DERIVING A DTM FROM A DSM BY UNIFORM REGIONS AND CONTEXT

Charles Beumier, and Mahamadou Idrissa

Royal Military Academy, CISS Department, Brussels, Belgium;
{beumier / idrissa}(at)elec.rma.ac.be

ABSTRACT

A Digital Terrain Model (DTM) is typically used in applications such as environment planning, flood risk evaluation and building detection. A Digital Surface Model (DSM) is usually acquired by LASER scanning or from stereoscopic pairs of aerial or satellite images. By contrast, a direct DTM acquisition traditionally requires more effort on the terrain or for scene object classification and the automatic derivation of a DTM from a DSM still faces difficulties.

We propose in this paper a new DTM from DSM approach that consists of three steps: DSM region segmentation, region selection and height interpolation. First the DSM is segmented into regions of limited slope. A gradient filter is applied to the DSM raster to highlight height transitions. A connected component algorithm labels the different regions separated by those transitions. Secondly, regions whose perimeter is on the average higher than their surrounding are discarded. Finally, a hierarchical interpolation procedure fills the holes in the DSM due to high gradients or discarded regions.

The proposed algorithm developed for urban areas has been applied to the Vaihingen dataset of the ISPRS benchmark and qualitatively validated by the results of its independent evaluation procedure for building detection. In the context of change detection for database revision, this new DTM extraction was applied to an area in Brussels for which digital aerial imagery and LIDAR measurements are available. This reference data allowed for some quantitative evaluation of the DTM errors.

INTRODUCTION

A Digital Terrain Model (DTM) is typically used in applications such as environment planning, flood risk evaluation, image ortho-rectification for map creation and building detection, to name a few. A common approach for building detection exploits the difference between a DSM (Digital Surface Model) and a DTM. This difference, called normalized DSM (nDSM = DSM-DTM), is the local elevation which can be thresholded to detect building candidates.

A DSM is usually acquired from LIDAR measurements or by photogrammetry with aerial or satellite stereo imagery. By contrast, DTM acquisition traditionally requires more effort due to the very slow procedure of terrain survey or the need for classification between ground and non-ground pixels. The derivation of a DTM from a DSM is attractive but still faces difficulties related to the quality of the DSM (1), the ambiguity of some scene objects (like bridges) and the difficulty to classify off-terrain points (2).

To derive a DTM from a DSM, off-terrain pixels have to be filtered out. A large bibliographic source is available from the LIDAR ground filtering research. In the review of Meng et al. (2), 6 categories of methods are identified. The book chapter written by Pfeifer about LIDAR data filtering and DTM generation (3) is more synthetic and lists 4 categories (morphological filters, progressive densification, surface-based filters and segmentation-based filters).

In our previous development (4) the DTM value assigned to a pixel was derived from the percentile 10 of the histogram of DSM values in a square neighbourhood of 50×50 m^2 around the pixel. This method corresponding to the morphological category gave satisfying results in case of small size buildings or woods and where the terrain is rather flat.

We propose in this paper a new DTM from DSM approach that consists of three steps: DSM region segmentation, region selection and height interpolation, all detailed in the next section. Several

papers described approaches based on segmentation, category 4 of Pfeifer's classification. Pérez-Garcia et al. (5) used the distance to a local plane to segment and guide ground point densification. Yuan et al. (6) exploited an intensity image to perform region growing. Yan et al. (7) segmented the DSM with a local slope constraint. Regions higher than their neighbours are filtered out. Compared to these works, our approach does not require seeds for DSM segmentation. All the regions are detected in two DSM raster scans rather than individually and successively for each seed.

The method was designed for urban areas with high resolution imagery. It was qualitatively evaluated for building detection in the context of the ISPRS test on urban classification. A more quantitative evaluation was performed in the framework of our cooperation with the Belgian National Geographical Institute (NGI) where the whole process of DSM acquisition and DTM derivation was compared against LIDAR measures, owned by CIRB (Centre d'Informatique pour la Région Bruxelloise).

