Generation of Digital Terrain Models from Digital Surface Models using a Watershed Transformation Approach

Thomas Krauss

DLR, Muenchener Str. 20, 82234 Wessling, Germany

ABSTRACT

The basis of Urban Digital Twins are 3D city models generated from airborne or spaceborne remote sensing imagery. These remote sensing data allows the extraction of digital surface models (DSMs) representing the absolute geographical height of each pixel together with orthorectified image mosaics. But for the derivation of a city model also the height of the urban objects and the height of the ground – the digital terrain model (DTM) – is needed. So in a first step there has to be a method to derive the digital terrain model (DTM) from the generated digital surface model (DSM). The extraction of DTMs from DSMs is already a basic task for many decades and there are many approaches but still there is no general fully satisfying solution for this demand. In this paper we present a novel method for deriving a DTM from a provided urban DSM based on the inverted watershed-transformation. The watershed-transformation normally fills up sinks in a DSM up to a ridge-line where two sinks meet. So the result is a segmentation of a DSM to distinct sinks. Using the inverted watershed-transformation delivers a segmentation to single elevated objects. The boundary points where each of these elevated segments meet are taken as possible candidates for ground areas. But such boundaries may also lie for example between adjacent roofs. Using a dynamically determined threshold in the height distribution of each watershed segment gives a possible ground candidate area of the segment. Merging all ground candidate areas and taking only larger areas with special properties into account gives all ground candidates. Keeping only the heights of these areas as real ground and filtering and filling them gives the final DTM. In the presented paper we describe the method in detail and apply the method to different DSMs of rather flat and hilly urban and non-urban environments, discuss the results and give finally an evaluation of the quality of the results.

Keywords: Digital Surface Model, Digital Terrain Model, DSM, DTM, Watershed-transformation, Urban Digital Twin, 3D City Model

1. INTRODUCTION

To generate a good city model we need to extract all urban objects which are located on the ground. Since optical remote sensing from airborne or satellite data delivers only the surface of the objects – the digital surface model (DSM) – we need a method to extract the ground – the digital terrain model (DTM) – from the DSM. There exist already many approaches but each of them has different problems. Fig. 1 shows a cross-section of an urban river. As can be seen the river is enclosed in steep walls with roads on top near the river.

Since the river-surface should be included in the DTM but also the roads on top of the walled river banks are part of a good DTM most existing approaches fail in such cases. As can be seen in fig. 1 the DSM (in blue) follows strictly the surface of the river and the river bank. The classic, morphological algorithm shown in green smooths out the border of the river. In this case the roads on the borders are higher than the ground and will be classified as e.g. buildings. A newly developed AI algorithm shown in orange performs better at the borders but give wrong results on the river itself.

So we developed a new DSM-to-DTM algorithm based on a watershed classification which is shown in fig. 1 in red. As can be seen the new proposed method follows strictly the road-surface on the left and the whole river. On the right the walled river-bank is covered by trees so the next ground can be only found on a road behind the trees and such the steep river edge smoothes out in this case.

After a short overview on existing DTM extraction algorithms in section 2 we describe the proposed algorithm in detail in section 3. This is followed by applying the method to different test datasets and evaluating them in section 4 and the final conclusion and outlook in 5.

Further author information: E-mail: thomas.krauss@dlr.de

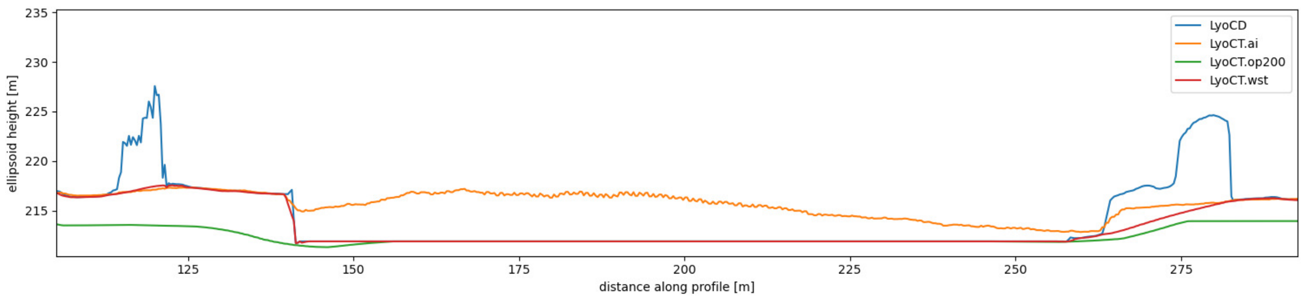


Figure 1. Digital terrain models across a river enclosed in an urban walled river bed (the Saône in Lyon), blue: DSM, red: proposed algorithm, green: classic algorithm, orange: AI algorithm

2. PREVIOUS WORK

Since the 1980s there is intense research[1] on digital terrain models (DTMs). Since then DTMs are generated from direct terrain measurements or extracted from digital surface models (DSMs). DSMs origin mostly from laser scanning, radar interferometry or from stereo processing of optical aerial or satellite imagery. DSMs represent the surface of all objects on the ground while a DTM should only represent the ground information without any objects located on it. But this leaves also room for individual interpretation – e.g. is the surface of a bridge a ground or is the ground the river below?

