59998 (accessed on 11 January 2017).

31. Granskog, M.; Assmy, P.; Gerland, S.; Spreen, G.; Steen, H.; Smedsrud, L.H. Arctic research on thin ice: Consequences of Arctic sea ice loss. *Eos* **2016**, *97*, 22–26.

32. Itkin, P.; Spreen, G.; Hudson, S.R.; Granskog, M.A.; Gerland, S.; Cheng, B.; Doble, M.; Sennechael, N.; Provost, C.; Haapala, J.; et al. N-ICE2015 Buoy Data, Norwegian Polar Institute: Tromsø Norway: 2015. Available online: https://doi.org/10.21334/npolar.2015.6ed9a8ca (accessed on 10 March 2017).

33. Itkin, P.; Spreen, G.; Cheng, B.; Doble, M.; Girard-Ardhuin, F.; Haapala, J.; Hughes, N.; Kaleschke, L.; Nicolaus, M.; Wilkinson, J. Thin ice and storms: Sea ice deformation from buoy arrays deployed during N-ICE2015. *J. Geophys. Res.* **2017**. in press.

34. Tonboe, R.; Lavelle, J.; Pfeiffer, R.H.; Howe, E. *Product User Manual for OSI SAF Global Sea Ice Concentration, Product OSI-401-b, Version 1.4*; Danish Meteorological Institute: Copenhagen, Denmark: 2016. Available online: http://osisaf.met.no/docs/osisaf_cdop2_ss2_pum_ice-conc_v1p4.pdf (accessed on 10 March 2017)

35. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110.

36. Brunelli, R. *Template Matching Techniques in Computer Vision: Theory and Practice*; John Wiley & Sons: Paris, France, 2009.

37. Jones, E.; Oliphant, T.; Peterson, P.; SciPy: Open source scientific tools for Python, 2001–. Available online: http://www.scipy.org/ (accessed on 11 January 2017).

38. Hansen, M.; Korosov, A.; Olaussen, T.; Vines, A.; Hamre, T. The Nansen-Cloud: a Scientific Platform as a Service for the Nansen Group. In *Proceedings of Living Planet Symposium 2016*; European Space Agency: Paris, France, 2016.

39. Hollands, T.; Linow, S.; Dierking, W. Reliability Measures for Sea Ice Motion Retrieval From Synthetic Aperture Radar Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 67–75.

40. Barton, I.J. Ocean Currents from Successive Satellite Images: The Reciprocal Filtering Technique. *J. Atmos. Ocean. Technol.* **2002**, *19*, 1677–1689.

41. Johansson, A.M.; Berg, A. Agreement and Complementarity of Sea Ice Drift Products. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 369–380.

42. Korosov, A.; Muckenhuber, S. SeaIceDrift-0.3, 2017. Available online: https://doi.org/10.5281/zenodo.376483 (accessed on 10 March 2017).

43. Kalderimis, J.; Meyer, M. Travis Continuous Integration Platform, 2017. Available online: https://travis-ci.org/ (accessed on 11 January 2017).

44. COVERALLS, DELIVER BETTER CODE, 2014. Available online: https://coveralls.io/ (accessed on 11 January 2017).