



HAL
open science

3D MODELING OF EARTH'S SURFACE: STUDY OF THE ANTARCTICA

Philippe Chiberre, Enric Meinhardt-Llopis, Carlo de Franchis, Gabriele
Facciolo

► **To cite this version:**

Philippe Chiberre, Enric Meinhardt-Llopis, Carlo de Franchis, Gabriele Facciolo. 3D MODELING OF EARTH'S SURFACE: STUDY OF THE ANTARCTICA. IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Jul 2019, Yokohama, Japan. pp.1088-1091, 10.1109/IGARSS.2019.8900179 . hal-02493725

HAL Id: hal-02493725

<https://hal.science/hal-02493725>

Submitted on 28 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

3D MODELING OF EARTH’S SURFACE: STUDY OF THE ANTARCTICA

Philippe Chiberre*

Enric Meinhardt-Llopis*

Carlo de Franchis*,†

Gabriele Facciolo*

*CMLA, ENS Cachan, CNRS, Université Paris-Saclay, France

†Kayrros, France

ABSTRACT

The evolution of the antarctic ice cap is a subject of the utmost importance for the climate science. For this reason exploiting the historic archive of SPOT 5 HRS binocular stereo imagery over the Antarctica has gained interest. However, estimating surface models of Antarctica from optical satellite imagery is a challenging task. Different factors contribute to its difficulty: the reduced contrast of the snow, the abrupt changes in elevation, and the persistent cloud cover that is often indistinguishable from the snowy ground, or is translucent, or projects shadows on the ground; all of them hinder stereo matching. For these reasons the direct application of existing satellite stereo pipelines on these images often yields unsatisfactory results. In this study we explore strategies to address these problems and improve 3D modeling on these regions. We adapt the S2P [1] pipeline and incorporate a new multiscale strategy that allows to deal with incorrect geometry estimation due to clouds or lack of texture. In addition, integrating existing very low resolution DSM (1 km per pixel) of the Antarctica allows to further filter the model to produce a cleaner 3D model.

1. INTRODUCTION

Earth observation satellites allow to map the poles providing scientists with a more precise picture of how glaciers are changing and to help them understand the impacts of climate change on our planet [2, 3]. Creating height maps from optical stereo images is key to follow the ice evolution over time, and make the best use of historical data.

In the context of the *SPIRIT* project [4] (*SPOT 5 stereoscopic survey of Polar Ice: Reference Images and Topographies*) during two years, from 2007 to 2009, a dataset of SPOT 5 HRS images was compiled over polar ice in order to produce DTMs [5] and ortho-images.¹

However, 3D reconstruction on these regions (see Figure 1) poses a number of difficulties for binocular stereo algorithms:

- There is no high resolution reference DSM model, such as SRTM, in this region. This impacts the performance of most stereo pipelines as the initialization allows to

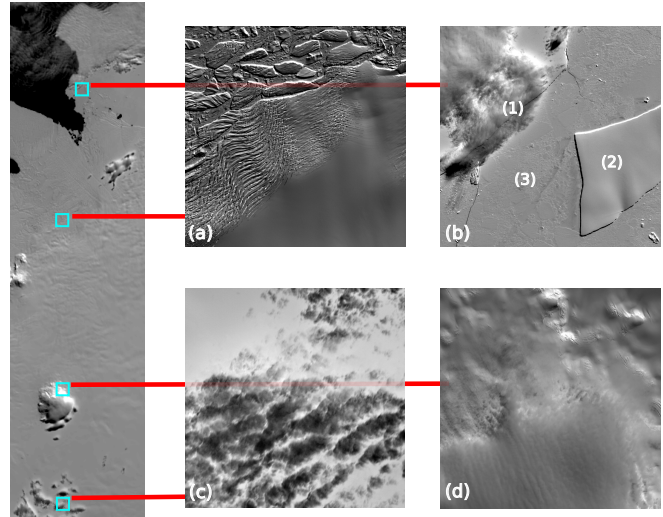


Fig. 1: SPOT 5 HRS image of the Getz ice shelf. A preview of the acquisition is shown on the left, on the right are shown 2000×2000 details from the large image. For all images the contrast has been adjusted to facilitate the visualization of details. (a) shows an ice sheet breaking off. (b) illustrates the specific difficulties to the polar region: transparent clouds that will disturb rectification (1) and large textureless regions (2) as opposed to textured regions which we can easily correlate (3). (c) shows cloud shadows projected on the ground. (d) shows a mountain peak covered from below by a cloud.

restrict the stereo matching range. Nevertheless, in the context of the *SPIRIT* project a reference topography with 1km resolution was produced, and it will be used in the present study to post-process our results.