A DTM can be derived from a DSM by detecting and removing all objects and filling these areas in an intelligent manner. The problem of generating a DTM from an existing DSM is therefore mainly the detection of off-ground objects. Beside the classical method of removing elevated objects manually also a wide variety of DSM filtering algorithms where developed.

In an overview paper on DTM extraction algorithms[2] the algorithms "Classical morphological approach",[3] "Geodesic dilation"[4] and "Steep edge detection"[5] were compared. Later there were proposed the "MSD filtering"[6] based on multi-directional processing as well as slope dependent filtering, "TERRA"[7] (Terrain Extraction from elevation Rasters through Repetitive Anisotropic filtering) and finally the end-to-end deep learning approach for DTM generation "DSM2DTM".[8]

As shown in 2 all algorithms has advantages and disadvantages depending on the situation of the surrounding. Actually the best multi-purpose algorithm will be the AI based algorithm as described in 8. But if there's not the possibility to use the individually trained special AI approach we searched also for a more simple approach which will work in comparable quality in urban regions and also satisfy the above mentioned problems near walled river-beds.

3. METHOD

In this section we present the newly developed approach of DTM generation based on the watershed-transformation. The proposed method works in six steps which are described in detail in the following sub-sections:

1. Calculate the inverted watershed-transformation of the DSM
2. For each watershed-segment make a height histogram, search first minimum of counts and mark all DEM values below as ground
3. Check all found ground-areas and mark areas with sizes smaller than a given limit as unsure ground
4. Check all marked unsure ground areas and decide if they connect properly to a nearby valid ground area, unmark them in this case
5. Grow the ground mask using the provided DSM by increasing it to neighbouring DSM values if these are below a given slope
6. Create a DTM by using only the DSM values of the ground mask and interpolate and optionally filter it

3.1 Step 1: Inverted Watershed Transformation

An inverted watershed transformation is applied to the DSM. This operation segments the DSM to separate elevated objects as shown in fig. 2.

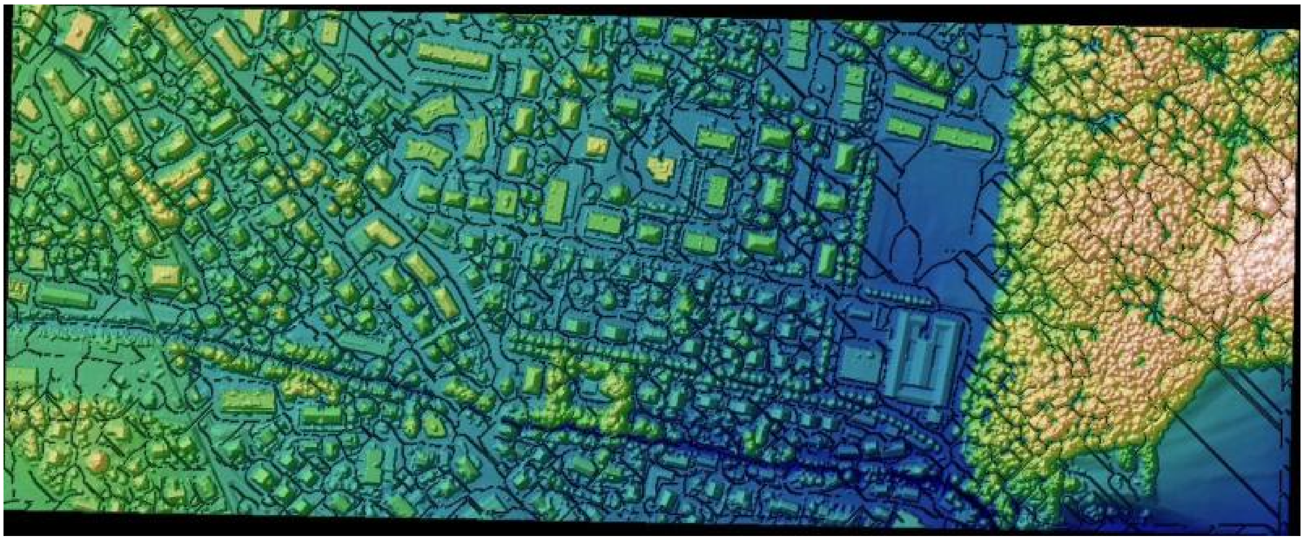


Figure 2. Shaded DSM of test region Bulle (Kanton Fribourg, Switzerland), 1.2×0.5 km^2 with marked watershed segments

The original description of the watershed transformation was given by Beucher in 1982.[9] We improved the method by using a runoff approach which starts with a pixel in the (sligthly Gaussian filtered) DSM and tracks it until it reaches a sink or an already visited pixel. Each sink is marked as a different watershed segment. Inverting the DSM will result in elevated objects like buildings or trees behaving like sinks. All pixels which "flow" to the same sink are supposed to belong to the same segment.

