



Cross-track satellite stereo for 3D modelling of urban areas

Thomas Krauß, Pablo D'Angelo & Lorenz Wendt

To cite this article: Thomas Krauß, Pablo D'Angelo & Lorenz Wendt (2018): Cross-track satellite stereo for 3D modelling of urban areas, European Journal of Remote Sensing, DOI: [10.1080/22797254.2018.1551069](https://doi.org/10.1080/22797254.2018.1551069)

To link to this article: <https://doi.org/10.1080/22797254.2018.1551069>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 10 Dec 2018.



Submit your article to this journal [↗](#)



Article views: 50



View Crossmark data [↗](#)

Cross-track satellite stereo for 3D modelling of urban areas

Thomas Krauß^a, Pablo D'Angelo ^a and Lorenz Wendt^b

^aDLR – German Aerospace Center, Oberpfaffenhofen, Germany; ^bDepartment of Geoinformatics – Z_GIS, University of Salzburg, Salzburg, Austria

ABSTRACT

Stereo imagery from very high-resolution satellite series can be used to create high-resolution digital surface models (DSMs) with sub-meter spatial resolution. Especially when in-orbit tri-stereo acquisitions are available as provided by the Pléiades constellation, buildings can be reconstructed well even in dense city centres. Unfortunately, such high-quality stereo data are often not available in the image archives and, thus, need to be newly acquired in most cases, leading to long delays and high cost. Especially for rapid mapping or crisis-response applications, high-resolution DSMs would be helpful for tasks such as city modelling and ultimately population estimation. In this work, we test the feasibility of producing DSMs by pairing two archive images, which were not acquired on the same day, in the same orbit or by the same satellite. We tested this approach for 14 images from Port-au-Prince, Haiti, and 12 images from Salzburg, Austria, from which we calculated a total of 136 DSMs. These were evaluated against Lidar DSMs. The resulting DSMs reach a completeness of up to 60%, and partly Median Absolute Deviation values comparable to in-orbit stereo pairs. For best results, the stereo convergence angle should be between 5 and 15°, and the difference in sun angles between the two images should be below 25–30.

ARTICLE HISTORY

Received 25 July 2018
Revised 15 October 2018
Accepted 18 November 2018

KEYWORDS

Photogrammetry; digital elevation model; satellite image; cross stereo; stereo processing; remote sensing

Introduction

Motivation

With more than half of the world's population living in towns and cities (United Nations, Department of Economic and Social Affairs, Population Division, 2014), urban areas get more and more into the focus of humanitarian relief organizations such as ICRC, Médecins sans Frontières (MSF), or SOS Children's Villages. Key information required for almost any intervention is an estimation of the population numbers for the towns and cities where these organizations operate in. As census data are often not available or outdated, population numbers have to be estimated by alternative methods such as remote sensing. The Department of Geoinformatics of the University of Salzburg Z_GIS has provided up-to-date maps of refugee camps and settlements of internally displaced people (IDP) to MSF and other organizations since 2011, which act as the basis for estimating the population numbers of these camps and settlements (e.g. Füreder, Lang, Rogenhofer, Tiede, & Papp, 2015a, Füreder, Tiede, Lüthje, & Lang, 2014, Lang, Tiede, Hölbling, Füreder, & Zeil, 2010). In the urban domain, such estimations are more challenging due to a more complex setting of multi-storey buildings and various building uses. Using digital surface models (DSMs) derived from very high-resolution (VHR) satellite imagery to account for the third dimension would be

helpful, but dedicated stereo- or tri-stereo image pairs/triplets acquired within a very short time or even in the same orbit by the same sensor are often not readily available in the image archives of satellite data providers when a disaster or crisis strikes. Acquiring such imagery takes valuable time and satellite resources, and might even be impossible, if a pre-event stereo model is needed, for example showing the situation *before* an earthquake.

In the X3D4Pop project (Wendt et al., 2018), researchers of Z_GIS, the Remote Sensing Technology Institute of the German Aerospace Centre DLR, UN Environment/GRID Geneva and AIT Austrian Institute of Technology tested in how far satellite image pairs acquired on different dates and/or by different satellites, here named cross-stereo models or cross-stereo pairs, are suitable for rapid DSM generation in crisis situations. To this end, mono, stereo and tri-stereo archive images acquired by Pléiades or WorldView satellites over Port-au-Prince, Haiti (14 individual images), Salzburg, Austria (12 images) and Mosul, Iraq (5 images) were collected and various possible stereo combinations were calculated. The results of these stereo processing experiments for Port-au-Prince and Salzburg are described in this paper. In a companion paper (Steinnocher et al., this issue), the effect of DSM/nDSM availability and quality under various scenarios of availability of additional data such as building footprints and land/building use information on the resulting population estimation is investigated. The tests over Mosul did

not aim at population estimation, but at the rapid assessment of (war) damages. They will be described in a separate paper.

The results of the cross-stereo processing were evaluated in terms of “completeness” (percentage of areal coverage of the test areas with valid elevation points) and “correctness” in comparison with DSMs derived from airborne laser scanning. For “correctness”, the Median Absolute Deviation (MAD) has been chosen as an outlier-tolerant deviation measure. It is detailed in the following section. The aim of this study is to find out what DSM quality can be expected from cross-date/orbit/sensor stereo pairs, and to define the parameters that lead to the best results in order to allow for a cost-effective choice of images to be paired from the totality of images that might be available in the image archives. Parameters that are expected to have an influence on the stereo-processing performance are the stereo-convergence angle between the two images, the sun angles and the difference of the sun angles between the images, the time between the two observations both in absolute terms and in terms of difference of day-of-year and the land cover. As the base to height ratio of a stereo pair is related to the stereo-convergence angle, we chose not to evaluate with respect the influence of the base to height ratio additionally. For the stereo convergence angle and the sun difference angle, the solid angles were used, not distinguishing between azimuth and elevation. For land cover, we distinguished between vegetated and non-vegetated areas.

