Review of Muckenhuber et al., doi: 10.5194/tcd-9-6937-2015

Summary

To compute sea ice drift on an operational basis, the authors present an algorithm which is basically an adapted version of ORB. The described algorithm is an open-source code written in Python. The authors apply this algorithm to a set of four image pairs, taken by a synthetic aperture radar on the Sentinel-1A satellite, to obtain sea ice drift vectors. For validation, the authors compare results of their algorithm to results computed with two other algorithms, SIFT and SURF. The authors conclude that their algorithm is superior in quality and quantity of detected drift vectors, and in computation time.

The manuscript addresses an important, topical issue, which is clearly suitable for The Cryosphere. Not many readers of the Cryosphere will be all too familiar with computer vision, and the authors have written the manuscript in a manner that fits The Cryosphere's audience. In my opinion, the title is not as accurate as can be. The title's subject is sea ice drift. However, the manuscript does not contain substantial sea ice drift results, which, I think, is also the authors' intention. The manuscript's subject is the tuned feature tracking algorithm. A rearranged title is more suitable, I think (see my specific comments for a suggestion). The manuscript length is appropriate, maybe even slightly too short. The image quality of the figures is very high. The open source distribution of the code will make it easy for others to replicate and build on the authors' work. I am not fully convinced that the results (or the way these are presented) corroborate the conclusions as good as they possibly could. The authors show mostly averaged results of four test image pairs, apart from Figs. 4 and 5. I am not convinced that the test image pairs are really representatitve of the investigated regions. Rather, the images seem representative of some sea ice conditions. The authors reach substantial conclusions, though these are written in the discussion section. I find the manuscript a little too descriptive, which is reflected by the very short discussion section. To make their case more convincing, I recommend the authors show—and discuss—results from all four test image pairs. If the authors decide to add such information, I believe it will strengthen the authors' rationale, instead of just making the manuscript longer.

I hope my comments and suggestions prove helpful, and look forward to trying out your software.

Kind regards
Tim Dunker

General comments

[1] I think the writing needs an overhaul: there are a number of spelling mistakes and grammatical errors, and physical terms and units are not used correctly. I think the title should be rearranged. Some statements are ambiguous, because you use a mix of active and passive voice. The validation (Sect. 4.5) and discussion (Sect. 5) parts are very short. There is, in my opinion, hardly any discussion. The headline mentions an outlook, but I do not think there is enough there to justify

this headline. Only the very last sentence of that section might be an outlook, but because you use passive voice, readers will not know if you recommend the combination of methods or if this is some general statement. An improved structure of discussion and conclusion (and outlook, if applicable) will not require too much work, but will help the reader to better navigate through the manuscript. I offer a number of specific comments and suggestions you may or may not consider. Hopefully, you will find some of them helpful.

[2] Though Fig. 4(c) supports your claim that the adapted ORB algorithm yields superior performance to SIFT and SURF, it would be instructive to see this comparison for a range of parameter choices (see Table 2). Furthermore, instead of showing an image of results in Figs. 4(b,c), the actual distribution function would reveal many more details of the algorithm's performance, I believe. Especially if there were a potential bias in any of the three algorithms towards large or small vector norms, this bias would stand out clearly in a plot of the distribution function or a quantile-quantile plot, for instance.

[3] Figures 1, 2, and 3 show average results from the four regions mentioned in Table 1. From these figures, it is not possible to judge where the algorithm performed better or worse, as indicated by the number of matches. Instead of having one aggregate figure, would it be better to have one panel per region per figure? In my opinion, this would clarify which parameter should be chosen for which sea ice condition. I expect the sea ice conditions in the Kara Sea and around Franz Josef Land to be quite different compared with the Fram Strait - namely, much less and slower motion in the Kara Sea and maybe more ridges around Franz Josef Land. I think these different conditions make possible a nice comparison of your algorithm with SIFT and SURF.

[4] I am curious to know how the algorithm performs if there are melt ponds on the sea ice surface. Unfortunately, the four test image pairs are all from March and April, when typically the sea ice is largely snow-covered and melt ponds are not as common as in late spring or summer. Therefore, I contest your statement that these image pairs are "representative [of your] region of interest" (p. 6940, l. 19; see also the Abstract, ll. 3 to 5). These images might be representative of the sea ice conditions mentioned in the same sentence (p. 6940, ll. 20 to 21), though.