METHODOLOGY

To derive a DTM from a DSM, ground pixels have to be identified. Automatic pixel classification into terrain and non-terrain is not obvious if no land cover database is available. Vegetation is a sensitive issue, as the distinction between grass (terrain) and trees (non-terrain) is often ambiguous and is further complicated by the recent trend to add vegetation on roofs.
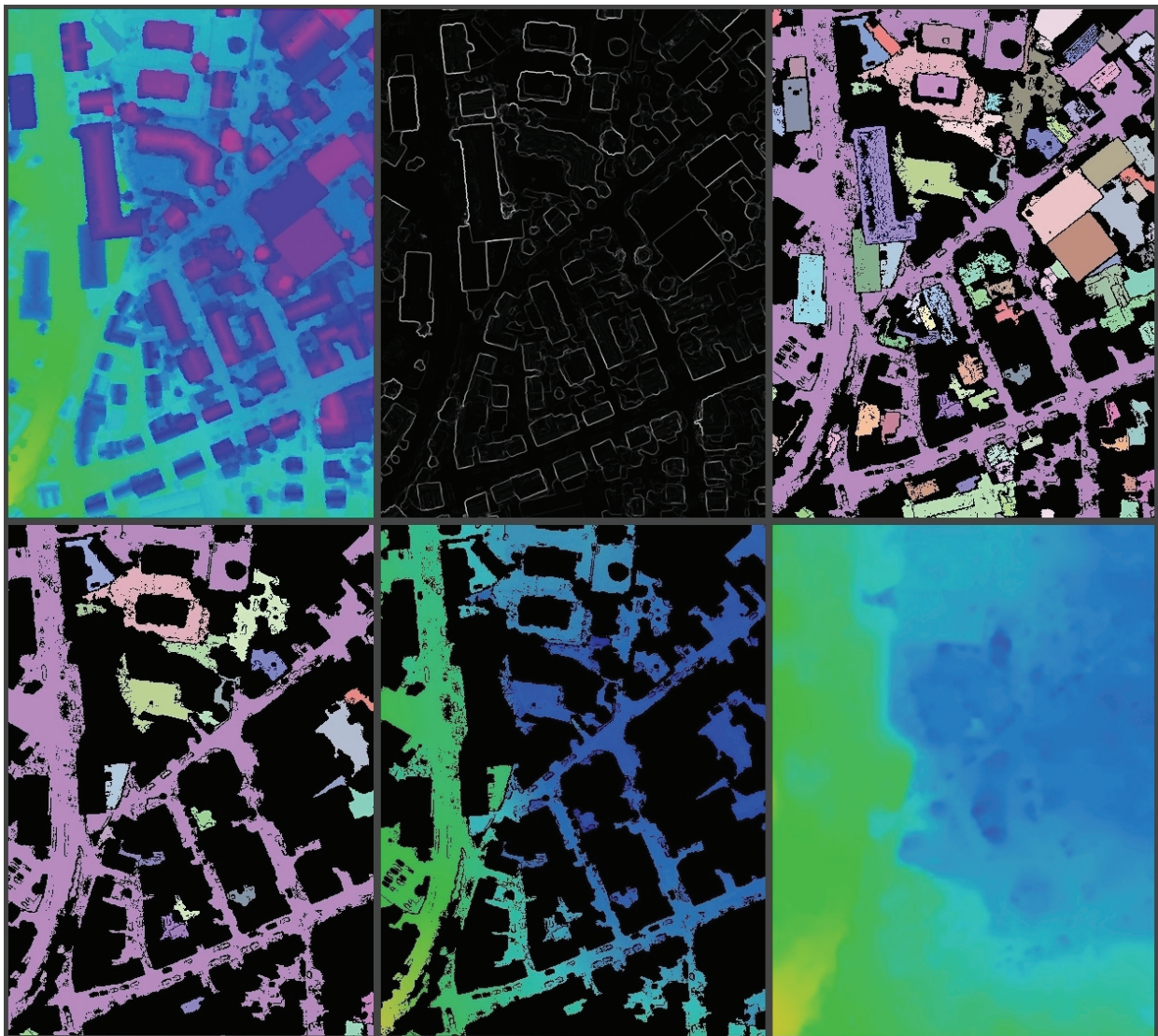

Figure 1: DTM from DSM: DSM for Vaihingen data (a); DSM gradient (b); Large regions with small gradients (c); Regions, except those locally elevated (d); Sparse DTM from DSM values (e); Full DTM from interpolation (f).

For the above reasons, we gave preference to geometrical rules. The constraints on size, slope and neighbouring height differences are easy to derive from the sole DSM and geometrical rules are intuitive and implemented with understandable parameters.

Our new DTM from DSM approach consists of three steps: DSM segmentation into uniform regions, selection of terrain regions and interpolation between terrain regions.

DSM segmentation

The DSM is first segmented into regions of limited slope. The gradient of the DSM raster is computed to highlight height transitions (Figure 1b). The gradient threshold is typically around 0.4 (25° from horizontal) as a compromise to get sufficiently large segments while limiting the variation within each of them.