Previous works

The use of across-track satellite images for stereo reconstruction has been studied in different contexts before. Choi, Kang, & Shin (2012) compare the geometric accuracy of in- and across-track stereo pairs using WorldView-2 and GeoEye-1 data using 17 GPS surveyed ground-control points. They found that in-track and across-track multi-sensor pairs have the same geometric accuracy but did not evaluate dense-matching or generation of DSMs. Lang (2014) evaluates DSMs generated from selected across-track datasets in Tripoli and Madrid. For these arid and semi-arid areas, different image combinations were evaluated, and guidelines for image selection were proposed. Multi-date and sensor image matching using pairwise matching and multi-view matching is evaluated by Ozcanli et al. (2015). Pairwise SGM (Hirschmüller, 2008) dense-matching followed by median DSM fusion is shown to provide the best results. Heuristics for the selection of stereo pairs as well as an improved merging of stereo pairs was performed by Facciolo, Franchis, and Meinhardt-Llopis

(2017) on the IARPA MultiView Challenge dataset (Bosch, Kurtz, Hagstrom, & Brown, 2016; IARPA, 2016), consisting of 47 WorldView-3 images. They show similar results to good in-track data can be archived when using the 38 diachronic images over urban areas in Buenos Aires. While most of the prior work aim to improve the DSM quality by merging as many mono datasets as possible, this is not always feasible, especially when commercial imagery needs to be bought. We, therefore, additionally emphasize investigating the influence of image acquisition parameters on the quality of a stereo pair.

Materials and methods

Input satellite imagery

Two study areas were chosen for this study, Salzburg, Austria and Port-au-Prince, Haiti. These areas were selected because for both areas, DSMs from airborne Laser scanning (ALS) are available as a reference. For Salzburg, gridded population data with a spatial resolution of 100 m was available to evaluate the population estimation described in Steinnocher et al.’s issue. Port-au-Prince shows a very varied city structure with multiple types of buildings from large industrial compounds via single family homes to informal settlements, providing a representative example for the cities where rapid information extraction due to a crisis situation may be required. All data used for this study was already available in the archives of Airbus DS and DigitalGlobe, respectively. The majority of the data used were acquired by the Pléiades constellation. The data were selected by the availability, giving priority to in-orbit stereo and tri-stereo acquisitions, to be able to test the influence of different stereo convergence angles, while keeping all other factors constant. An overview to the test these areas is provided in Figure 1. The test area in Port-au-Prince (to be more precise, between Port-au-Prince and Delmas) has an extent of approximately 2×2 km, the test area for Salzburg is about 6×6 km in size. Observation parameters of the imagery used are provided in Tables 1 and 2. Selected parameters of the individual image pairs are presented in the following chapters, while the full list of input parameters (and results) is found in the supplemental material to this article.

Reference data

The various optical DSMs derived from pairwise combinations of the above data were evaluated against DSMs from Lidar. For Salzburg, a DSM with a spatial resolution of 1 m, acquired in 2016, was used. It has a cell size of 1 m and can be obtained from the State of Salzburg for a fee.¹ The ALS-data for Port-au-Prince was acquired following

¹https://www.salzburg.gv.at/verwaltung_/Seiten/als-daten.aspx.

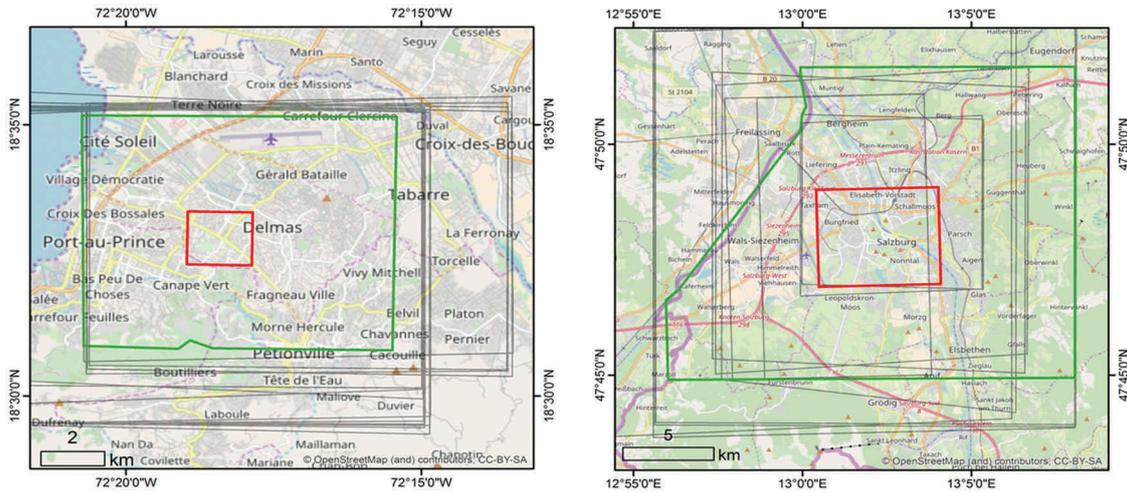


Figure 1. Test areas in Salzburg, Austria and Port-au-Prince/Delmas, Haiti.

Table 1. Satellite imagery used for Salzburg.

