PFG 2012 / 4, 0317–0330
Stuttgart, August 2012

Article

Digital Surface Model Extraction and Refinement through Image Segmentation – Application to the ISPRS Benchmark Stereo Dataset

Daniela Poli & Pierre Soille, Ispra (VA), Italy

Keywords: high resolution, data fusion, digital surface models, segmentation

Summary: This paper proposes a methodology for the geometric refinement of a digital surface model (DSM) using high or very high resolution satellite scenes through an advanced hierarchical image segmentation and shows the results obtained on a dataset of Catalonia, Spain, provided by ISPRS WG I/4 through the project "Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data". During this experiment a WorldView-1 quasi-nadir scene was used for the enhancement of the DSM generated by a Cartosat-1 stereopair. The original and final DSMs are compared to the lidar DSM on the same area for quality analysis.

Zusammenfassung: Erstellung von Oberflächenmodellen und deren Verbesserung durch hierarchische Bildsegmentierung am Beispiel der ISPRS Benchmark-Stereodaten. Dieser Artikel stellt eine Methode zur geometrischen Verbesserung eines digitalen Oberflächenmodells (DSM) durch hierarchische Bildsegmentierung von hoch bzw. sehr hoch auflösenden Satellitenbildern vor. Die Methode wurde an einem Datensatz aus Katalonien in Spanien getestet, der im Rahmen des ISPRS WG I/4 Projektes „Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data" zur Verfügung gestellt wurde. In diesem Experiment wurde ein aus einem Cartosat-1-Stereopaar erstelltes DSM durch eine WorldView-1 Quasi-Nadir-Szene verbessert. Zur Bewertung der Methode wurde ein Lidar DSM als Referenz verwendet. Der Artikel beschreibt die Methode sowie die Ergebnisse des Tests und der Qualitätsanalyse.

1 Introduction

In the last years the number of Earth-observation platforms equipped with sensors delivering very high resolution (VHR) images with ground sample distances (GSD) smaller than 1 m has increased. Those sensors are characterized by different geometric, radiometric, spectral and operational specifications, and in most cases allow for the acquisition of multi-angle imagery for automatic digital surface models (DSM) generation and 3D feature extraction. One of the tasks of the Joint Research Centre of the European Commission in Ispra (Italy) is to assess changes in land cover after events like natural disasters or conflicts. VHR satellite images are generally used for this scope, as they almost cover any area of the planet, in particular remotely located areas. To describe the pre-event and post-event scenarios the available data are analysed and eventually fused in order to achieve the most accurate and reliable information. With respect to automatic 3D information extraction, the availability of accurate and detailed DSMs is a crucial issue for automatic building detection and subsequent damage estimation. While the acquisition of VHR stereo scenes can be planned after the event and used to model the post-event 3D scenario, VHR stereo scenes are generally not available for data acquisition before the event; on the other hand, single

www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0120
1432-8364/12/0120 $ 3.25

VHR scenes are likely to be available in the archives, as well as DSMs at medium or low resolution from multi-line optical sensors with simultaneous along-track stereo acquisition, like SPOT-5/HRS, Cartosat-1, ALOS-PRISM, or other sources. In this context we developed a strategy for data fusion, in order to enhance the surface models using information from single scenes. It is enhanced in the sense that the final DSM is sharper and therefore likely to be more suitable for subsequent 3D shape detection. We tested the method on the dataset provided by ISPRS Working Group I/4 within the project "Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data" (Reinartz et al. 2010). The approach and first results were presented at the Workshop "High-Resolution Earth Imaging for Geospatial Information" held in Hannover (Germany) in June 2011 (Poli & Soille 2011); in this work the latest developments in the methodology and the achieved results are reported and discussed.

2 Methodology

2.1 *Related Works*

The combined use of digital imagery and surface models mainly finds applications in automatic object extraction and 3D reconstruction from aerial data or very high resolution satellite sensors (Baltsavias et al. 1995, Jaynes et al. 1996, Paparoditis et al. 1998, Lu et al. 2002, Tao & Yasuoka 2002, Lu et al. 2006, Li et al. 2008, Arefi 2009). Few works aim at using satellite imagery for DSM enhancement. In Maire (2010), user-defined semantic contents (sea, lakes, buildings, roads, etc.) are extracted using a supervised classification in high resolution satellite imagery; then in correspondence of each class the surface is modelled by plane surfaces with geometric constraints given by the topological properties of each class and neighbour regions. This approach was applied, for example, for the enhancement of SRTM-X using a 2.5 m ground resolution SPOT-5 scene. In the computer vision community, several matching approaches employ image segmentation during the generation of the DSM (Klaus et al. 2006, Krauss & Reinartz 2010). For example, in the approach proposed by Krauss & Reinartz (2010), one stereo image is segmented and transferred to the disparity map, then for each segment the original disparity map is filled with suitable interpolation of the disparities. In the proposed method, segmentation of a VHR scene is used to refine a given DSM at a coarser resolution.