A two-pass connected-component labelling algorithm, similar to the one described and analysed in Walczyk et al. (8) for real-time video processing, identifies the different regions separated by the transitions (Figure 1c). This labelling algorithm is particularly fast, as it only scans twice the DSM raster with very few pixel accesses (globally 9 per pixel). Since they are likely to be less reliable, small regions are filtered out thanks to the label pixel counts obtained in one raster scan.

Region filtering

Secondly, the regions identified in the first step whose perimeter is on the average higher than their surrounding are discarded.

A global image transform is performed to highlight valid regions. The difference between the original DSM and a blurred version (with a 4×4 m^2 box filter) reveals large differences D at region borders (Figure 2b), where the DSM slope is strong by construction (see previous subsection). The sign of the difference is dependent on the relative height of the neighbouring regions. Each region is assigned a count PD for positive (D > 2 m) and a count ND for negative (D < -2 m) pixel difference D. A region is rejected, if its count PD exceeds half of its count ND. Figure 1d shows kept regions, corresponding mainly to the road network, squares and gardens.
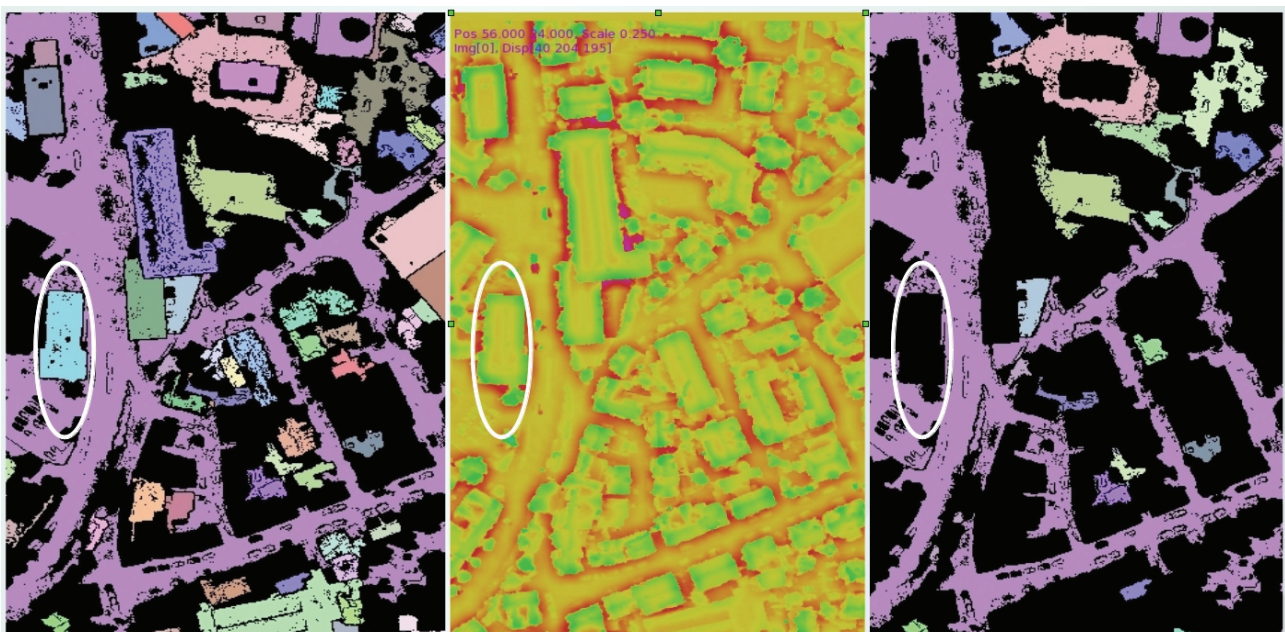

Figure 2: Region filtering. Uniform regions, in false colours (a); DSM transformation to highlight higher (in green) and lower (in red) border regions (b); Lower regions of DSM (c).

Height interpolation

Finally, a hierarchical interpolation procedure fills empty areas corresponding to high gradients or discarded regions.