Sensor	Acquisition date	Code	Mono/stereo/ tristereo	Azimuth	Elevation	Sun Azimuth	Sun Elevation
Pléiades	09.09.2012	120909p	Mono	325.8°	62.3°	166.5°	46.8°
Pléiades	12.10.2012	121012r	Mono	244.6°	68.8°	171.5°	34.5°
Pléiades	28.05.2013	130528l	Stereo	230.77°	72.7°	157.9°	62.5°
Pléiades	28.05.2013	130528n	Stereo	335.67°	74.6°	157.4°	62.5°
Pléiades	01.09.2015	150901b	Tristereo	353.9°	78.1°	160.4°	49.3°
Pléiades	01.09.2015	150901d	Tristereo	310.82°	85.2°	160.8°	49.4°
Pléiades	01.09.2015	150901f	Tristereo	213.4°	76.1°	160.8°	49.4°
WorldView2	22.07.2015	150722f	Mono	251.16°	61.57°	159.90°	61.50°
Pléiades	31.10.2015	151031j	Mono	141.38°	72.0°	167.5°	27.5°
Pléiades	01.11.2015	151101h	Mono	72.54°	61.42°	165.48°	26.91°
WorldView2	19.12.2015	151219b	Mono	139.09°	81.57°	166.00°	17.80°
WorldView2	27.12.2015	151227d	Mono	271.83°	88.97°	166.20°	17.90°

Table 2. Satellite imagery used for Port-au-Prince.

Sensor	Acquisition date	Code	Mono/stereo/ tristereo	Azimuth	Elevation	Sun Azimuth	Sun Elevation
Pléiades	05.07.2013	130705B	Tristereo	233.17°	81.8°	73.51°	70.82°
Pléiades	05.07.2013	130705C	Tristereo	280.6°	84.7°	73.52°	70.82°
Pléiades	05.07.2013	130705D	Tristereo	330.24°	82.1°	73.51°	70.82°
Pléiades	06.07.2013	130706v	Tristereo	127.88°	76.2°	74.60°	68.97°
Pléiades	06.07.2013	130706w	Tristereo	73.77°	75.9°	74.70°	68.72°
Pléiades	06.07.2013	130706x	Tristereo	100.61°	77.5°	74.69°	68.74°
Pléiades	20.09.2014	140920Mk	Stereo	237.9°	81.2°	134.15°	65.79°
Pléiades	20.09.2014	140920Ml	Stereo	323.23°	81.7°	134.15°	65.79°
Pléiades	21.09.2014	140921Mg	Stereo	123.65°	76.7°	131.55°	64.18°
Pléiades	21.09.2014	140921Mh	Stereo	72.23°	75.8°	131.55°	64.18°
Pléiades	17.11.2014	141117Mp	Tristereo	242.44°	71.4°	159.05°	50.09°
Pléiades	17.11.2014	141117Mq	Tristereo	319.02°	72.4°	158.70°	50.00°
Pléiades	17.11.2014	141117Mr	Tristereo	274.31°	75.5°	158.70°	50.00°
Pléiades	29.04.2015	150429Md	Mono	302.23°	84.16°	99.59°	71.87°

the severe earthquake in 2010 (Worldbank, 2010). It also has a spatial resolution of 1 m/cell.

(Cross)-stereo processing

DSMs were calculated using the fully automatic processing chain CATENA at DLR (Krauß, 2014). CATENA is a multi-purpose, multi-sensor processing environment developed at the Remote-Sensing Technology Institute of the German Aerospace Centre (DLR). One of the most prominent general purpose chains is the multi-stereo processing chain supporting all VHR

stereo satellite scenes. In the project X3D4Pop, this chain was extended to also allow the processing of data from different sensors, dates and orbits instead of only the standard in-orbit-(multi)-stereo scenario.

The multi-stereo processing chain is shown in Figure 2. First, all of the requested stereo scenes are imported and tie points between the all scenes are matched using SIFT-matching and refined and transferred to all other images using local least squares matching. As local least squares-matching models the geometric and radiometric transformation between the images, high-quality tie points are available even

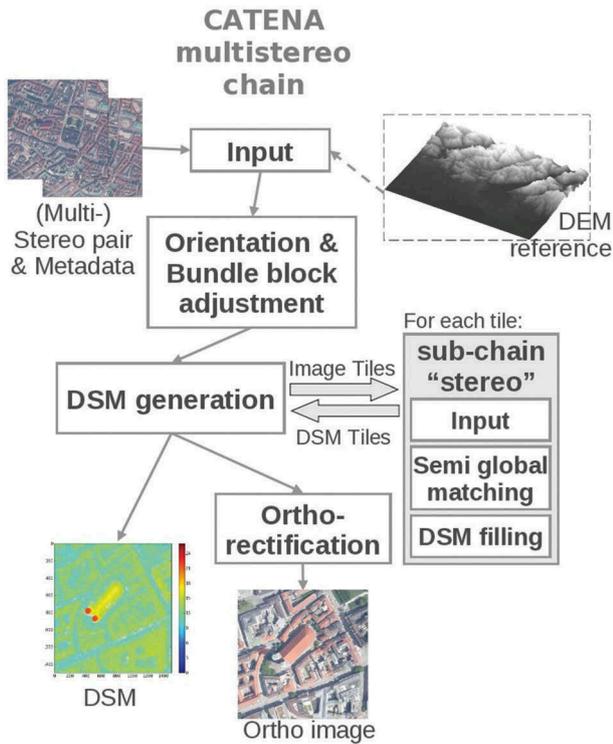


Figure 2. The CATENA multi-stereo processing chain.

for large viewing angle differences. Then, a bundle block adjustment estimates RPC correction parameters to ensure a consistent orientation of all images. Afterwards, all overlapping parts of the panchromatic images are cut to smaller tiles and re-ingested in CATENA using the stereo-chain. All these tiled stereo jobs are now processed again in parallel on many processing nodes.

The stereo chain takes all co-registered, overlapping images of its tile and does a forward- and backward-matching of each possible image pair using a variant of the semi-global-matching algorithm (SGM) (Hirschmüller, 2008), optimized for extraction of fine detail from satellite stereo imagery (d’Angelo & Reinartz, 2011). It includes the weighted use of Census and Mutual Information as stereo-matching costs, as well as a locally adaptive disparity search ranges. For each stereo pair, DSMs in the resolution of the input imagery (0.5m) are produced. Afterwards, in the case of more than only one stereo pair available, all pairwise results would be combined to a best-of-DSM, by computing the pixel-wise average of all height values within 5 m of the median. Holes in the unfilled DSM originate from occlusions (i.e. areas which are only visible in one of the stereo images), changes in the object appearance due to different viewing conditions, mismatches by the SGM algorithm or gridding effects due to re-projection of the DSM from image geometry to UTM coordinates. A filled (interpolated) version of the resulting DSM is generated and used for orthorectification of the image data, resulting in orthoimages with 0.5 m pixel size.