- Images contain thin clouds with visible texture covering large parts of the scenes. These clouds often lead to stereo mismatches, introducing a bias difficult to correct in the elevation maps.
- Some clouds and icebergs move between the acquisitions, introducing a movement that is inconsistent with the parallax motion. This leads, in these areas, to estimations of the binocular stereo geometry (i.e. fundamental matrix) that are inconsistent with the true cameras configuration.

¹SPIRIT data from: <http://polardali.spotimage.fr:8092/wstools/IPY/>

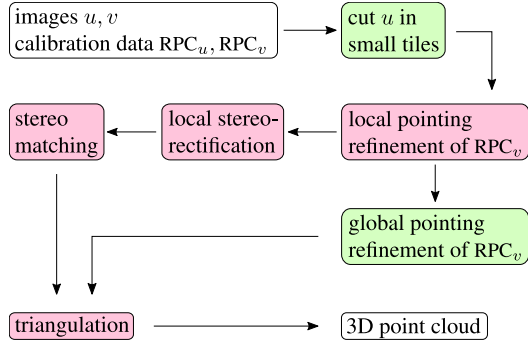


Fig. 2: S2P stereo pipeline overview. The input is a pair of images with their respective camera models (RPC), and the output is a digital elevation model given as a georeferenced 3D point cloud. Green blocks are applied to the whole images, while pink blocks are applied on small independent tiles. They are processed in parallel.

- Images contain large areas of snow, lacking of texture, In these areas no stereo matching algorithm can give meaningful results.

In this study we use the *S2P satellite stereo pipeline* [1, 6], an open-source modular pipeline that allows to produce digital elevation models from satellite images. This pipeline uses off-the-shelf computer vision tools while abstracting from the complexity associated to satellite imaging. To do so, it proceeds as shown in Figure 2. Images are first cut in small tiles for which the pushbroom geometry is accurately approximated by a frame camera model. These tiles are then processed with standard stereo image rectification and stereo matching tools. As the preferred stereo matching algorithm [7] for S2P is based on the census transform, the results are robust to contrast changes between the images. The specifics of satellite imaging such as pointing accuracy refinement, estimation of the initial elevation from SRTM data, and geodetic coordinate systems are handled transparently by S2P. The final DSM is finally triangulated from the correspondences computed on all the tiles.

The S2P stereo pipeline overcomes some of the above mentioned difficulties, but struggles with large regions (spanning several tiles) containing moving clouds and/or textureless areas. We will therefore build on the existing S2P to improve the DSM reconstruction on Antarctica, proposing a novel multiscale approach and post-processing for addressing these issues.

2. LIMITATIONS OF LOCAL STEREO RECTIFICATION

Regions with low texture and clouds such as the Antarctica create problems for tile processing used by S2P, concretely during the tile rectification step.

First, let us review the S2P tile processing from Figure 2. Every pair of tiles (composed of a tile from the first image and the corresponding tile from the second image) is independently processed as follows: (i) match SIFT keypoints; (ii) locally correct the camera models using the matches; (iii) use the corrected camera models to compute the *fundamental matrix* F and stereo rectify the tile; (iv) then apply the stereo matching algorithm.

Stereo rectification permits to restrict the search for corresponding image points from the entire image plane to a single line. For any point \mathbf{x} in a view of the pair, the corresponding point \mathbf{x}' in the other view, if it exists, lies on the *epipolar line* of \mathbf{x} denoted by $\text{epi}^{\mathbf{x}}$. Conversely \mathbf{x} lies on $\text{epi}^{\mathbf{x}'}$. The rectification aims to resample the images in such a way that corresponding points are located on the same row, thus simplifying the matching task and permitting to use all classic stereo matching algorithms.

Although stereo rectification shouldn't be possible for pushbroom cameras [8], de Franchis et al. [1] have shown that satellite images can be precisely rectified on small tiles (of about 1000×1000 pixels) by locally approximating the pushbroom cameras with pinhole cameras.