The DTM values are first initialised with DSM values in retained regions (Figure 1e). This DTM with holes is used to construct a multi-resolution pyramid with a scale decrement equal to 2, thanks to a recursive lowpass filtering and subsampling to create each level (Figure 3a). The recursion stops when a level with no hole is created. Then the pyramid is traversed in the other direction, towards the higher resolution. Starting from the penultimate lowest resolution, pixels with no DTM value receive the interpolated values from opposite direct neighbours (8-connectivity) at the same level or, if there are none, from the corresponding neighbours in the directly lower resolution. The holes in the DTM are thus progressively filled from lower to higher resolutions, until the initial DSM resolution is reached.
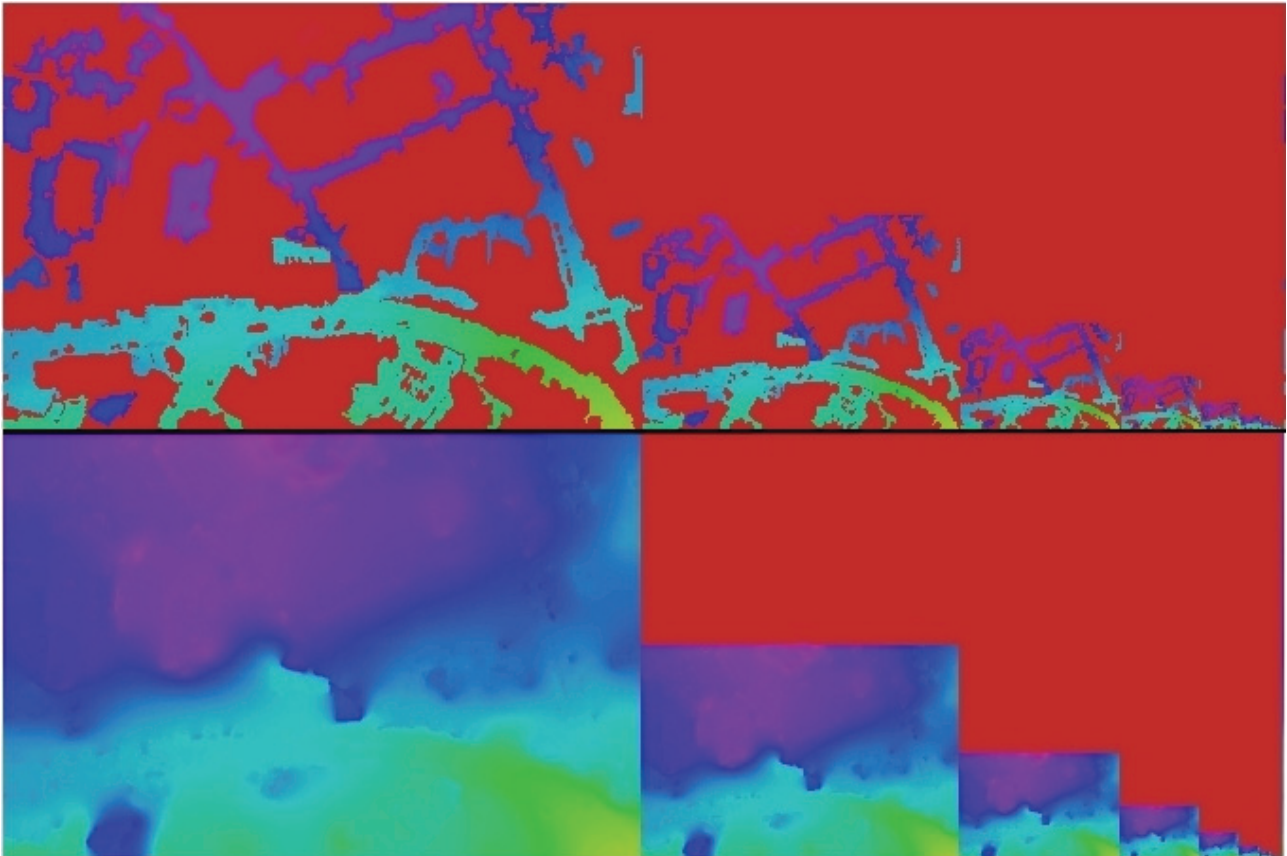

Figure 3: Image pyramid for DTM interpolation. Pyramid construction towards lower resolution (a); Pyramid filling from lower resolution (b).

This DTM from DSM approach is particularly well suited to very high resolution imagery (0.1m), since the segmented regions are sufficiently large. In 1 m imagery, small objects like little streets, terraces and small gardens have little chance to be retained as ground surfaces. However, even at these scales, small objects like cars or garden huts may be erroneously kept in the DTM so that post filtering could improve results. These artefacts are of limited amplitude and normally do not penalize applications such as building detection.

RESULTS

The proposed algorithm was applied to the Vaihingen dataset of the ISPRS benchmark and was qualitatively validated by the results of its independent evaluation procedure for building detection.

In the context of change detection for database revision for the Belgian National Geographic Institute, this new DTM extraction was applied to digital aerial imagery of Brussels (GSD=7.5cm), for which ground truth values from LIDAR are available. This reference data allowed for some quantitative evaluation of the DTM errors.