For this study, 34 GB of input satellite data were processed to 605 GB of cross-stereo DSMs and ortho-images in a total processing time of 26 days 23:33:20 on 12 hosts (64 processing nodes). A total of 136 individual DSMs were further evaluated, 91 for Port-au-Prince from Pléiades images, 36 for Salzburg from Pléiades images and nine for Salzburg from mixed pairs of Pléiades and WorldView-2 images.

Evaluation

To evaluate the quality of the DSMs, the un-interpolated (filled) output of the stereo processing was used without the involvement of ground control points (GCPs), as no GCPs were available, and thus relying on the attitude and positioning data provided with the raw satellite imagery and the SRTM DSM used during the image orientation, as the ALS data were not available in the early stages of the project. Image orientation was performed using all images of each site in a single block, using 4131 tie point in Salzburg and 19,816 tie points in Port-au-Prince, resulting in a tie point RMSE of 0.2 pixels. Due to the use of SRTM as main vertical and horizontal reference, a vertical shift was noticed between the (cross-) DSMs and the reference DSM especially for Salzburg, which are reflected by the mean and median quality metrics, but not in the MAD (see below), is independent to such global shifts. Only negligible horizontal shifts were observed. Also, the relatively long time between ALS data acquisition and image acquisition might have contributed to the quality metrics reported here.

Several quality metrics were considered and viewed. The most interesting metrics were found to be the completeness as defined as the fraction of areal coverage of the DSM of the study area, and the Median Absolute Difference MAD (Equation (2)) of the valid (not “no-data” points) (Höhle & Höhle, 2009). As both metrics describe relative changes, no further vertical co-registration between the stereo DSMs and the ALS data was required. Both quality metrics were evaluated both on the entire study area and on the vegetated and non-vegetated areas separately. This distinction was made using a manually defined threshold on the Normalized Differential Vegetation Index NDVI. In the case of Salzburg, water areas (mostly the Salzach river) were also extracted based on NDVI and added to the vegetation class. For Port-au-Prince, water areas did not play a role.

$$\Delta h = |h_{DSM} - h_{Ref_height}| \quad (1)$$

$$MAD = median(|\Delta h_i - median(\Delta h)|) \quad (2)$$

Results

Visual interpretation

To get a first impression of the effects controlling the DSM performance, it is worthwhile to have a look at a few selected DSMs for a qualitative, visual analysis. Figure 4 shows examples from Salzburg. In Figure 3(a), the subset of the most complete in-orbit DSM generated from one image pair is displayed, with an inset shown in Figure 3(d). This DSM has a completeness of 91% and a MAD of 0.93 m. The study area is almost completely covered, including the vegetated areas on the hills, with almost the only exception being the Salzach River. Minor data gaps also occur in shadowed or occluded areas directly next to building walls (Figure 3(d)). Further typical errors are slightly distorted building outlines due to a mismatch between backward and forward matching, a fine cross-hatching pattern caused by re-projection from image to map coordinates, and gaps caused by specular reflection of metal roofs or roof windows. In Figure 3(b,e), the most complete cross-stereo DSM for Salzburg is shown which has a time difference of more than 1 day. In this case, the time

difference is 826 days, or 96 days in terms of day-of-year (doy). The completeness is 53%, the MAD is 2.63 m.

The figure shows that in parts of the city building, roofs are quite well represented, in particular east of the river, in other parts the buildings contain many gaps. Large data gaps also occur in the streets and on squares and places, on shadowed areas like the northern face of the hill east of the river and on vegetated areas. The reason for the relatively well performance of this pair despite the long time gap between the two observations is the relatively low sun angle difference of only 13.3° and the low stereo angle of 5.5° . The image 150901b is part of a triplet, the image 130528n is part of a pair; so that in total, six combinations were calculated for this date combination. A comparison of the results shows that higher stereo angles, here between 5.5° and 25.8° , lead to better correctness (MAD drops to 0.79 m), but lower completenesses (highest: 53%, lowest: 14% overall completeness; see supplemental material for these numbers).

Figure 3(c) shows the ALS DSM for comparison. In Figure 3(e), cross-sections through the three DSMs