The epipolar geometry of an image pair is then captured by the fundamental matrix $F \in \mathbb{R}^{3 \times 3}$. F is a rank-2 matrix such that

$$\mathbf{x}^T F \mathbf{x}' = 0, \quad (1)$$

where $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^3$ are matching points represented in homogeneous coordinates. Then, for a point \mathbf{x} the epipolar constraint on the points of the second image is given by $\text{epi}^{\mathbf{x}} \mathbf{x}' = 0$, where $\text{epi}^{\mathbf{x}} = \mathbf{x}^T F$ (similarly $\text{epi}^{\mathbf{x}'} = F \mathbf{x}'$). This means that for each point observed in one image the corresponding point on the other image must lie on its epipolar line.

The problems on the Antarctic regions come from the fact that often there are entire tiles not textured enough for allowing a correct local rectification. A different problem occurs when the area is perturbed by moving clouds, where the matches correspond to a displacement that is inconsistent with the parallax motion. In both cases the estimated fundamental matrices are inconsistent with the scene geometry and ultimately result in incorrect altitude estimations. This is the biggest drawback of this tiling strategy.

3. MULTISCALE RECTIFICATION SOLUTION

The proposed solution is to work in a multiscale fashion. We use lower resolution images to progressively estimate the most consistent fundamental matrices for each tile (step (iii) of the pipeline). Once the rectification at the last scale is done we proceed as usual with the S2P chain.

We first create multiple images (and adapt the corresponding RPCs) from the original image. We iteratively downsample the image by a factor 2 until scale 64. The resulting low resolution images will certainly contain points of high con-

Algorithm 1: DSM post processing

Input : DSM_{in}: DSM to be processed
Input : DSM_{ref}: coarse reference DSM
tmpDSM \leftarrow NaN // invalid height
tmpDSM \leftarrow DSM_{in} **where**
|DSM_{in} - DSM_{ref}| < 300
DSM_c \leftarrow remove_small_CC(tmpDSM, 40)
return: DSM_c

trasts on the ground such as mountains, cliffs or snow-water interfaces, as a 1000×1000 pixels tile at scale 64 represents 320 km^2 . We therefore are able to create a first approximation of the fundamental matrix using matches that are with high probability anchored on the ground.

For each new tile at each new scale we estimate a new fundamental matrix F from the SIFT matches [9]. Let $(\mathbf{x}_i, \mathbf{x}'_i)$ denote the SIFT matches in homogeneous coordinates computed at the current scale (we have n matches). At this point we decide whether to keep the estimated F or to use the one estimated at the previous scale F' . For that, we evaluate the matches by computing the point-line distance $d()$ to the epipolar lines (reprojection errors) associated to each fundamental matrix. We compute the following ratio:

$$r = \frac{\sum_{i=1}^n d(\mathbf{x}_i^T Z F' Z, \mathbf{x}'_i) + d(Z F' Z \mathbf{x}'_i, \mathbf{x}_i)}{\sum_{i=1}^n d(\mathbf{x}_i^T F, \mathbf{x}'_i) + d(F \mathbf{x}'_i, \mathbf{x}_i)}, \quad (2)$$

$$\text{where } Z = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3)$$

Since F' was computed at half resolution we needed to scale the image coordinates with Z . If the ratio r is high, the current matches do not fit the initial model F' , thus F is probably incorrect because it was estimated from erroneous SIFT matches. Therefore, we replace it with F' and use it for the subsequent scales. On the other hand if the ratio is close to 1 then the reprojection errors for both matrices are comparable thus we keep the finer matrix F and use it for the subsequent scales. For the experiments below we set the threshold on the ratio to 40.

DSM post-process. We filter the DSM resulting from the previous process using the low resolution reference DSM (1 km per pixel). This allows us to remove mismatches due to the lack of texture and incorrectly estimated altitudes resulting from matching the cloud cover. Algorithm 1 summarizes the post processing steps.