3.2 Step 2: Extraction of ground regions per segment

For each watershed segment a height histogram on the DSM is calculated as shown in fig. 3, right ("hist"). This histogram is filtered with a small Gaussian filter (size 2 m, "histg") and the local minima and maxima are searched ("lo" and "hi"). The first minimum and the last maximum are taken as thresholds for a "low" and a "high" class. All values between are marked as "unsure". All DSM values in the masked segment are now set to one of the classes "low", "high" or "unsure" as shown in fig. 3, center. Fig. 3, left shows the associated DSM of the masked watershed segment.

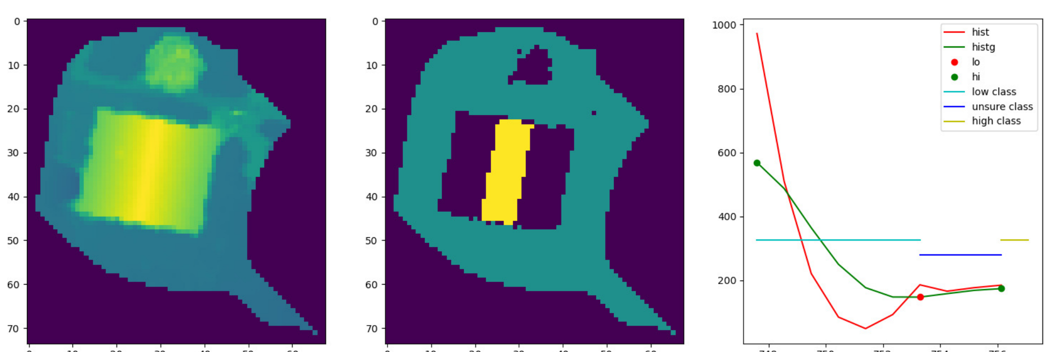


Figure 3. DSM (left), classes (center, yellow: high, cyan: ground, dark-blue: unsure) and height histogram of one watershed segment

3.3 Step 3: Select ground by size

After processing all watershed segments there are many potential ground areas in the class image spreading across multiple watershed segments. But there may also be building structures (mainly on top of buildings) where the detected low area of the watershed segment belongs not to a real ground but rather to an other roof or tree top. To detect and elimate these cases in this step the detected ground areas are filtered by size.

Ground area regions where both widths or heights are below a given size (defaulting to 200 m) or only one is below 1/5 of this size are also marked as unsure for the next step.

3.4 Step 4: Fix unsure areas

In this step for all previously found unsure ground class areas it is checked if they are connected properly to a nearby valid ground area and they get marked also as ground areas.

For this the 10 and 90 % percentiles of the heights of the unsure ground and a sure ground in a surrounding of the same size as in the previous step are compared and the unsure area set to ground if the low value of the unsure is near the high value of the ground.

3.5 Step 5: Grow ground areas

After gaining the sure ground regions in the previous steps an iterative increasing of the ground mask is performed by checking the slope between ground and non-ground neighbour pixels. If the slope is below a given value (e.g. 10 % – which means 3 cm height difference at a image resolution of 30 cm between neighbouring pixels) the ground mask is grown to these pixels.

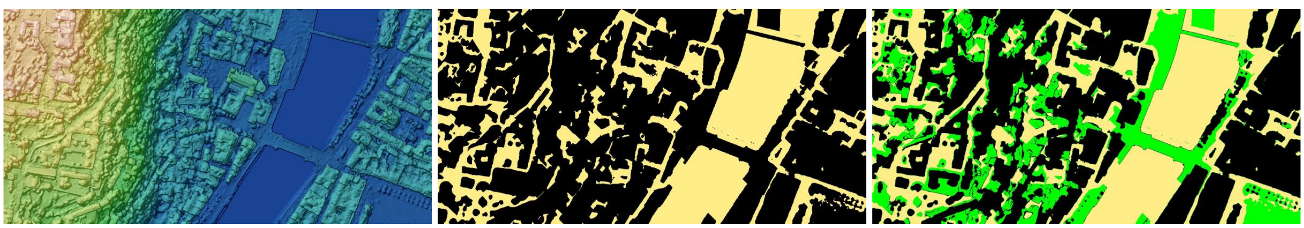


Figure 4. Detected ground areas; left: DSM of a test area 1500×750 m^2 in Lyon, France, center: original ground areas from wst, right: ground areas after growing

Fig. 4 shows the ground areas derived from a DSM of Lyon (left). In the center the original ground mask after extraction from the watershed segments and fixing unsure areas is shown. On the right side the iteratively grown groundmask from this step is shown with green areas depicting filled ground.