[5] I do not think I fully understand the quality measure and how you applied it, maybe I have just misunderstood it. In Sect. 4.5 you describe the RMSD: "... combination of the manually produced error and the displacement variation betweeen the manual and calculated vector." In Figure 4(c), you show the RMSD for SIFT and SURF, but you do not mention any manually-drawn vectors that you might have compared to SIFT or SURF results. How did you compute the RMSD for SIFT and SURF? What do you mean by "manually produced error"? You also compute the numbers of correct and wrong matches, but you do not use these to assess quality. Are these figures not be appropriate for qualtiy assessment? Given that the manuscript is about sea ice velocity, should the quality measure not be a velocity either? Have you attempted to quantify the uncertainty of the drift speed?

[6] Before publication, this comment must be addressed: Parts of Section 3.1 are a basically a repetition with negligible alterations of Section 4.1 in Rublee et al. (2011), most notably page 6943, line 19, to page 6944, line 8. One might get the impression that the mathematics is part of your work. However, the mathematics is exactly as published by Rublee et al. (2011). If I am not mistaken, Rublee et al. (2011) already used brute-force matching and the Hamming distance, so Rublee et al.

(2011) should also be cited on page 6942, lines 16 to 17. There are some direct quotes (p. 6942: ll. 7 to 9; ll. 14 to 16; ll. 17 to 18) of Rublee et al. (2011) with minimal changes, but are not marked as such. In the last of these sentences, you have changed the meaning of the sentence, because you have substituted "of" for "unlike". Any direct quote must be enclosed by quotation marks, and modifications must be clearly indicated. Either use quotation marks or reword entirely, but always cite Rublee et al. (2011). Another solution could be to drop most of Section 3.1 as it is now, and instead cite Rublee et al. (2011) for the ORB algorithm and concentrate on the adaptations you have made.

Specific comments

In most instances throughout the manuscript, "computational" must be an adverb, not an adjective. On page 6941 (lines 9, 15, and 18), page 6944 (line 18), page 6947 (lines 4 and 9), and on page 6948 (line 6), the use of an adjective is correct. Please search for "computational" and replace it by "computationally", where appropriate. Please also replace "pixel" with "pixels", where appropriate. In general, I vote for the use of basic SI units, such as $\mathrm{m\,s^{-1}}$ instead of $\mathrm{km\,h^{-1}}$, though I see that the latter appears more practical when two radar images are separated by many hours or days.

Title: You may want to think about rearranging the title. Now, readers might be led to believe the manuscript is mostly about sea ice drift estimated from the feature tracking method. Instead, the manuscript is about the feature tracking method applied to sea ice drift detection. Would not "An open source feature tracking algorithm to detect sea ice drift from Sentinel-1 SAR imagery" be more suitable? I think so.

Page 6938, lines 3 to 5: I do not think it is justified to call the image pairs simply "representative". They might be representative of the sea ice conditions you investigate, but not necessarily representative in general. Please specify what you consider the images to be representative of.

Page 6938, lines 5 to 6: Comma after "introduced". I generally prefer writing "use" instead of "utilise" (also on page 6941, line 7; page 6941, line 22; page 6943, line 11; page 6949, line 13), unless there is an urgent need for exactly that word. Besides, I do not think that this sentence is strictly necessary, and I think you overstate the quality measure's novelty and importance a little. From lines 11 to 13, page 6938, it is not clear that this refers to the quality measure.

Page 6938, line 10: Decimal in "2.7 GHz"

Page 6938, line 12: Comma after "vectors"

Page 6938, line 13: "significantly"

Page 6938, line 14: "four times"

Page 6938, line 18: I propose to change "observe from earth observation data" to "observe from remote sensing data".

Page 6938, lines 23 to 24: I disagree. A little bit further into Sect. 1 (page 6939, lines 9 to 12), you write the opposite: data with a spatial resolution of 400 m have been available for some years. Maybe you should clarify what you mean by "sufficient resolution".

Page 6938, lines 23 to 24: I do not think "convergence field[s]" or "divergence field[s]" is correct, because divergence is a scalar. The divergence operator is applied to a field, for example to a velocity field. I suggest to simply write "divergence" and/or "convergence".

Page 6938, line 25: "The regions of interest are".

