Response to referee comments

Precise DEM extraction from oblique imagery of Svalbard in 1936

Girod, L., Nielsen, N.I., Coudrette, F., Nuth, C., Kääb, A.
Geoscientific Instrumentation, Methods and Data Systems Discuss., doi:10.5194/tc-2016-228, 2016

We would first like to thank the three reviewers for their constructive contribution and input that helped clarify and focus the paper. Detailed responses are provided below (in bold), together with a mark-up manuscript version where the changes made are highlighted.

Simon Buckley (Referee Comment #1)

This paper is a valuable contribution that examines the "classical" photogrammetric orientation problem associated with non-standard image blocks. In this case the high oblique configuration is one where analogue, analytical, and early digital photogrammetric workstations would have struggled to achieve orientation solutions. It is therefore interesting to see that modern techniques are able to resolve the non-optimal overlap and perspectives to achieve higher quality (and, importantly, extremely usable) results.

The paper itself is largely straightforward to read and outlines the method carried out in an easy to follow manner. I would have liked to see some more reference to earlier photogrammetric literature, as much work has been done on archive photography to update results according to (then) state-of-the-art techniques, even if outside of cryosphere research. E.g. a quick search of The Photogrammetric Record's archive ("archive photography") reveals relevant articles (Chandler and Clark, 1992; Fox and Cziferszky 2008; Perez et al., 2014). Some descriptions of processing method components could be simplified or updated based on photogrammetric textbook references, e.g. the standard interior orientation step.

Thank you for your suggestion of literature, the text has been edited with more references to earlier work.

Title: a suggestion for improvement is "Precise DEM extraction from Svalbard using 1936 high oblique imagery"

Title changed to your suggestion.

Abstract could do with a final sentence with "conclusion" statement

The abstract now includes a short summary of the conclusions.

Line 20: photogrammetric scanners used in the 1990s can maintain good geometric geometry? It's not clear what you mean by "normalized" here.

The scans were not performed in a consistent way, and the fiducial marks are not consistently found at the same pixel coordinates. The interior orientation needs to be standardized before proceeding to the calibration and exterior orientation. Text has been clarified, and the word "normalized" replaced by "standardized" throughout.

*Stereo-overlap is low (typically 60% along-track and 20% cross-track) compared to modern standards, *which are: : :?**

For SfM surveys, I would say 80/60, or at least around these values. Depending on the terrain, higher or lower overlaps can be chosen, but never as low as 60/20. The 80/60 value is added to the text.

In addition, I have marked areas in the manuscript where some minor rephrasing or editing is required (attached PDF).

Thank you for these comments, they have been taken into account in the revised manuscript, and a final language editing is provided by the journal.

Please also note the supplement to this comment: https://www.geosci-instrum-method-data-syst-discuss.net/gi-2018-25/gi-2018-25-RC1-supplement.pdf

Anonymous Reviewer (Referee Comments #2)

Structure from Motion photogrammetry offers today the possibility to process or reprocess old images, obtaining new information about the past. This paper analyzes the challenge of using 1936 high oblique imagery to generate a DEM of Svalbard. These images have peculiarities that can make their processing more problematic compared to nadir images, as correctly illustrated by the authors in the introduction. In this light the work is particularly interesting. The text is easy to read and to follow, and the research is presented in an appropriate way.

I'm agree with Simon Buckley comment: the references must be expanded, adding more articles about "archive photogrammetry", even if in other fields of application. Moreover the abstract appears cut in the final part: a sentence about the "conclusion" of the work must be added.

More references are added. The abstract now includes a short summary of the conclusions.

Some questions and observations:
- Which software was used for processing? A commercial one or a software developed "in house" by the authors? It's not clear.

The free open source photogrammetric suite MicMac (see http://micmac.ensg.eu) was used. This is now referenced in the first paragraph of section 3.

- What do the authors mean using the term "normalization"? In photogrammetry the collimation of fiducial marks is the phase of interior orientation; in this case study their positions are available? If not, how was the thing solved? This point must be better explained in paragraph 3.1.

"Normalization" is replaced with "interior orientation" as this is indeed the appropriate term. In the case study, a series of measurement on a sample of the original films was averaged into approximate fiducial mark coordinates. In the generic case, if no data is available, a theoretical approach has to be made, considering the film size and the apparent position of the points. The text is edited to explain this further.

- Paragraph 3.3: I understood that the camera calibration was performed processing 5 images and then extending the solved parameters to the other images...is it correct? Maybe this point must be better explained and some observations about the calculated parameters added.

This is correct. The text has been edited to add clarification.

- Figure 2: it is "mosaicking" I think...

Yes, corrected.

- I didn't understand: in order to obtain the point cloud triplets of images have been processed; in the overlapping areas between a triplet and the next one the extracted points are consistent?