As can be seen areas on bridges and beneath riverbeds are usually not marked as ground since there are lower regions in the surrounding. But in this case the detected ground near the buildings can grow to the areas above the river.

3.6 Step 6: Create DTM from ground areas

In this final step first a sparse DTM is created by using only DSM values of the obtained ground mask containing the heights of detected ground areas and no-data values otherwise. Finally this sparse DTM is interpolated and optionally filtered with a small Gaussian filter to receive the final DTM as shown in fig. 5.

4. EVALUATION

For the evaluation of our algorithm in comparison with ground truth and other DTM generation algorithms we use following datasets:

- FriRefCity: A reference DSM and DTM of SwissTopo for the city of Bulle in the Kanton Fribourg at a ground sampling distance of 50 cm and a size of 1164×477 m^2

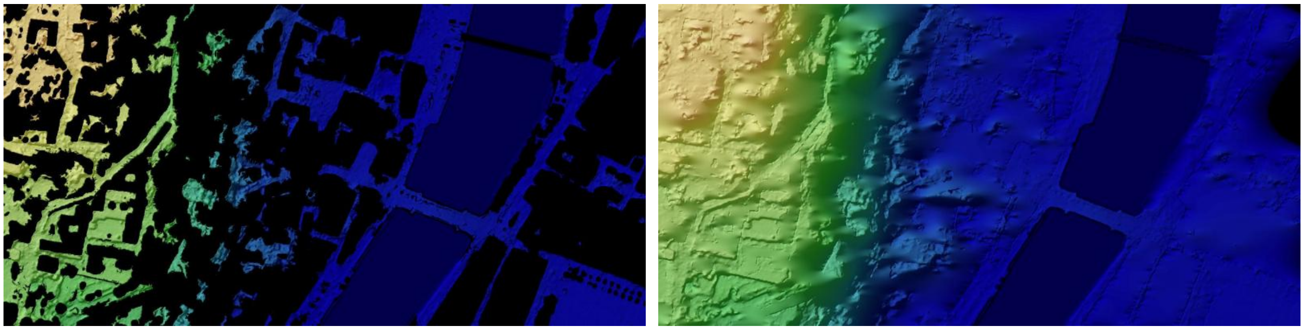


Figure 5. Final DTM of test area Lyon, left: DSM values on ground mask, right: interpolated final DTM

- FriRefBerg: A reference DSM and DTM of SwissTopo for a mountain range around the Dent de Vounetse (1812 m) east of Bulle also in the Kanton Fribourg at a ground sampling distance of 50 cm and a size of 2754×1771 m^2
- LyoC: A DSM of a section of the city of Lyon (France) derived from WorldView-1 satellite data acquired on 2018-10-18 at 13:40 UTC (© European Space Imaging 2018) covering 1500×750 m^2 with a ground sampling distance of 30 cm. For this test region no ground truth DTM was available.

In this evaluation we compare our results for the three mentioned test areas with the following DTM extraction algorithms:

- "Classical morphological approach"[3] ("op" for morphological opening)
- "TERRA"[7] ("ter" for TERRA)
- "DSM2DTM"[8] ("ai" for the AI-based approach)
- "wst" for the watershed based approach presented in this paper

4.1 Comparison of DTMs

In our first evaluation we calculate the DTMs using the above mentioned algorithms and compare them with the ground truth or with the AI algorithm in case of missing ground truth at the test site Lyon.