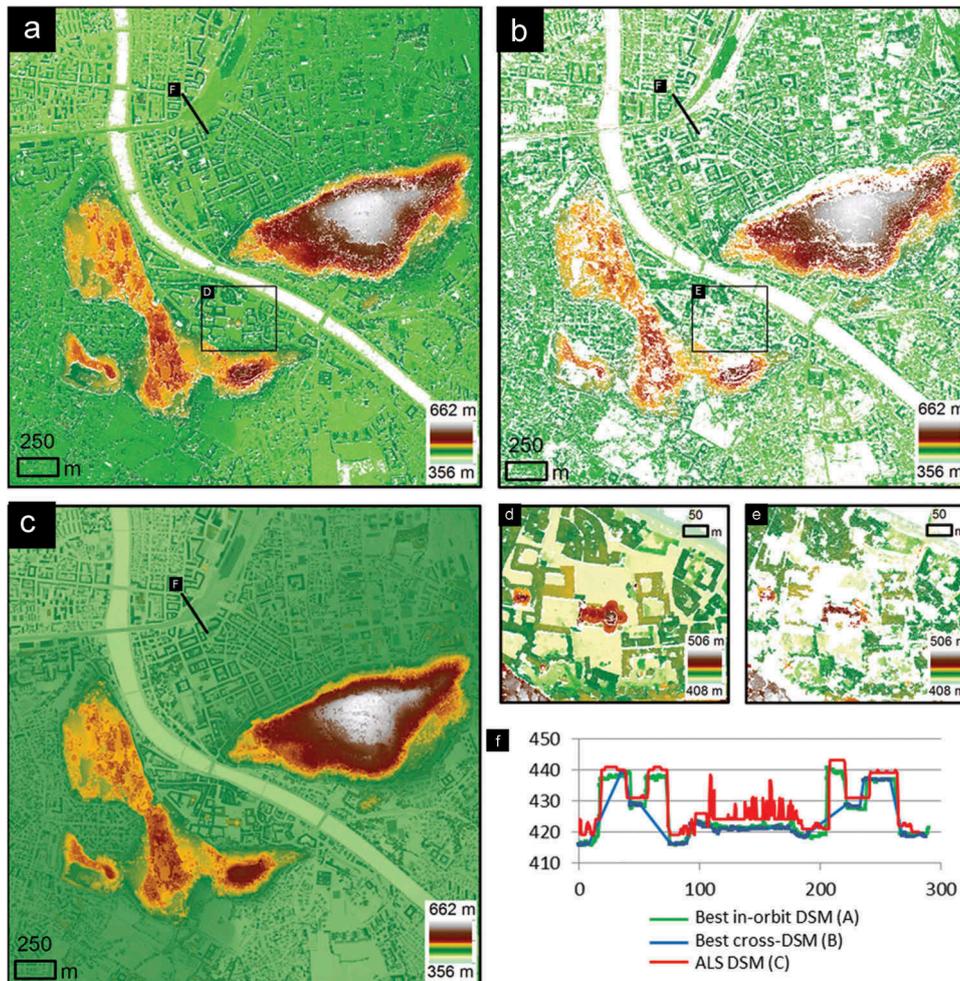


Figure 3. Comparison of DSMs from Salzburg. (a): best in-orbit DSM (150901b-150901d; overall completeness 91%, MAD 0.93 m). (b): most complete cross-stereo DSM for a time difference of at least 2d (150901b-130528n; overall completeness 53%, MAD 2.63 m). (c): Lidar DSM for comparison. (d): Subset of best in-orbit DSM (a), showing the cathedral of Salzburg. (e): Subset of best cross-stereo DSM (b) Note the data gaps. (f): Profiles of (a), (b) and (c). Data gaps are interpolated by straight lines.

are displayed. It shows an apparent systematic offset between ALS and optical DSMs and some minor discrepancies along building edges, in particular between the best in-orbit DSM (green) and the ALS DSM (red). The gaps in the cross-DSM (blue) are here interpolated with straight lines, emphasizing the missed building parts. The profiles also show that the reference DSM contains fine details like electrical overhead lines of the railway, which are not represented in the optical DSMs. Some smaller changes in the city area lead to additional minor changes between the optical DSMs and the reference.

Another example is shown in Figure 4. The top row shows the pair 140920l-141117r, which resulted in the highest overall performance for cross-stereo pairs with more than 1 day of time difference, for the study area in Port-au-Prince. The stereo-matching was performed on the panchromatic channel, but for easier recognition, the pan-sharpened images are shown here. The time difference here is 59 days, the convergence angle is 11° , sun angle difference 20.2° . The resulting DSM has a completeness of 61% and a MAD of 1.38 m. As in the example of Salzburg, buildings are quite well represented, with the exception of the large roofs showing specular reflection. Roads are also not well represented, possibly either because of building shadows or because of too smooth surfaces with a lack of distinctive features the stereo-matching algorithm can recognize. Areas with changing vegetation such as in the south-west of the subset also lead to data gaps. The lower row of Figure 5 shows a subset of the worst-performing pair for this study area, and the resulting DSM with an overall completeness of only 10%. The absolute time difference is 500 days (134 days-of-year) and the sun angle difference is 42.4° .

The unfavourable timing of the two images which results in the high sun angle difference leads to very distinctly appearing buildings with shade and shadow occurring at different sides in the two input images. The stereo angle of 29.4° is also relatively high and leads to many occlusions. The two images also show changes in built-up area (south-west of image), but this appears not to be the controlling factor, as also the apparently unchanged parts of the image do not provide a good DSM. Only few roofs are unaffected by the changes in illumination and result in a successful matching.

Evaluation of quality metrics

Overall performance

The best results in terms of overall completeness among all cross-DSMs were achieved by image pairs which were only 1 day apart. Here, the completenesses reach up to 76% (Port-au-Prince, 130705B-130706v). The MAD of this pair is 0.7 m. For comparison, in-orbit stereo pairs reach completenesses up to 97% (Port-au-Prince 130705B-130705C), the MADs of these pairs are between 0.38 m and 1.02 m. For stereo pairs with a larger time difference, the maximum overall completeness reached is 61%.

The influence of convergence angle and sun angle difference on completeness and correctness

The convergence angle appears to have an opposing effect on completeness and correctness of the resulting DSMs. Low convergence angles lead to higher completenesses (because the two images are more similar, resulting in more matches), but the correctness is lower (higher MAD). The MAD is particularly high if