Vaihingen benchmark

The data used for development and test consists of the Vaihingen set of the ISPRS benchmark (9). This set contains CIR aerial images (Red, Green and Near Infrared channels), a derived DSM map and an ortho-rectified CIR image with 9 cm spatial resolution. A DSM was also produced by laser scanning at 25 cm resolution, but we have not considered this source. We tested our own DSM generation tool, taking advantage of the 11-bit image data and multiple views.

Thanks to the DSM, the CIR images were geometrically corrected (ortho-rectified). Then the Normalized Difference Vegetation Index was computed at each pixel thanks to the red (R) and near infrared (NIR) channels by $NDVI = (NIR\text{-}R)/(NIR\text{+}R)$. The vegetation mask, obtained by pixels for which NDVI > 0.2, was adapted to deal with shadow pixels, for which vegetation is usually underestimated due to the reduced spectral dynamic (10).

A DTM was derived from the DSM by the presented procedure. The segmentation with gradient threshold at 0.4 generated regions with a good balance between spread and uniformity. Their minimum size was set to 50 m^2 to reject a few off-terrain regions occurring to be lower than their neighbours. The Vaihingen areas contain enough large and middle-sized terrain regions to populate a sparse DTM for proper interpolation.

The building mask for the ISPRS benchmark was obtained as pixels out of the vegetation mask, where the DSM exceeds DTM values by at least 2.0 m. The building masks for the three Vaihingen study areas were analysed by the ISPRS benchmark procedure to deliver quality measures in terms of completeness and correctness and a quality factor balancing both criteria.

Area1 is situated in the city centre and consists of buildings with rather complex shapes and some trees. Area2 has a few high residential buildings surrounded by trees. Area3 is a residential area with small detached houses. The quality measures are given for the three areas for all participants in web page tables (11). This benchmark, further detailed in (12), confirmed our confidence in the qualitative correctness of our DSM to DTM approach for building detection in the difficult case of urban areas. Table 1 shows an excerpt of these results that situate the performance of our building detection development in the middle of the range of 26 other results (11).

Table 1: Quantitative results of building detection per pixel for Area1, Area2 and Area3.

Evaluation method	Area1	Area2	Area3
Per-area Completeness	91.6%	95.4%	91.3%
Per-area Correctness	92.4%	85.9%	92.4%
Per-area Quality	85.2%	82.6%	84.9%

Brussels data

Estimating the correctness of a DTM derived from a DSM is quite tricky, since errors may emanate from different origins like inaccurate DSM values, wrong filtering when deriving the DTM and inaccuracies of the reference values. Usually, small differences from the reference DTM are polluted by large discrepancies related to DSM errors and erroneous filtering so that global measures like the standard deviation is no valid accuracy measure. We computed the Normalized Median Absolute Difference (*NMAD*) as motivated in the work of Höhle et al. (13) about accuracy assessment of digital elevation models and we localised large errors by observation. We neglected the possible planimetric (*XY*) errors, but they should have limited impact, since DTM surfaces are smooth and mainly with limited slope in urban areas.

A DTM was derived from the DSM by the proposed approach with a slope limit equal to 0.4. The minimal region size was set to 400 m2 as many city blocks include small off-terrain regions that are nonetheless lower than their neighbourhood (garage, house annexes).