2.2 *Proposed Method*

The principal steps of the proposed approach are summarized in Fig. 1. The initial DSM can be any surface model generated by matching techniques from stereo images or other sources. Unless already orthorectified, the VHR scene is oriented using ground control information and orthorectified on the initial DSM. The generated orthophoto is segmented with alpha-omega connectivity (Soille 2008) and the segments are overlaid to the existing DSM. After calculating the statistics (mean, median, sigma, minimum and maximum values) of the heights of the points falling into each segment, the new surface model is calculated from the existing one by imposing that the height values that fall into the same segment follow a certain mathematic function, i.e. constant value or planar surface. In the initial version of the algorithm, the new height values for each segment were calculated as the minimum, mean or maximum values of the original DSM heights, or as the result of planar fitting.

A more advanced approach for the calculation of the enhanced height values was then developed. It is based on exploiting not only the radiometric information of the VHR scene, but also the information contained in the DSM shape. The aim is to enhance the surface model using the maximum value in segments deemed to correspond to a building, and using the minimum value for segments supposed to match ground areas between buildings. Buildings are generally represented in the DSM as convex surfaces while ground areas between buildings are likely to match concave surfaces (the term convex and concave are used in the sense "convex from above" and "concave from above", respectively). Therefore, the original DSM is segmented to extract concave, convex

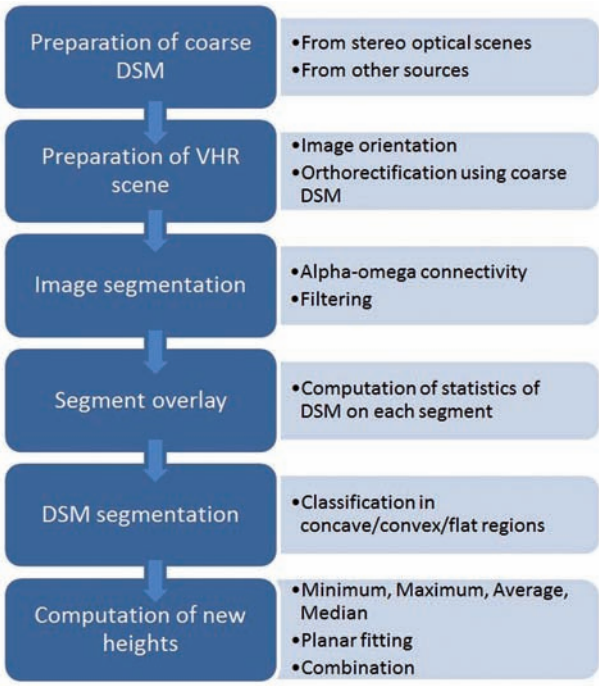


Fig. 1: Proposed workflow for DSM enhancement.

and flat regions, and then the new height values are computed differently, depending on whether the segment falls in a concave, convex or flat region. The decision rule is:

- Use minimum value if the segment falls in a concave region;
- Use maximum value if the segment falls in a convex region;
- Use median value or planar fitting if the segment falls in a flat region.

In practice, to exclude possible outliers, the 10 % (resp. 90 %) percentile is used instead of the minimum (resp. maximum) value. In the next paragraphs the algorithms used for image and DSM segmentation are described.

2.3 *Image Segmentation by Constrained Connectivity*

Given a logical predicate, the segmentation of an image is defined as the partition of the image definition domain into non-overlapping regions called segments, such that the logical predicate returns true on every segment but false on the union of any pair of adjacent segments (Horowitz & Pavlidis 1976). Recently, a segmentation technique producing a hierarchical partitioning of the image definition domain under connectivity constraints corresponding to logical predicates was proposed (Soille 2008). Given the input constraints, the resulting partition is uniquely defined, contrary to most segmentation techniques. By increasing the threshold levels associated to each connectivity constraint, one obtains a series of fine-to-coarse partitions with the pixels at the lowest level and the whole image definition domain at the highest level of the hierarchy.

Constrained connectivity relies on a dissimilarity measure computed between each pair of adjacent pixels. The absolute difference between the intensity values was used for all experiments hereafter (see Soille 2011 for other dissimilarity measures). Two pixels are said to be alpha-connected if there exists a path linking these two pixels in such a way that the dissimilarity between all pairs of adjacent pixels of the path does not exceed the

value of alpha. The relation 'to be alpha-connected' is an equivalence relation (reflexive, symmetric, and transitive relation) and therefore it leads to a unique partition of the image definition domain into regions of maximal extent called alpha-connected components.