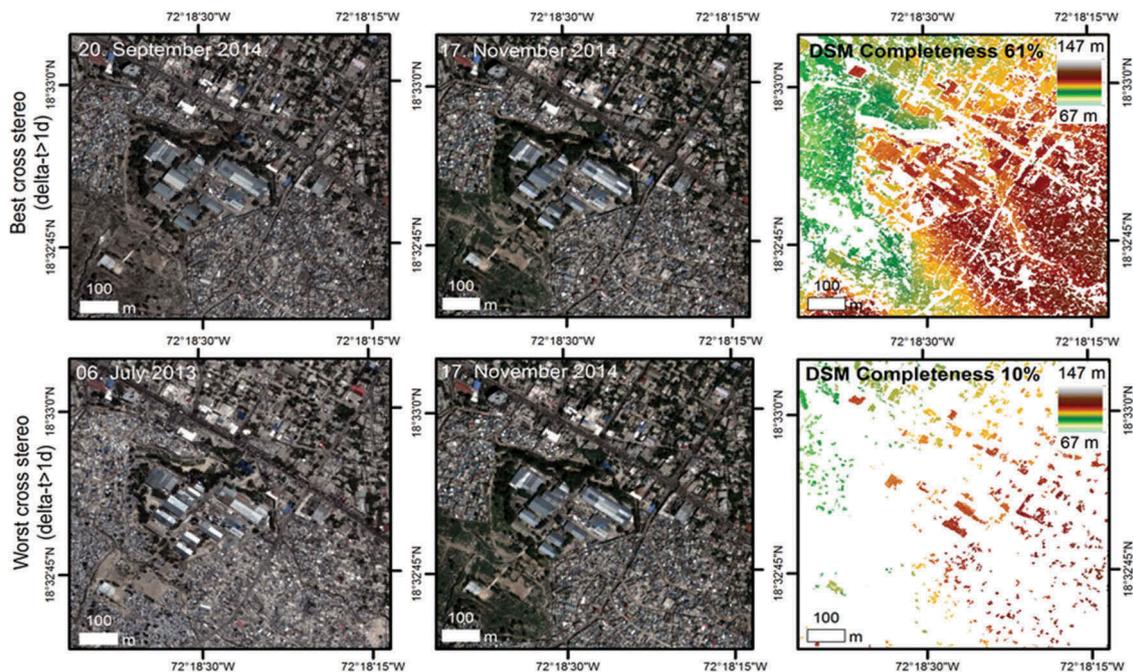


Figure 4. Comparison of well-matching cross-stereo pair 140920l-141117r (overall completeness 61%, MAD 1.38 m; top row) and worst cross-stereo pair 130706x-141117p for Port-au-Prince (overall completeness 10%, MAD 0.42 m; bottom row) .

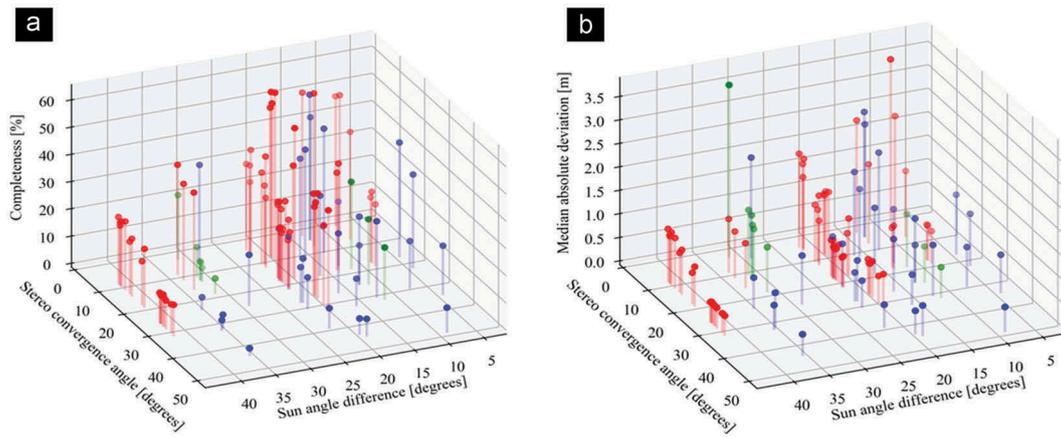


Figure 5. Dependency of completeness (a) and MAD (b) on-stereo convergence angle and sun angle difference. Red: results from Salzburg (Pléiades), green: results from Salzburg (WV-2 and Pléiades cross-sensor), blue: results from Port-au-Prince (Pléiades). Only pairs with convergence angles $> 2^\circ$ and a minimum time difference of 2 days are shown.

the stereo convergence angle gets too low. This is evidenced by a set of image pairs from Port-au-Prince, stereo convergence angles below 1° which all have comparably high completenesses of up to 50% and display MAD values of up to 13.4 m (see supplemental material). For convergence angles higher than $\sim 5^\circ$, the MADs reduce to values of approximately 3 m and lower. For high convergence angles, the completeness increases because of more occlusions in “urban canyons” and more differences between the two images.

The same behaviour is also shown in Figure 5. It shows the dependency of completeness and MAD on stereo convergence angle and sun angle difference for stereo angles higher than 2° and time differences larger than 1 day. Experiments from Port-au-Prince (blue), Salzburg, Pléiades only (red) and Salzburg, cross-sensor experiments Pléiades-WV2 (green) are shown together. The plot shows that with increasing stereo convergence angle, the MAD drops further, but also the completeness decreases further to values even below 10%.

A further important influencing factor is the sun angle difference, also shown in Figure 5. The higher the sun angle difference, the lower the completeness. For example, to reach completenesses above $\sim 30\%$, the sun angle difference should not be higher than $\sim 25^\circ$. The correctness is less influenced by the sun angle difference. MADs below 1 m are found for all sun angle differences in this experiment set. This suggests that those points that get matched despite unfavourable sun angles are likely to be correct.

Thus, the ideal combination to reach both high completeness and high correctness (MAD) appears to be sun differences as low as possible, and stereo convergence angles around $5\text{--}15^\circ$. This corresponds to approximate B/H-ratios of 0.087 to 0.26. For lower stereo angles, the MAD rises strongly, and the results are likely to be not

useful. For higher stereo angles or higher sun difference angles, the completenesses drop, but the remaining points are likely to be valid. The optimal stereo angles (or B/H ratios) found here are actually lower than those recommended by Airbus for Pléiades acquisitions,² which range from ~ 14 to 22° (B/H 0.25 to 0.4).

