Introduction

Thank you for your detailed reply to my comments and the significant improvements in your paper. It is nice to see this tempting idea evolving and over all the paper is now worth being publishing after some more revisions to be done. The only fundamental problem I see is the parallel publication on the same topic from two other authors (actually two research group leaders!) from the same institute but this should not be a problem of an individual PhD student especially since the this respective discussion paper by Muckenhuber and Sandven has been published earlier and the associated code has been made publically available. At the end, it is a decision only the editorial board of the Cryosphere can make.

This having said, I'm afraid that I have still some comments. During the last revisions I asked to be more careful regarding some claims by asking for "softer" phrases and adding a more detailed analysis. The three main claims of the paper are:

1. The authors use feature tracking for a first guess to be computational more efficient than pure pattern matching
2. The first guess is based on independent vectors – therefore the authors are more capable to resolve deformation zones
3. HV polarization is better for feature tracking and therefore potentially as well for pattern matching
4. The algorithm can handle rotation

It is obvious that the setup of the algorithm allows a better handling of rotation but it is less obvious and not shown in the paper, that the algorithm is actually faster than pure pattern matching algorithms, that they can handle deformations and shear regions in a better way or that the use of HV data is more advantageous for pattern matching as well (see as well my thought in the last review which the authors included in the paper as well). I won't say that these claims are not potentially true but the authors neither show nor prove it. It's merely a question of the wording and no extra science that needs to be done but the first three claims are not as obvious as the authors suggest. (e. g. it might be better to say that the first guess by feature tracking potentially allows a computational efficient computation than claiming that it is more efficient that pure pattern matching. If you want to claim it: prove it!). It might not sound that sexy but it would be more precise. The same holds true for claim 2 and 3.

One last thing before I deal with the comments in detail: Personally I think that the authors stress the computational efficiency of their algorithm a bit too much and avoid every operation that could hamper this goal. That's okay but it should be discussed in the text (e.g. "a better way to determine d would be to take into account the variability of the vectors as well we avoided it for the sake of computational efficiency") In this way the readers would be aware of these limitations. Just scan through your last answers to my comments for the term computational efficiency and add some explanation at the respective position in the

manuscript. Besides that the DTU shows that operational products a possible with pure pattern matching as well.

Comments on Response to Referee #1

The authors answered most of my questions and solved many of the problems I mentioned. Thanks for that. In the following I will focus on the points only, where I do not agree with the answers which thereby might sound more negative than actually meant.

Page 1 Line 14ff the authors are of course right that the noise level depends on the incident angle. However, there is a pattern of the noise level (e.g. https://doi.org/10.1016/j.procs.2016.09.247 and ESA mentions -25 dB and -22 dB). But if the authors basically use a fixed threshold for the scaling of all scenes, I would like to ask for a reference that suggests doing so or containing information about the variation of noise level over the swath width. It might hold true for some scenes but definitely not for all and would lead to a scaling over a value range containing only e.g. 50 % meaning full values or throw away 50 % of values containing information in the case the authors showed and used to make the point that there approach is correct . But if the distribution has so many values in in the range below the noise level it is at least worth checking if the calibration routines of Nansat work properly (the scaling itself provides somehow has to provide a reasonable image otherwise, there would not be any data.

Page 2 Line 40 "… an even better parameter set, because a higher number Is considered …" It is not the same, it is a completely different approach and therefore not better but just different, providing different information but I'm okay with it.

Page 2 Line 50 "… more feature tracking vectors are found on image pair with a smaller time span" – given a sufficient number of linear features visible in HV, only and not in the less textures areas where the feature tracking algorithm tends to fail I would like to add but as I said, I guess, that's okay.

Page 3 Line 80: I understand the point of the authors and agree with them! Having said this, I'm not happy with it and don't think that things like this should happen since it is bad science and from an outside perspective creates the impression that NERSC has more expertise in the field of sea ice drift than there is actually available plus increases the noise level for publications in this field in general, which makes it more difficult to keep an overview of relevant publications. However, as I stated in my introduction it is up to the editorial board of the Cryosphere to decide.