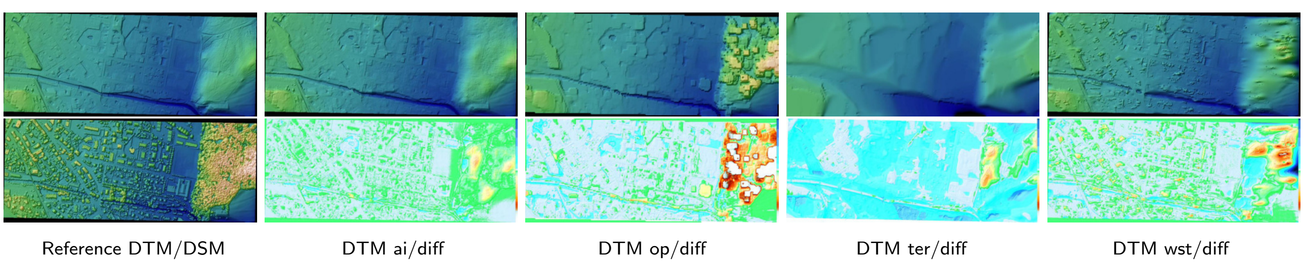


Figure 6. Derived DTMs for test region FriRefCity, top row l.t.r.: ground truth DTM, DTM ai, DTM op, DTM ter, DTM wst, bottom row l.t.r.: original DSM, differences of DTM ai, op, ter and wst to ground truth, color range from blue (DTM is 20 m lower than reference) to white (DTM is 20 m higher than reference, see also color bars on the right edges)

The first comparison for the test region FriRefCity (fig. 6) shows the best results for the AI algorithm followed by the proposed watershed algorithm for the urban area. The classical algorithm fails for large area buildings while the terra algorithm encounters larger problems on the bottom of the DTM. In the forest area on the right also the AI algorithm scores best followed by terra, wst and the worst for the classical algorithm. The proposed wst algorithm encounters problems in the forest region due to too small detected ground regions.

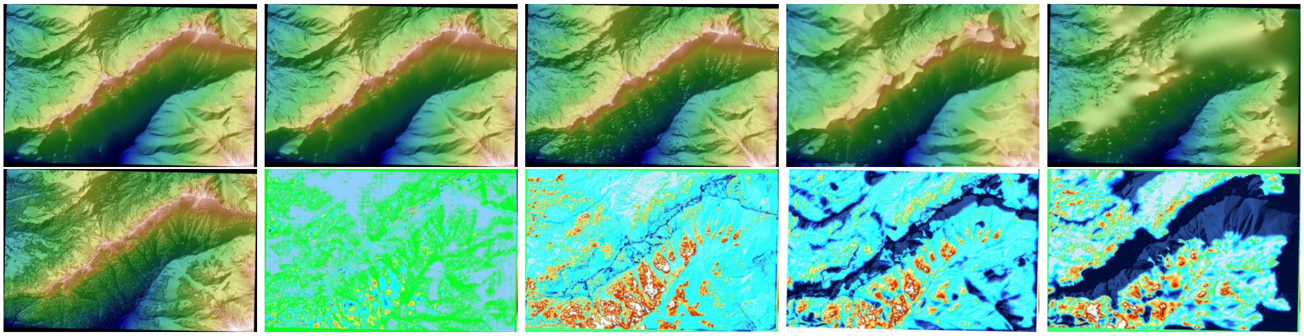


Figure 7. Derived DTMs for test region FriRefBerg, top row l.t.r.: ground truth DTM, DTM ai, DTM op, DTM ter, DTM wst, bottom row l.t.r.: original DSM, differences of DTM ai, op, ter and wst to ground truth, color range from blue (DTM is 20 m lower than reference) to white (DTM is 20 m higher than reference, see also color bars on the right edges)

The second test region FriRefBerg (see fig. 7) shows problems for some of the algorithms. While the AI approach reflects nearly perfectly the ground truth (since it's trained on such regions) the classical approach gives the best results for the other methods but estimates the ground up to 20 m too high for the forest areas on the slopes of the mountains. The summit areas are underestimated by all three non-AI algorithms: for the classical approach by about 14 m, for terra by 60 m and for the proposed wst even by 230 m. This is due to non-exising ground-classes on the top areas of the mountains. Since wst derives the ground from the watershed classes and in this mountain case the watershed classes go far down the mountain slopes the ground areas are massively underestimated. Increasing the allowed slope for the region growing of the ground areas in the wst algorithm to 100 % or more improves the results for mountaineous areas enormous, but makes it nearly useless in urban ares as shown in fig. 8.