The influence of absolute sun elevation on completeness

Another controlling factor for both in-orbit and cross-stereo pairs is the absolute sun elevation at the time of image acquisition. This is shown by the image pair 151031j-151101h for Salzburg. Although these images were taken only one day apart and thus show a very low sun angle difference of 1.9° , the low sun elevation at this time of the year (31 Oct/1 Nov) of only 27.5° leads to long shadows, in which the both images show very little texture, so that the matching does not work properly. This leads to a relatively low completeness of only 53%. The high convergence angle of 27.3° is also unfavourable, as it leads to occlusions in the narrow alleys of the old town of Salzburg, but this appears not to be the controlling factor, as data gaps also occur in flat, open spaces which are well visible in both images, but are covered by cast shadows.

The effect of vegetation on completeness correctness

To test which influence vegetation has on the performance on the stereo-matching, all quality metrics were calculated both for the entire study areas and separately for the vegetated and non-vegetated areas (with water added to vegetated areas in the case of Salzburg). The completenesses for Salzburg are shown in Table 3, which are representative also for the study area in Port-au-Prince. The table shows that for the in-orbit stereo pairs, the completenesses for vegetated and non-

²Pleiades Imagery User Guide. Astrium, 2012 <http://www.intelligence-airbusds.com/en/4572-pleiades-technical-documents>.

Table 3. Comparison of completenesses for vegetated and non-vegetated areas for Salzburg.

Overall completeness Salzburg												
	Color scale 0–100											
Image	120909p	121012r	130528l	130528n	150901b	150901d	150901f	151031j	151101h	150722f	151219b	151227d
120909p		28	10	40	42	34	18	6	6			
121012r			35	18	22	29	36	17	9			
130528l				57	14	23	45	4	2			
130528n					53	45	24	5	3			
150901b						91	68	16	13	19	5	11
150901d							82	19	11	24	7	23
150901f								19	7	32	5	6
151031j									53			
151101h												

Non vegetation completeness Salzburg												
	Color scale 0–100											
Image	120909p	121012r	130528l	130528n	150901b	150901d	150901f	151031j	151101h	150722f	151219b	151227d
120909p		34	18	49	52	44	25	12	12			
121012r			43	27	31	38	44	28	17			
130528l				59	24	36	53	10	6			
130528n					59	51	33	10	8			
150901b						93	70	27	24	27	10	17
150901d							84	32	22	32	12	28
150901f								33	15	37	9	11
151031j									57			
151101h												

Vegetation and water completeness Salzburg												
	Colour scale 0–100											
Image	120909p	121012r	130528l	130528n	150901b	150901d	150901f	151031j	151101h	150722f	151219b	151227d
120909p		24	6	36	37	29	15	3	3			
121012r			31	15	17	25	32	12	5			
130528l				56	9	17	41	2	1			
130528n					50	42	21	2	1			
150901b						90	66	11	8	11	2	6
150901d							81	13	6	15	3	16
150901f								12	3	23	2	3
151031j									51			
151101h												

vegetated areas are very similar, with only a few percentage of difference. For cross-stereo DSMs, the completenesses for vegetated areas are always up to 15% lower. The MAD is lower for non-vegetated areas than for vegetated areas. As vegetation changes with the seasons, both effects have been expected.

The effect of cross-sensor image pairs

We also obtained WV-2 images for Salzburg for three different dates, which we paired with the three Pléiades tristereos scenes of 1 September 2015 (Table 3 and supplemental material). These image pairs resulted in overall completenesses between 5 and 32%. Again, the highest MAD of 3.64 m was found for a low convergence angle of 4.1°. Thus, the cross-sensor experiments appear to be in line with the results obtained from Pléiades-Pléiades image pairs, suggesting that mixed pairs deliver approximately similar results. Note, however, that for only nine cross-sensor experiments have been performed.

Further effects influencing the DSM-generation performance

The image pair 130528l-130528n from Salzburg is unusual among the pairs of images investigated in this study. This pair was acquired as in-orbit stereo

pair on the same day, thus with virtually no sun angle difference and a sun elevation of 62.5°.

Among the image pairs used in this study was one pair from Salzburg. Despite these favourable conditions, the overall completeness only reached 56%, which is in the order of cross-stereo pairs. The overall MAD is 0.52 m. The reason for the relatively low performance might be the rather high convergence angle of 25.8°. Another reason may be an unfavourable orientation of camera-viewing angles, rooftops and illumination, which leads to specular reflection so that the two images become too different to be properly matched. In any case, this observation, however, shows that although in most cases, in-orbit stereo pairs perform much better than cross-stereo pairs, this is not always the case.

Merging of cross-track datasets

One way to improve the quality of cross-track data is to merge the results of multiple stereo pairs (Ozcanli et al., 2015). This leads to dense DSMs with good completeness and accuracy values. For Port-au-Prince, a completeness of 99.8% and a MAD of 0.85 m are archived. Slightly lower values of 94% completeness and a MAD of 1.2 m have been computed for the Salzburg across-track scenes, mainly due to larger differences in the vegetated areas.

Discussion and conclusions

The above described experiments have shown that cross-stereo processing is possible, but the resulting DSMs generally have a lower completeness. The best cross-stereo DSMs reached completenesses around 60%, many were in the order of 30–40%, but also values as low as 10% were observed. At the same time, the MAD values are partially significantly higher than for in-orbit DSMs. The most important controlling factors appear to be the convergence angle and the difference in sun angles. The convergence angle should be in the order of 5–15°. For smaller convergence angles, the correctness, here calculated as MAD, increases strongly. For higher convergence angles, the completeness drops strongly. The sun angle difference should be lower than approximately 25°, while the sun elevation of both images should be as high as possible to avoid hard-to-match cast shadows, which contain little image information.

Cross-sensor-matching (matching between images from different satellites) is possible and performs as well as matching with data from the same sensor, provided both have approximately the same spatial resolution.