Page 6939, line 1: "These seas" or "These areas"?

Page 6939, line 8: "which acquires data"

Page 6939, lines 18 to 19: One day is not a frequency.

Page 6940, line 1: "... has been considered" by you or by Rublee et al. (2011)? Please clarify.

Page 6940, line 18: Either 400 km × 400 km or (400 × 400) km^2.

Page 6940, line 13: One day is not a frequency.

Page 6940, line 19: "representative of"

Page 6940, line 23: You provide all relevant links in Appendix A already. I suggest to either move the link to a footnote here or write "(see Appendix A; Korosov et al., 2015)". With either option, you get rid of the two subsequent parentheses.

Page 6940, line 24: "scientist-friendly"

Page 6940, line 26: "a simple"

Page 6941, line 7: For consistency: "centre" and "centred"

Page 6941, line 10: "at high latitudes"

Page 6941, lines 13 to 14: I think you overstate a little here regarding the quality measure.

Page 6942, lines 1 to 2: Commas after "measure" and "cell".

Page 6942, line 11: "applies a Harris"

Page 6942, lines 11 to 12: What does "them" refer to—pyramid levels or multi-scale features? Your

description, specifically the order of steps, seems to be different from the one given by Rublee et al. (2011). Please clarify.

Page 6942, line 19: "applied to"

Page 6942, line 20: The variable σ^0 does not show up in Eq. (1). I suppose σ in Eq. (1) should read σ^0. I am not exactly sure if I understand the formulation "i range [0,255]" correctly. I think you mean that the intensity $i \in \mathbb{R}$ is defined on the interval [0,255]. If so, you could write something like $0 \leq i \leq 255$ for $i \in \mathbb{R}$.

Page 6942, lines 25 to 26: "for both channels."

Page 6943, lines 3 to 4: "with a perimeter"

Page 6943, line 4: "nine contiguous pixels"

Page 6943, line 5: "recognised"

Page 6943, line 6: "as a keypoint"

Page 6943, line 7: What do you mean by "retained keypoints"?

Page 6943, line 8: "keypoint is assigned"

Page 6943, line 9: Something went wrong with the citation here. I suppose you want a semicolon after "Harris corner measure".

Page 6943, line 16: "intensity-weighted"

Page 6944, lines 8 to 9: Did you do the same as Rublee et al. (2011)? Because you use passive voice here, this statements does not say anything about what you did or did not do. Please clarify.

Page 6944, line 19: "40 m to 80 m"

Page 6945, line 2: No comma after parenthesis

Page 6945, line 8: No comma after "rejected"

Page 6945, line 9: "Unreasonably"; "$10\,\mathrm{km} + 1\,\mathrm{km\,h^{-1}}$" is not a velocity.

Page 6945, line 15: "equal"

Page 6946, line 1: This sentence confused me. You do not say which parameters you are going to test, so it is not clear to me what the "remaining parameters" are. At first, I understood this statement such

that you did not test pyramid levels, scale factor, HH and HV limits, and ratio test, but set these to the mentioned values instead. From Table 2 and Figures 2 and 3, however, it becomes clear that you tested these settings over some range. Please clarify which parameters you tested and over what range (refer to Table 2, if appropriate) and which parameters you did not test.

Page 6946, line 4: A "low speed filter with 2.5 km" does not say much about the filter's nature, nor is 2.5 km a velocity. From what you write on the same page in lines 6 to 7, I first thought you used a high-pass speed filter. However, the filter rejects unreasonably high velocities, so I think you constructed a bandpass filter with a passband between your chosen cut-off speeds. Is this correct?

Page 6946, lines 6 to 7: 2.5 km is not a velocity.

Page 6946, line 13: "Taking this"

Page 6946, line 15: I assume a pixel covers an area. Do you mean 2.72 km^2 or (2.72×2.72) km^2?

Page 6946, line 16: Comma before and after "respectively"

Page 6946, lines 21 to 22: "... image pairs, we suggest to set the maximum backscatter $\sigma_{\max}$ to 0.08 and 0.013 for HH and HV."

Page 6946, lines 23 to 24: This statement ("decreasing towards higher values") is a bit confusing at first. To avoid misunderstandings, you could write something like "... , because the number of matches decreases for increasing values of $\sigma_{\min}$ (not shown)."