First, we observed the difference (Figure 4a) between the DTM produced by the new approach and a LIDAR raster (DSM) created from LIDAR points. Since the LIDAR density is 30 points per m^2, we adopted a planimetric resolution of 0.2 m and adapted the DTM consequently. In Figure 4a, red pixels correspond to points with Z difference in the [-1m..+1m] range, while white pixels, having larger differences, normally relate to buildings and trees, present in the LIDAR raster but not in the DTM. Pixels with no LIDAR value are black. The figure shows a globally correct quality in line with building detection. Roads and buildings are well separated. Trees along the streets are also highlighted. Most building blocks have inside areas marked as ground (in red).
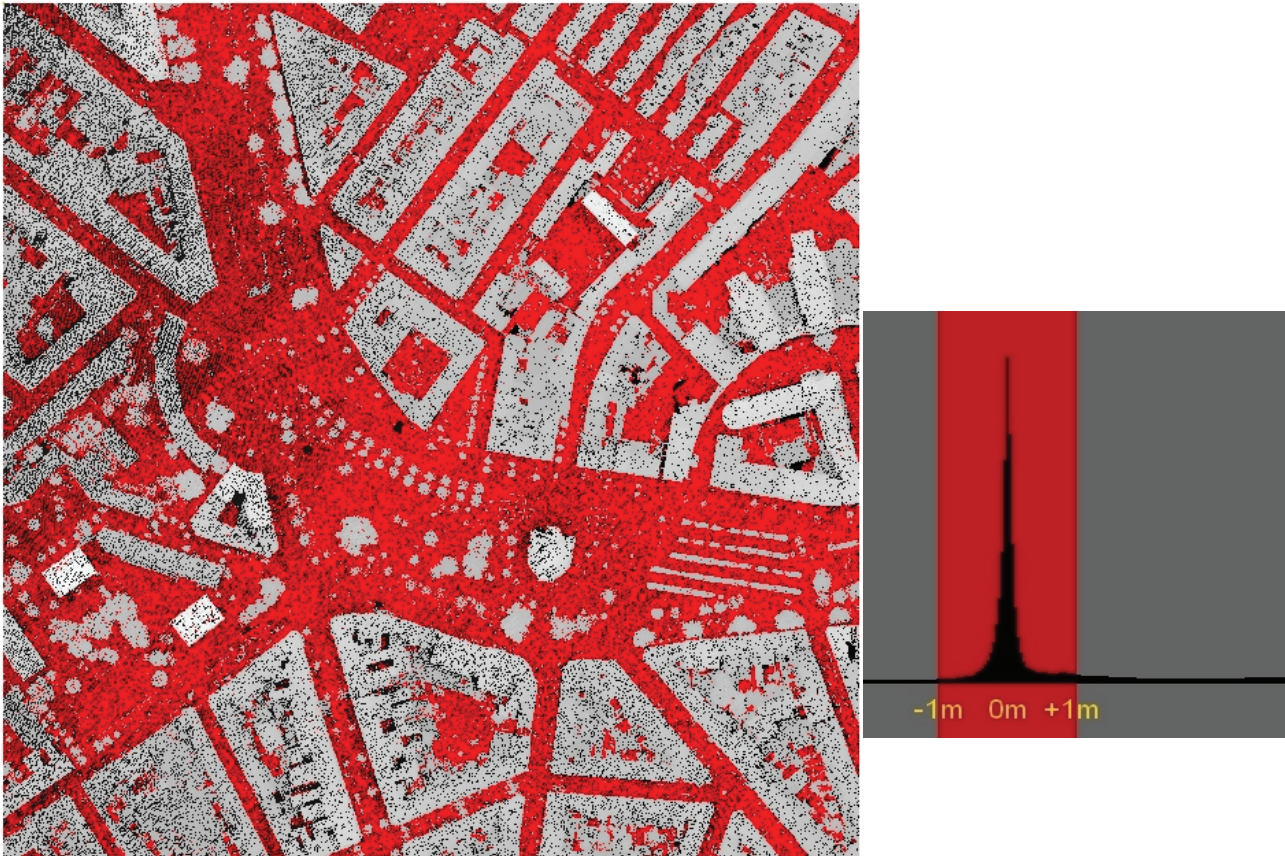

Figure 4: Difference LIDAR-DTM in Brussels area: Raster, [-1m..+1m] in red (a); Distribution (b).

Next, we considered the difference between our proposed DTM and a reference DTM purchased by the CIRB. In Figure 5, corresponding to 0.5 km^2, the zones with deviation larger than +1 m (green) or -1 m (red) account for 7% of the total area. The zones with deviation larger than +/- 2 m, accounting for 2% of the total area, originate from a difference in city blocks either due to erroneous ground points or to interpolation in areas lacking ground points.

Green circles relate to missed ground regions, under the level of surrounding streets. Most of these regions were rejected by the minimum area condition, which aimed at avoiding off-terrain regions falsely classified as ground points, since they are surrounded by higher regions. In dark green circles, the difference is very local (e.g. pit for ventilation), but got extent by interpolation. In the red circle a region above the street level was detected but rejected by the minimum area condition.

In some situations, ground points were missing in the DSM due to difficult stereo matching conditions. The errors in blue circles relate to missed small ground parts between high and near buildings.