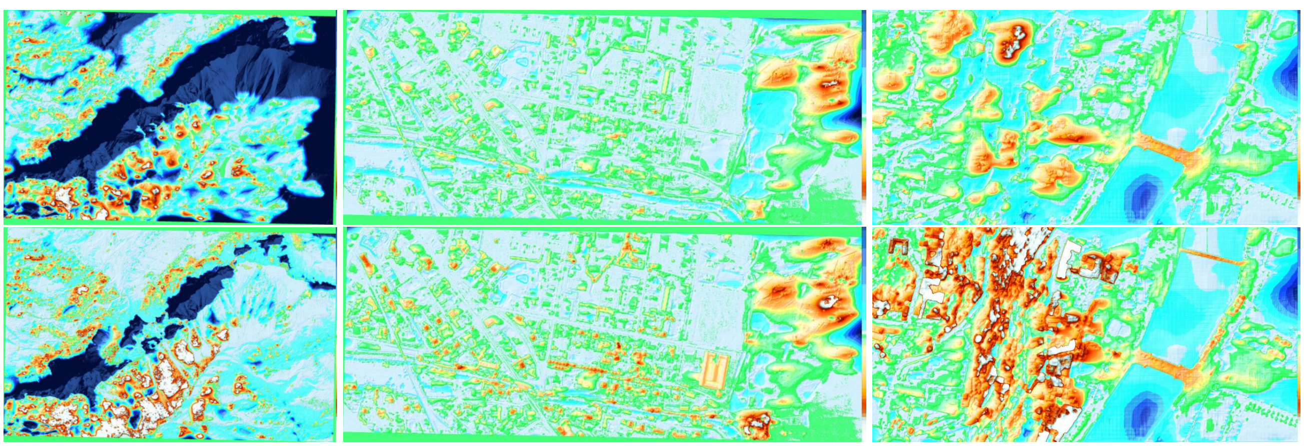


Figure 8. Comparison of difference images derived for the proposed watershed-algorithm, slope for region growing of ground areas: top row: 10 %, bottom row: 100 %

In the third test area Lyon no DTM ground truth was available. So we selected – based on the good experiences from the other test regions – the results of the AI algorithm as reference. The results for this test region are shown in fig. 9. The classical approach fails in such dense urban areas due to too large building complexes. Using a filter radius of 200 m instead of the default 100 m gives much better results in this case (3rd column in fig. 9. The results of the terra algorithm look really good while with the proposed watershed algorithm some areas shown in green are about 20 to 70 cm too high. Since in this algorithm the bridge is included in the "ground" the surface of the bridge is 9 m above the other algorithms due to the fact that they return the water level below the bridge. As can be seen all non AI-methods show a hole of about 7 m in the lower part of the river. This is an due to error in the ai result.

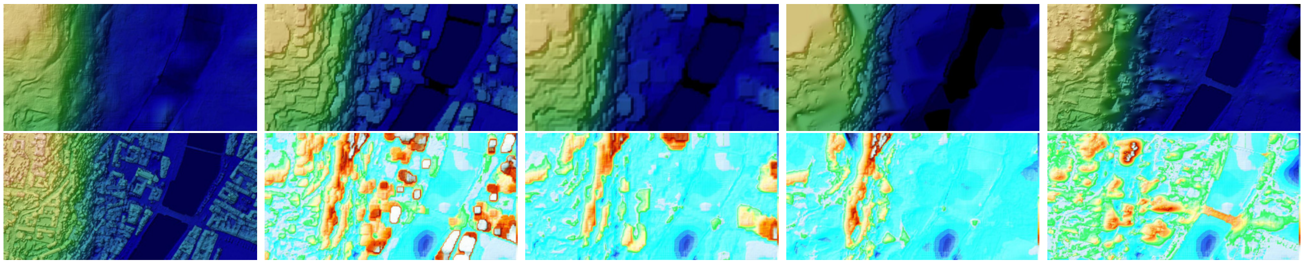


Figure 9. Derived DTMs for test region Lyon, top row l.t.r.: DTM ai, DTM op, DTM op200, DTM ter, DTM wst, bottom row l.t.r.: original DSM, differences of DTM op, op200, ter and wst to ai, color range from blue (DTM is 20 m lower than reference) to white (DTM is 20 m higher than reference, see also color bars on the right edges)

4.2 Comparison of nDEMs

Since for the generation of urban models the main issue is the detection of urban objects above ground we perform a special evaluation for this purpose. In this second type of evaluation we compare the resulting building masks. In fact they are high-objects masks since they also contains trees, walls and other high objects on the ground. But for simplicity we call them "building masks". First a normalized digital elevation model (nDEM) is calculated as $nDEM = \max(0, DSM - DTM)$. Afterwards the building-mask is calculated by thresholding the nDEM with a height threshold of 3 m.

These building masks are now compared to the building mask of the reference nDEM and IoU (intersection over union) masks and values are calculated. Fig. 10 shows the resulting IoU masks for the test region FriRefCity. White is the intersection of the buildings masks of the reference and the tested method. Green are false positives – i.e. areas which are marked as buildings in the nDEM of the tested method but did not exist in the reference. Red are false negatives – i.e. areas which are marked as high objects in the reference nDEM but not found by the method. For this test region the classic method "op" delivers the same results independent of the selected radius (100 or 200 m), so only one is shown.