By assuming a maximum slope of the terrain of 30%, we derive a 300 m threshold as the maximum terrain difference allowed for a 1 km pixel of the reference DSM. That is, if the absolute difference between the estimated DSM and the reference DSM is greater than 300 m we consider the point to

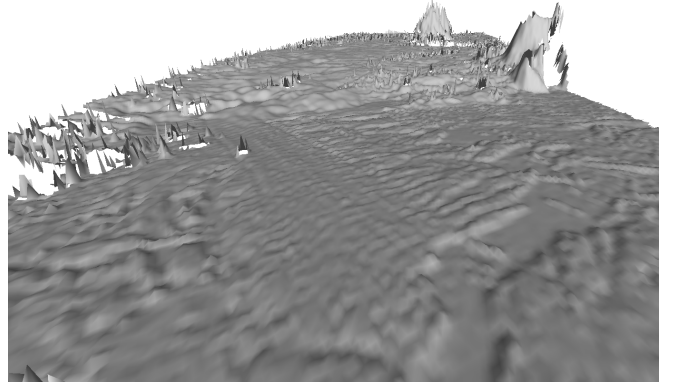


Fig. 3: 3D visualization of the estimated surface with the proposed multiscale method.

be wrong and set it to *no data*. This allows to remove most of the mismatches resulting from matching the cloud cover which is usually more than 500 m above the ground.

A second post-processing step allows to deal with most of the mismatches coming from textureless regions that were not filtered with the first step. Since elevation in mismatched regions is usually very irregular, following [6] we implement a speckle filtering (denoted `remove_small_CC`) that removes connected disparity components smaller than 40 pixels.

4. EXPERIMENTS

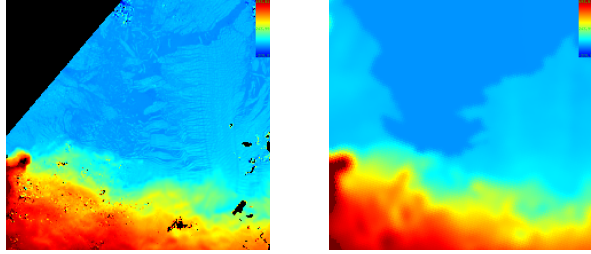
Our study focuses on the exploitation of optical data from SPOT 5 HRS sensor to monitor the Antarctic ice cap provided by the SPIRIT project. We selected two sites: one site on the Antarctic peninsula and one part of the Getz ice shelf. The imagery considered in our study is:

- The pair SPOT5/HRS acquired 2006-11-25 at 1:16pm. Coords. S66° 31', W62° 55' (Antarctic peninsula)
- The pair SPOT5/HRS acquired 2008-01-09 at 3:10pm. Coords. S75° 25', W108° 13' (Getz ice shelf)

However, we will only show results on the Getz ice shelf because of its diversity in cloud cover and large textureless regions. The results on the peninsula region are similar.

In Figure 4 we compare a detail of the reference DSM with the corresponding area produced by the proposed multiscale procedure. Note the high level of detail that can be attained using the SPOT 5 HRS images.

In Figure 5 we compare the result of the proposed method with the output of S2P [1] on the entire image over the Getz ice shelf. First, we should note that the small squares correspond to processing tiles. We observe that many tiles in the result of S2P (Figure 5a) have an unusually high value (bright). These tiles were poorly textured and affected by a thin cloud cover which lead to the estimation of the altitude



(a) High resolution DSM (our method) (b) Reference DSM (1km per pixel)

Fig. 4: We see that even though we have a reference DSM (b), the analysis of optical images is extremely important to obtain precise measurements of the surface heights (a).

of the cloud cover. The proposed multiscale method (Figure 5b) on the other hand managed to recover the small elevation changes over the large snowy areas in the central part of the image. Moreover, in the lower left part of the image it correctly detects and removes the cloudy areas. In both cases the post-process allows to remove most of the mismatches and clouds. Figure 3 shows a 3D visualization of the surface. Notice that there are only few outliers and that the surface is rather smooth.

5. CONCLUSIONS AND FUTURE WORK

This article addresses the problem of surface reconstruction from optical stereo satellite images over the Antarctica. To overcome difficulties specific to this region, we develop a stereo rectification approach using multiple scales of the original images. This enables a simple, reliable and automatic method, solely based on stereo pairs of satellite images, to create digital elevation models in polar regions with the highest possible quality.