The magenta circle indicates a strong but local difference between our photogrammetric DTM and the LIDAR DTM due to a tunnel. Other difficult cases may arise from elevated terraces or bridges wrongly classified, if they are not smoothly connected to a ground part. We found in the test area only one obvious example of erroneous ground estimation in the reference DTM. It is circled orange, at the bottom of the image.
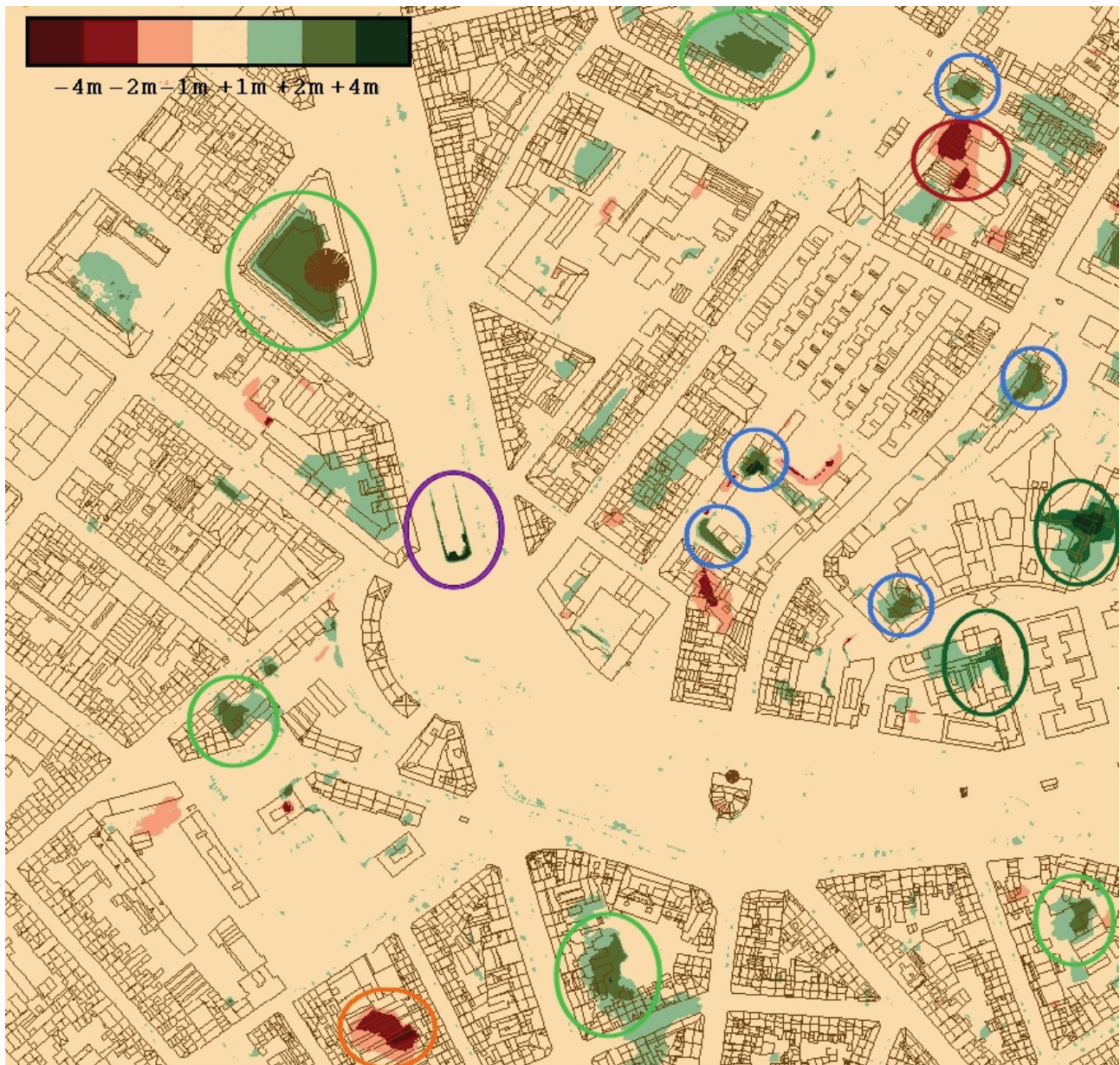
Figure 5: Difference in false colours between our DTM and the reference based on LIDAR data.

Finally, the histogram of the difference between the extracted DTM and the reference DTM was analysed. The distribution is similar to the one of Figure 4b except that the tail to the far right, corresponding to buildings and trees, is not so important. The sharp peak at 0 supports the idea that many pixels of the derived DTM are close to the reference. The distribution in [-1 m..1 m] led to a NMAD of 0.22 m. This corresponds to 3 times the ground sampling distance of the stereo imagery from which our DTM values were derived. This accuracy is sufficient for applications with a threshold around 2 m, like building detection. The influence of local DTM errors is spread by the interpolation process in built-up areas. A much better consistency of DTM values is encountered in the open areas and on road segments.