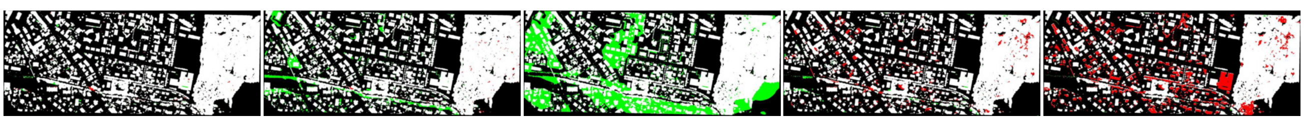
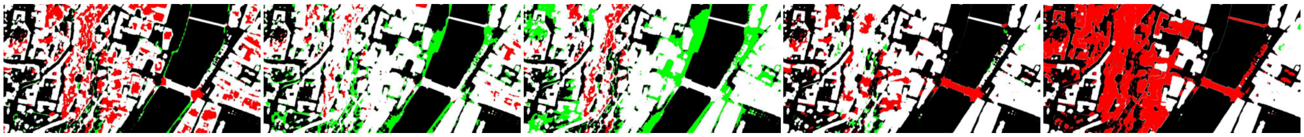


Figure 10. Intersections over Unions for test area FriRefCity (wst1 is wst with 100 % slope)

Fig. 11 shows the IoU masks for the test region Lyon. The reference is taken from the AI results as in the previous evaluations for this test region. At this test area we can see a major influence of the choosen radius (100 or 200 m) for the classic method "op". While the small radius misses most of the large buildings the larger radius increases the IoU significantly from 73.3 % to 83.0 %. As discussed earlier the proposed WST method marks also the bridges as ground. So the bridges which are marked as high objects in the reference and the other algorithms are classified as ground with the newly proposed method. For this the bridges are marked as "false negative" (high object not found) in the wst result. Considering this effect the IoU for the proposed WST method in the Lyon test area raises from 85.3 % to 87.4 %. Using the 100-%-slope for ground filling in the proposed algorithm ("wst1") makes most of the buildings and trees part of the ground due to the smoothed out borders of high objects in DSMs derived from satellite imagery and thus delivers a huge amount of false negatives and a very bad IoU of only 47.4 %. As can also be clearly seen in the IoU maps is the missing ground areas near the walled river which results in large false positives in the classic and terra methods.

4.3 Comparison of profiles

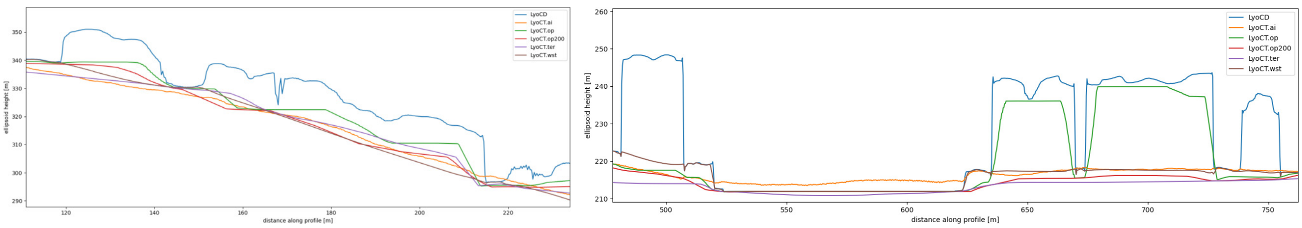
Fig. 12, left shows a 120 m long cross-section along the hill over the old town in Lyon. The DSM is shown in blue, all other profiles are DTMs from the tested methods. As can be seen the proposed method shown in dark brown fits best the supposed ground in this region. It corresponds well with the road areas cutting through the



op $IoU = 73.3$ % op200 $IoU = 83.0$ % ter $IoU = 74.6$ % wst $IoU = 85.3$ % wst1 $IoU = 47.4$ %

Figure 11. Intersections over Unions for test area Lyon (wst1 is wst with 100 % slope)

forests and interpolates smoothly the forests between. The classical method with a small radius gives too high values while using the large radius gives too small heights. The terra method works very well in this case since it was designed for landscapes of this type. The AI algorithm works also very well but it over- and underestimates the real positions of the roads.



hill river

Figure 12. Profile along the hill slope and across the Saône in test area Lyon

In fig. 12, right a profile across the Saône is shown. As can be clearly seen the classical algorithm with the small radius of 100 m misses the larger buildings (green) while using the large radius of 200 m smooths out the river edges more (red). The proposed method (brown) shows the best behaviour in retrieving the grounds between and below the buildings and also on the river and it's walled edges. The normally comparable or better AI method (orange) shows in this case a worse reconstruction of the river and it's edges while the terra method (purple) underestimates the ground for the whole profile shown.