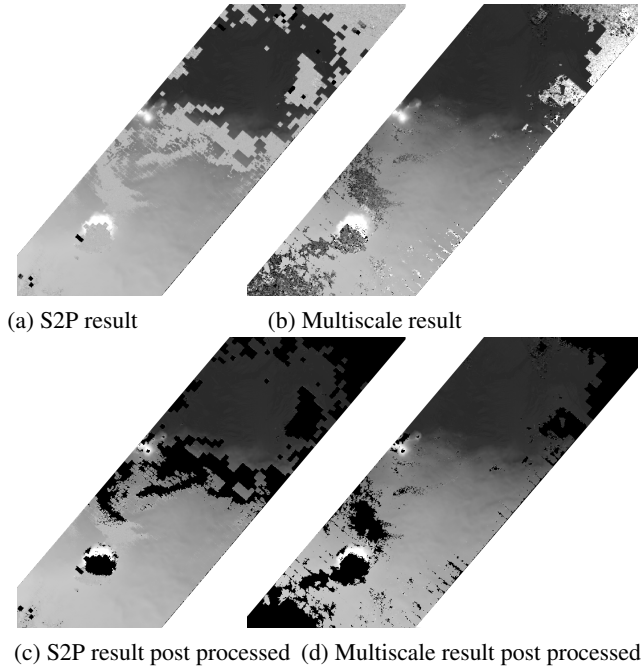
This will enhance the S2P pipeline enabling for large textureless regions to be handled correctly and automatically. It will also, using archive stereo images, enable for the creation of time series of elevation models over such regions. For future work we could reduce the disparity range search, using the reference DSM or the result obtained at coarser scales.

Acknowledgements. Work partly financed by Office of Naval research grant N00014-17-1-2552, DGA Astrid project “filmer la Terre” n° ANR-17-ASTR-0013-01, CNES (MISS Project).

6. REFERENCES

[1] C. de Franchis, E. Meinhardt-Llopis, J. Michel, J.-M. Morel, and G. Facciolo, “An automatic and modular stereo pipeline for pushbroom images,” in *ISPRS Annals*, 2014.

[2] R. Marti, S. Gascoin, T. Houet, D. Laffly, and P. René, “Evaluation du modèle numérique d’élévation d’un petit glacier de montagne généré à partir d’images stéréoscopiques Pléiades: cas du



(a) S2P result (b) Multiscale result (c) S2P result post processed (d) Multiscale result post processed

Fig. 5: Comparison of the results of S2P and our multiscale approach (black areas correspond to no data values). We notice that both smooth and abrupt height transitions are correctly handled in multiscale and clouds (bottom left) are correctly assigned a no data value, which is not the case in the standard approach.

glacier d’Ossoue, Pyrénées françaises.” *Revue Française de Photogrammétrie et de Télédétection*, vol. 208, pp. 57–62, 2014.

[3] E. Berthier, C. Vincent, E. Magnússon, A. P. Gunnlaugsson, P. Pitte, E. Le Meur, M. Masiokas, L. Ruiz, F. Palsson, J. Belart, and P. Wagnon, “Glacier topography and elevation changes derived from Pléiades sub-meter stereo images,” *The Cryosphere*, pp. 2275–2291, 2014.

[4] J. Korona, E. Berthier, M. Bernard, F. Rémy, and E. Thouvenot, “SPIRIT. SPOT 5 stereoscopic survey of Polar Ice: Reference Images and Topographies during the fourth International Polar Year (2007–2009),” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, no. 2, pp. 204–212, 3 2009.

[5] D. Poli, L. Zhang, and A. Gruen, “Spot-5/hrs stereo images orientation and automated dsm generation,” *ISPRS*, vol. 35, no. B1, pp. 421–432, 2004.

[6] G. Facciolo, C. Franchis, and E. Meinhardt-Llopis, “Automatic 3D Reconstruction from Multi-date Satellite Images,” in *IEEE CVPRW*, 2017.

[7] G. Facciolo, C. de Franchis, and E. Meinhardt-Llopis, “MGM: A Significantly More Global Matching for Stereovision,” in *BMVC 2015*, 2015, pp. 1–90.

[8] R. I. Hartley and R. Gupta, “Linear Pushbroom Cameras,” in *IEEE Trans. Pattern Anal. Mach. Intell.*, 1997, pp. 963–975.

[9] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *IJCV*, vol. 60, no. 2, pp. 91–110, Nov 2004.