Page 6947, line 1: "1 to 16". Figure 3 only shows data for a scale factor of 1.2, but not for the other scale factors mentioned here.

Page 6947, line 5: "seven pyramid"

Page 6947, line 9: "four representative"

Page 6947, line 11: "Figures"

Page 6947, line 12: "significantly"

Page 6947, line 24: "longitude"

Page 6948, line 8: "sufficiently computationally efficient"; Still, it sounds a bit awkward, but I have not come up with a better formulation. What do you consider sufficiently efficient?

Page 6948, lines 9 to 10: "The processing times shown ..."

Page 6948, line 12: "36 % and 67 %"

Page 6948, line 13: Comma after "SURF"

Page 6948, lines 21 to 24: This sentence is not related to validation, so it should be moved to another section.

Page 6949, line 4: "This proves"; I think you can be more confident about your conclusion here.

Page 6949, line 6: "Proofing" is a noun only. What about "showing" or "demonstrating" or similar?

Page 6949, line 9: "independent of"

Page 6949, line 13: What does "This" refer to? I assume you want it to refer to the "not evenly distributed" sea ice drift fields, while "This" actually refers to "performance". On a different note: Have you really used the uneven distribution of vectors in a quality measure in your manuscript? Maybe I misunderstood your quality measure. Please clarify.

Page 6949, line 21: "at no cost"

Page 6949, line 24: Comma after "algorithm"

Page 6949, line 25: Comma after "data"

Figures

Figures 1, 2, and 3: "the four test..."

Figures 1 and 3: The captions do not mention the combination HH + HV.

Figure 2, caption: Please clarify what you mean by "chosen maximum". I think it should read "chosen parameter", because the value indicated by the grey line does not seem like the maximum to me. Mathematically, there seems to be only one maximum, therefore "chosen maximum" confuses me.

Figure 4: The use of colour for the ocean and for Greenland and Svalbard is unnecessary. This makes the colour scale hard to interpret, because the colours on the map become a mix between the background colour and the colourbar. I think the background (that is, the ocean) should just be white. In the caption: "grid cell", and comma after "(c)". The figure size should be maximized, or the figure should be divided. Panel (a) is hard to interpret in print—even more so when I printed it in black/white—because of its many features. In the PDF version, magnifying is unproblematic because you use high-quality vector graphics, which is great.

Figure 5: For the figure to be useful, there must be a velocity scale. I assume the data in Fig. 5 is the same as in Fig. 4(a). However, some vectors present in Fig. 4(a) are not present in Fig. 5. Is

that correct? If so, how and why have these been removed? See my comments on colour regarding Fig. 4. The use of red and green together is not a good choice and should be avoided.

References[1]

- Bay et al. (2006): This is actually a book chapter. See here for the full bibliographic information: DOI: 10.1007/11744023_32.
- Korosov et al. (2015): I think the conference was called "World Ocean Science Congress". I could not find this reference or any proceedings from that conference. Can you make the presentation/poster/article available in the appropriate Nansen Center Github projects or on the Nansen Center's website?
- Kwok (1998): The original title is all capitalized. Editors of the book are C. Tsatsoulis and R. Kwok.
- Kwok (2010): "sea-ice"
- Kwok and Sulsky (2010): The abbreviated (and also full) journal title is "Oceanography".
- Low (2004): The abbreviated journal title is "Int. J. Comput. Vision".
- Rosin (1999): The abbreviated journal title is "Comput. Vis. Image Und."
- Rosten and Drummond (2006): Similar to comment on Bay et al. (2006). See DOI: 10.1007/11744023_34 for details.
- Rublee et al. (2011): I do not know what "Willow Garage" is, but the paper was published in the journal IEEE I. Conf. Comp. Vis., see here: DOI: 10.1109/ICCV.2011.6126544.

Tables

Table 2: I do not understand the speed filter value. The setting you recommend is not a velocity. Besides, you apply both a low-speed and high-speed filter. What do you mean by "tested range (resolution)"? I assume you mean "tested range or resolution". If this is what you mean, please do not use parentheses. The recommended settings for HV and HH channel brightness are not maxima in Fig. 2, even if the figure caption says so (see my comment below on the caption of Figure 2).

Table 3: "Create two Nansat..."; "from two Nansat objects"

[1]For the abbreviated journal titles, see the Web of Science, e.g. mirrored by the Caltech Library.