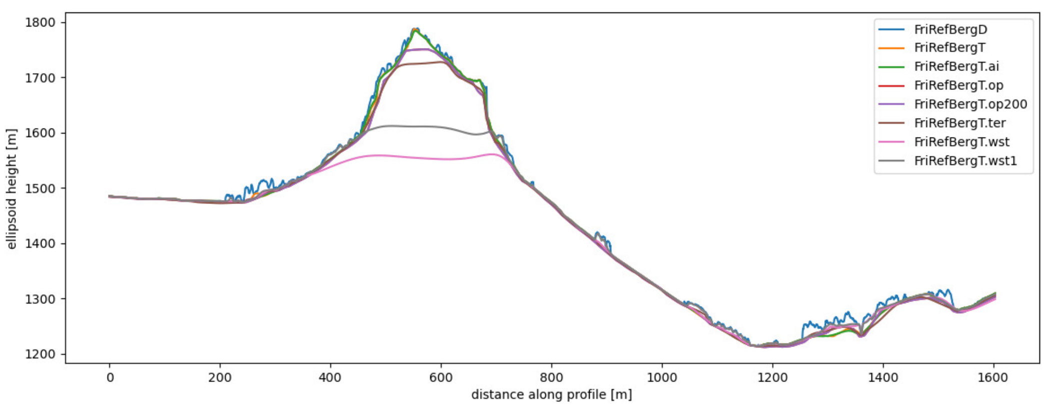


Figure 13. Profile across the mountain range in test area FriRefBerg

Fig. 13 shows a profile of the DSM and all resulting DTMs across the mountain range. As can be seen the proposed method shown in purple cuts off the mountain completely since it finds only ground areas near the mid and bottom slopes of the mountain. Increasing the slope for the region growing of the ground areas from 10 % to 100 % results in the dark gray profile (wst1) which is better but still not comparable to the other classic methods "op" (purple) and "terra" (brown) or even the – in this case – nearly perfect AI method.

5. CONCLUSION AND OUTLOOK

In the presented paper we proposed a new approach for the derivation of digital terrain models (DTMs) from digital surface models (DSMs) which can be created from airborne or satellite optical stereo imagery. We compared the results of the proposed method with an also recently developed AI algorithm and also with some other existing "classical" – i.e. non-AI – approaches.

The proposed method performs in urban test areas similar to the AI method and better than all existing classic methods. In mountainous areas the new method does not work very well due to the steep and smooth slopes of the hills. Also the classic methods can not model the mountain DTMs correctly but better than the new method.

Especially in urban areas containing small hills with buildings and rivers the method performs best in comparison to all classic methods due to the improved detection of ground areas. Especially the road levels above walled rivers and even bridges over the rivers are classified correctly as ground areas.

Future work should deal with large non-ground regions like on mountains and in forests which could actually not be detected correctly.

REFERENCES

[1] Li, Z., Zhu, Q., and Gold, C., [*Digital Terrain Modeling - Principles and Methodology*], Routledge Chapman and Hall (11 2004).

[2] Krauß, T., Arefi, H., and Reinartz, P., "Evaluation of selected methods for extracting digital terrain models from satellite born digital surface models in urban areas," in [*International Conference on Sensors and Models in Photogrammetry and Remote Sensing (SMPR 2011)*], *SMPR* (5 2011).

[3] Weidner, U. and Förstner, W., "Towards automatic building extraction from high resolution digital elevation models," *ISPRS* **50 (4)**, 38–49 (1995).

[4] Arefi, H., Engels, J., Hahn, M., and Mayer, H., "Automatic DTM generation from laser-scanning data in residential hilly area," *Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* **36**(4) (2007).

[5] Krauß, T. and Reinartz, P., "Urban object detection using a fusion approach of dense urban digital surface models and VHR optical satellite stereo data," in [*ISPRS Istanbul Workshop 2010, WG I/4 Modeling of optical airborne and space borne sensors*], *ISPRS* **39** (10 2010).

[6] Perko, R., Raggam, H., Gutjahr, K., and Schardt, M., "Advanced DTM Generation from Very High Resolution Satellite Stereo Images," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **II-3/W4** (2015).

[7] Pijla, A., Bailly, J.-S., Feurer, D., Maaoui, M. A. E., Boussema, M. R., and Tarolli, P., "TERRA: Terrain Extraction from elevation Rasters through Repetitive Anisotropic filtering," *Int J Appl Earth Obs Geoinformation* **84**(2020) (2020).

[8] Bittner, K., Zorzi, S., Krauß, T., and dAngelo, P., "DSM2DTM: An End-to-End Deep Learning Aproach for Digital Terrain Model Generation," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* **2023**, 1–9 (9 2023).

[9] Beucher, S., "Watersheds of Functions and Picture Segmentation," in [*IEEE International Conference on Acoustics, Speech and Signal Processing*], *IEEE International Conference on Acoustics, Speech and Signal Processing* , 1928–1931 (1982).