In our approach, the points are rasterized into DEMs before the merging of the different triplets, and the mosaicking step selects which point to record for the final product. In some rare cases, a mosaicking line can be observed.
Blending the point clouds before rasterization is possible, but the file size of a merge point cloud makes processing even more challenging. We did however test this in a subset of the data, and in the overlapping areas (that are quite small in our case study), the points from both triplets were put to contribution for the DEM value. The variation of the point values from different triplet is well within the noise level so their cross contribution does not have a measurable effect. As explained in the text, for rasterization*, *"each cell of the rasters is given the mean value of the points in a radius around the cell center"*, so the values would be mixed then.

- Is it possible to add some numeric values indicating the quality achieved in the processing? For example the reprojection error or the ground control points RMSE.

For the Prins Karl Forland dataset, the average residual of the tie points was 0.803 pixels, the GCP reprojection error was of 4.108 pixels with an RMSE of 18.15m. These values have been added in the text. The relatively high RMSE value is due to the relatively soft DEM used as an input for GCP identification. As written in the text, the quality of the data is better estimated by pulling statistics from various Difference of DEMs (DoD) between our data and ground truth DEMs.

Anonymous Reviewer (Referee Comment #3)

The paper presents an application of SfM to a dataset of historical oblique images with the aim of deriving a DEM of the area with the best achievable accuracy. The application is interesting and the case study is meaningful, however the methodological part of the paper requires some major clarifications.

Specific comments
Line 32: please, clarify how the image were resampled.

The text was edited to provide more information on the normalization of the interior orientation.

Paragraph 3.2: please, state how many ties did you look for and which statistical criteria did you adopt to accept or reject points. Please, state also which software did you use.

(idem RC #2) : The free open source photogrammetric suite MicMac (see http://micmac.ensg.eu) was used. This is now referenced in the first paragraph of section 3.
SIFT++ and ANN found between 1000 and 10000 tie points per image (a note is added to the text), depending on content (more land and more contrasted texture yields more points than water and contrast free surface) and overlap (the edge images, having only overlap on one side, yield much less points than images in the center).

Par 3.4: it is not clear what is the "modern source" used and its accuracy.

At the time of processing the orientation, the 1990 DEM from the Norwegian Polar Institute was the best data available, and was therefore used. The GSD of this product is 20m and the standard deviation with the 2008 DEM is below 2.5m. This information is added to the text.
From a methodological point of view, any "good" data source is a candidate for GCP identification.

Par. 3.5, line 6: the choice of working with three images only is quite unusual for SfM and should be discussed better.

The SfM part of the processing chain happens before the division of the data into triplets and therefore uses all the images of a line of flight (as SfM is merely the computerized and automatized computation of the relative orientation). Three images are enough for correlation, and, while more images (at least 5) would be preferable, the approximate 60% overlap means that no such data exist.

Par 3.5, l 14-15: Please, reformulate this part. It is not clear in this context what you considered "area of interest", what you considered background and foreground. It is an important aspect for process automation.

The area of interest is literally the area the user is interested in. In our case, it is the strip of land closest to the camera positions, as the background is of far too low resolution. A comment is added to the text.
The definition of background is indeed a bit hazy, the hardest to automatize (see Fig 4.f, where the statistics for foreground and background are overlapping), and is to be defined by the user.

Par 4.1, l 14-17: what is the nominal accuracy of these products?

According to the metadata (https://doi.org/10.21334/npolar.2014.dce53a47), 2-5m for the NPI DEMs. This value is added to the text.

Par 4.1, l 18: please, be more precise: what did you mean with low-slope?

Low slope is <20°. This is added to the text.

Par 4.1, l 1-2 (pag10): did you compare with all the products mentioned at the beginning of the paragraph? Are these values an average among the three? How did you make the comparison between DEMs at different resolutions?

The text has been edited, as the values correspond to the comparison with the 2008 NPI DEM (by far the most complete, high resolution, and accurate ground truth available).

Par 4.1, l 4-5: this comparison is not meaningful since the investigated areas are very different in extent; can you evaluate the elevation error of your product in the same small area investigated by Mertes et al.? Furthermore, a few details about their methodology would ease the comparison.

Mertes et al. did not disclose the exact area considered in their statistics, but the selection of their data seems to be similar to ours. It is our opinion that the values are comparable, and we added a note on this to the text.

Par 5, l11: it appears to me that the amount of manual editing was relatively large and it was crucial for the accuracy.

The amount of manual editing (clean-up of the point clouds) is indeed time consuming, especially because of the loading time involved when working with point clouds consisting of tens of millions of points. However, that process is really just a clean-up of the data so that the areas of interest are not polluted by large chunks of erroneous data. The actual conserved data points are not edited. Clarification has been added to the text.