Discussion

The proposed DSM to DTM approach was applied to two different urban types: mostly detached houses in the case of Vaihingen and mostly city blocks in the case of Brussels. The validity of the approach for building detection has been shown for Vaihingen and for Brussels (Figure 4). The comparison with a LIDAR DTM has highlighted differences in city blocks that required higher values of the minimum region size parameter to filter out small off-terrain regions falsely taken as ground, as they are surrounded by higher neighbours.

Apart from the possible improvement on DSM generation, we plan to first extract a DTM based on very large low regions (typically the road network) and refine this DTM by integrating smaller regions close in elevation to the first DTM approximation.

The different image processing tasks have been designed for time performance, since the solution is likely to be run on very large images. A 10 Megapixel DSM is processed in about 12 s on an i7 CPU at 2.67 Ghz using 1 core, 8 GB RAM.

CONCLUSIONS

This paper presents an original approach for DTM estimation from a DSM. It segments the DSM into regions with limited slope thanks to connected-component labelling applied to the DSM gradient and rejects the regions clearly higher than their neighbours. The DSM values of the remaining regions form a sparse DTM that is completed by a hierarchical interpolation procedure.

The DTM is quickly derived and with a quality dependent on the DSM correctness. It was developed for urban areas with very high image resolution (around 0.1 m), since it relies on uniform regions like the road network. It was applied successfully to normalise the DSM for the detection and localisation of building candidates.

The algorithm encounters difficulties in large areas with no ground pixel (forest, very large building blocks) and in other specific situations like the interior of city blocks with a low off-terrain structure falsely taken as ground when surrounded by higher objects. Some DSM errors may be rejected by the approach, but they need in general additional information to be properly dealt with. In this respect one possibility is to fuse filtered DSM values with other elevation sources (old DTM values or elevation from a vector database). We plan to adapt the method to first concentrate on very large regions likely to be the ground (like the road network) and add smaller regions with elevation similar to the surrounding large regions.

ACKNOWLEDGEMENTS

REFERENCES

1   Nex F & M Gerke, 2014. Photogrammetric DSM denoising. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol XL-3: 231-238

2   Meng X, N Currit & K Zhao, 2010. Ground filtering algorithms for airborne LiDAR data: A review of critical issues. Remote Sensing, 31(2): 833-860

3   Pfeifer N, 2009. LiDAR data filtering and DTM generation. In: Topographic Laser Ranging and Scanning: Principle and Processing, edited by Jie Shan and Charles K. Toth, CRC/Taylor & Francis: 309-333

4   Beumier C, 2008. Building verification from disparity of contour points. In: Image Processing Theory, Tools & Applications, IPTA 2008 (Sousse, Tunisia) IEEE, 6 pp.
DOI: 10.1109/IPTA.2008.4743771

5   Pérez-Garcia J L, J Delgado, J Cardenal, C Colomo & M A Ureña, 2012. Progressive densification and region growing methods for LIDAR data classification. International Archives of Photogrammetry and Remote Sensing, Vol. XXXIX-B3: 155-160

6   Yuan F, J.X Zhang, L Zhang & J Gao, 2009. DEM generation from airborne LIDAR data. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol XXXVIII-7/C4: 308-312

7   Yan M, T Blaschke, Y Liu & L Wu, 2012. An object-based analysis filtering algorithm for airborne laser scanning. International Journal of Remote Sensing, 33(22): 7099-7116

8   Walczyk R, A Armitage & T D Binnie, 2010. Comparative study on connected component labeling algorithms for embedded video processing systems. The 2010 International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV 2010 (Las Vegas, Nevada, USA) 7 pp.

9   ISPRS_test, 2013. ISPRS Test Project on Urban Classification and 3D Building Reconstruction (last date accessed: 28 Jan 2015)

10 Beumier C & M Idrissa, 2014. Building detection with multi-view colour infrared imagery. EARSeL eProceedings, 13(2): 77-84

11 ISPRS_results, 2013. Results of the ISPRS Test Project on Urban Classification and 3D Building Reconstruction (last date accessed: 28 Jan 2015)

12 Rottensteiner F, G Sohn, M Gerke, J Wagner, U Breitkopf & J Jung, 2014. Results of the ISPRS benchmark on urban object detection and 3D building reconstruction. ISPRS Journal of Photogrammetry and Remote Sensing, 93: 256-271

13 Höhle J & M Höhle, 2009. Accuracy assessment of digital elevation models by means of robust statistical methods. ISPRS Journal of Photogrammetry and Remote Sensing, 64: 398-406

14 Cramer M, 2010. The DGPF test on digital aerial camera evaluation – overview and test design. PFG, 2010(2): 73-82; DOI: 10.1127/1432-8364/2010/0041