Page 4 Line 93: The mentioned independence is not only reduced by overlapping / close features but as well by very characteristic and bright features that might be represented in all resolution levels of the ORB algorithm but I guess that's okay

Page 4 Line 105: The feature tracking algorithm chooses the best detectable features automatically but this means that there can also be vast areas with no feature tracking

vectors at all if there are enough well defined ridges is one corner of the scene. For the first guess the motion field is now interpolated over these vast “vector”-less areas as well. I have my doubts if this is better or more independent than a fine tuned resolution pyramid in the case of pattern matching. I suggest skipping this comparative statement

Page 5 Line 134 f It is not necessarily true that a scaling that favours feature tracking favours pattern matching as well. 1. The NCC needs no scaling at all since it is normalized (the N of NCC) and 2. Imagine an extreme scaling that would separate ridges in white from the surrounding background in black. For feature tracking, this would be perfect, since it would highlight the linear feature. For pattern matching it would be a problem, especially in the regions outside the ridged areas since it is based on areal pattern variation.

Page 8 Line 242 Thanks for the explanation. Just mention it in the manuscript

Page 9 Line 291 “We believe … “ I suggest to show it or be more careful with such a statement. The ice drift of the e.g. for example is pure pattern matching without any fancy resolution pyramid and runs pretty fast to provide the ice drift within the framework of CMEMS.

Page 12 Line 409 / Page 13 Line 437 the authors promise to bring feature tracking performance of HH closer to HV by an increased coverage of S1 images and a better pre-processing. I would claim that HH and HV show different characteristics and that’s it. You won’t change the stronger dependency of linear features in HH polarised images from the incident angle by increasing the temporal resolution but may see other features. Having said this, a better pre-processing and a higher temporal resolution never hurts and will improve the results for sure.

The Manuscript itself

Most of the remaining points I already handled in my answers to your comments, meaning there is not much left.

Still open points from previous reviews

Page 4 Line 119 f – see as well my comment above: the independence is difficult to assess and its effect on resolving deformation zones is not shown in this paper – may be more careful

Page 7 Line 210 – 215 See my comment to your answers

Page 16 Line 393 I think we had this discussion already in the first draft. I think it is difficult to use a buoy dataset for validation if you used the same dataset beforehand to optimise the setup of the algorithm (Page 16 Line 370-377).

It does not result in a real assessment of the performance of the algorithm but in the potentially performance in the case that the algorithm is tuned in the right way to the respective ice conditions.

Technical comments

Page 3 Line 62: “On the one hand… on the other hand” – there is “the one hand” missing

Page 4 Line 105 since you mention in line 104 that the area has a revisit time of less than one day it is obvious that your area of interest is monitored on a daily basis

Page 13 Line 324 “accurate drift” – it not necessary a more accurate drift but merely a more localised/ locally tuned/ locally adapted one. The term “accurate” sounds good but has a meaning which has nothing to do with the case at hand – nevertheless we might hope that a locally adapted drift vector represents the drift at a certain position more accurately than the first guess but that is not necessarily the case.

Page 15 Lin 363 “since the uncertainty increases with distance d (Figure 7)” – may be rephrase the sentence – the authors are correct but it suggests that Figure 7 shows the increase of “uncertainty” depending on the distance. Having said this: “uncertainty” is a statistical term – may be choose something less well defined.

Page 15 Line 368 / Page 17 Line 405 – it is more or less the same sentence, may be skip one

Page 21 Line 482 “accuracy of the introduced algorithm” I suggest to add “for given image pairs, given ice conditions, given region and given time” since the results are not transferable to other ice conditions or regions…

I hope my comments help you improving your paper and I’m sorry for the spot of work I left you with. It is an interesting topic and I’m really looking forward to the final paper!