Vegetated areas are more difficult to match in cross-stereo pairs than non-vegetated areas, but the difference is only around 10–15%, or in other words, while the cross-stereo-matching works better for non-vegetated areas, it also is far from perfect and reaches completenesses of only up to ~60%.

With these limited completenesses and elevated errors, the question arises how the resulting DSMs can be used. If the application at hand can tolerate patchy data, because only a rough estimation of prevalent heights is sufficient, using the data as-is may be feasible. Otherwise, it has to be combined with other data to create a more complete dataset. This other data can further DSMs from other cross-stereo pairs, as shown above. Another option might be the combination with a digital terrain model of lower resolution to extract building heights, or the combination with other types of data, such as building footprints, which would also assist in reconstructing the structure of a city.

Acknowledgments

This work was part of the research project “Suitability of mixed-sensor derived 3D data for rapid urban population estimation in crisis situations (X3D4Pop)”. It was partially funded by the Austrian Research Promotion Agency (FFG) within the Austrian Space Applications Program (ASAP) under grant FFG-ASAP-859796, the Swiss Space Office SSO and the German Aerospace Center DLR. We thank European Space Imaging (EUSI) for providing the WorldView-2 scenes of Salzburg.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Österreichische Forschungsförderungsgesellschaft [859796] and the German Aerospace Center.

ORCID

Pablo D'Angelo  <http://orcid.org/0000-0001-8541-3856>

References

- Bosch, M., Kurtz, Z., Hagstrom, S., & Brown, M. (2016). A multiple view stereo benchmark for satellite imagery. In *2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)* (pp. 1–9). Washington, DC: IEEE. doi:10.1109/AIPR.2016.8010543
- Choi, S.Y., Kang, J.M., & Shin, D.S. (2012). A comparison of accuracies of the RPC models of homo- and hetero-type stereo pairs of geosy and worldview images. *ISPRS Annales Photogrammetry Remote Sensor Spatial and Information Science, I-4*, 65–69. doi:10.5194/isprsannals-I-4-65-2012,2012
- d'Angelo, P., & Reinartz, P. (2011). *Semiglobal matching results on the isprs stereo matching benchmark*. ISPRS Hannover Workshop 2011 High-Resolution Earth Imaging for Geospatial Information, Hannover, Germany.
- Facciolo, G., Franchis, C. D., & Meinhardt-Llopis, E. (2017). Automatic 3D reconstruction from multi-date satellite images. *IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 1542–1551), Honolulu. doi:10.1109/CVPRW.2017.198
- Füreder, P., Lang, S., Rogenhofer, E., Tiede, D., & Papp, A. (2015a). Monitoring displaced people in crisis situations using multi-temporal VHR satellite data during humanitarian operations in South Sudan. In A. Car, T. Jekel, J. Strobl, & G. Griesebner (Eds.), *GI_forum 2015 – geospatial minds for society* (pp. 391–401). Wien: Verlag der Österreichischen Akademie der Wissenschaften. doi:10.1553/giscience2015
- Füreder, P., Tiede, D., Lühje, F., & Lang, S. (2014). Object-based dwelling extraction in refugee/IDP camps - challenges in an operational mode. *South-Eastern European Journal of Earth Observation and Geomatics*, 3(2S), 539–544.
- Hirschmüller, H. (2008). Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2), 328–341. doi:10.1109/TPAMI.2007.1166
- Höhle, J., & Höhle, M. (2009). Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(4), 398–406. doi:10.1016/j.isprsjprs.2009.02.003
- IARPA. (2016). Multi-view stereo 3D mapping challenge. Retrieved from <https://www.iarpa.gov/challenges/3dchallenge.html>
- Krauß, T. (2014). Six years operational processing of satellite data using CATENA at DLR: Experiences and recommendations. *Kartographische Nachrichten*, 64, 74–80.
- Lang, F. (2014). *Analyse digitaler Oberflächenmodelle aus multi-temporalen hochauflösenden optischen Satellitendaten* (Masterarbeit). Hochschule für Angewandte Wissenschaften München.
- Lang, S., Tiede, D., Hölbling, D., Füreder, P., & Zeil, P. (2010). Earth observation (EO)-based ex post assessment

- of IDP camp evolution and population dynamics in Zam Zam, Darfur. *International Journal of Remote Sensing*, 31 (21), 5709–5731. doi:[10.1080/01431161.2010.496803](https://doi.org/10.1080/01431161.2010.496803)
- Ozcanli, O.C., Dong, Y., Mundy, J.L., Webb, H., Hammoud, R., & Tom, V. (2015). A comparison of stereo and multiview 3-D reconstruction using cross-sensor satellite imagery. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 17–25). Boston, MA: IEEE. doi: [10.1109/CVPRW.2015.7301292](https://doi.org/10.1109/CVPRW.2015.7301292)
- United Nations, Department of Economic and Social Affairs, Population Division. (2014). *World urbanization prospects: The 2014 revision*. New York, NY: United Nations.
- Wendt, L., Krauß, T., D’Angelo, P., Tiede, D., D’Oleire-Oltmanns, S., Kulesa, K., ... Lang, S. (2018). Bevölkerungsabschätzung in Städten anhand von Stereo-Bildpaarungen, die nicht am selben Tag aufgenommen wurden. *AGIT ? Journal für Angewandte Geoinformatik*, 4, S.264–S.270. Wichmann Verlag, VDE VERLAG GMBH Berlin, Offenbach. ISBN 978-3-87907-647-5, ISSN 2364-9283, eISSN 2509-713X. doi:[10.14627/537647034](https://doi.org/10.14627/537647034)
- Worldbank. (2010). World Bank - ImageCat Inc. - RIT Haiti Earthquake LiDAR dataset. *Survey date*. Retrieved January 21– 27, 2010, from <http://opentopo.sdsc.edu/datasetMetadata?otCollectionID=OT.072010.32618.1> doi: [10.5069/G96Q1V50](https://doi.org/10.5069/G96Q1V50)