The constrained connected component of a given pixel is then defined as the largest alpha-connected component of this pixel that satisfies a series of constraints. Any number of constraints may be considered. Given a local dissimilarity defined as the absolute difference, the most natural constraint is defined in terms of a threshold level on the range of the grey level values of the connected components. This constraint is called the omega-constraint and its associated threshold value is the omega threshold. Denoting by α and ω the values of the alpha and omega thresholds, respectively, the (α,ω)-connected component of a pixel is defined as the largest α'-connected component such that α' is less than or equal to α and the range of the α'-connected component is less than or equal to ω (Soille 2008). The omega constraint prevents linking through transitions, a well-known problem of single-linkage clustering.

A typical choice is to consider the same values for the alpha and omega thresholds. By increasing this value, a fine-to-coarse hierarchy of connected components is obtained. In practice, it is desirable to obtain segmentation as coarse as possible but retaining all structures of interest in the image given the considered application. In this study, since we are interested in buildings and other man-made structures, the parameters need to be selected in such a way that these objects are simplified as much as possible, i.e. ideally are matched by one connected component, but not merged with other objects.

To further favour linking within homogeneous regions while preventing it through transitions, an image pre-processed by the edge sharpening technique described in (Soille 2011) is applied before segmentation. The pre-processing consists in considering as seeds all local maxima of the input image and propagating their values so as to cover the whole image definition domain. The propagation is controlled by the original image values (the smaller the difference between a seed and a non-seeded adjacent point, the greater the speed). Alternatively, rather than pre-processing the image, the dissimilarity measure proposed in (Soille 2011) could be used to prevent linking through transitions while favouring it within homogeneous regions.

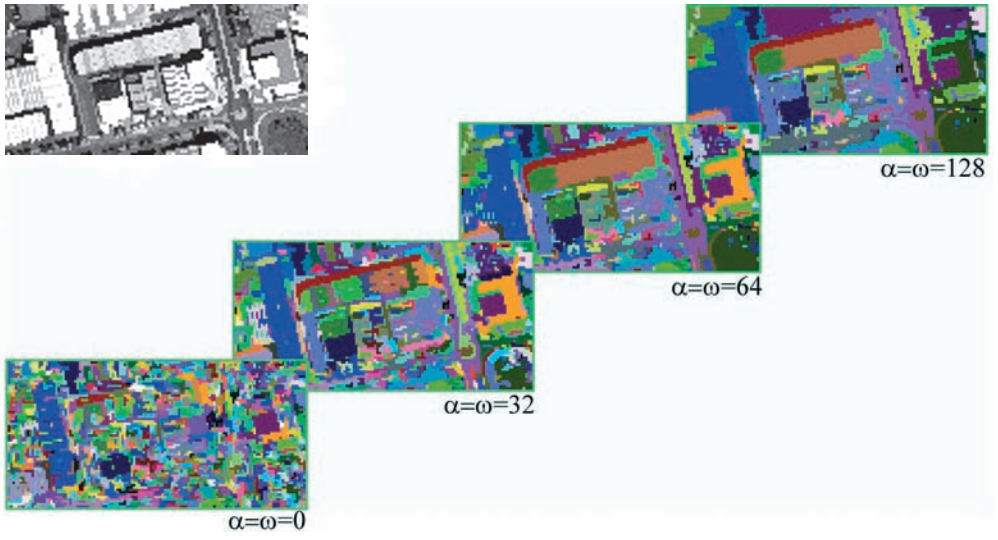


Fig. 2: Example of constrained connectivity partitioning of a 512 × 256 sample of a WorldView-1 image resampled at 2.5 m, rescaled to byte data type, and contrast enhanced. The partitions are displayed for the alpha and omega thresholds values equal to 0, 32, 64, and 128 respectively (from fine to coarse partitions).

An example of constrained connectivity partitioning is shown in Fig. 2. In this figure, the original image is partitioned with increasing values of alpha and omega parameters. To highlight the hierarchical property of the resulting partitions, random colours were given to the connected components of the finest partition while the colour of a given connected component of a subsequent level is inherited from the largest connected component of the previous level, contained by this connected component.

2.4 *DSM Segmentation*

The approach used for DSM segmentation and identification of concave, convex and flat regions was proposed by Pesaresi & Benediktsson (2001) and applied to satellite images. The method is a morphological segmentation by the derivative of the morphological profile, based on the use of residuals from opening and closing (levelling) by reconstruction.

A structure or an "object" in an image is defined as a connected component of pixels sharing the same morphological characteristics, i.e. "flat", "concave", "convex" in case of the local curvature of the grey level function surface. The residuals between any opening by reconstruction (or closing by reconstruction) and their original function are interpreted as a measure of the relative brightness of the structure (or relative darkness). Then, one membership function can be defined corresponding to the class "convex" and another membership function corresponding to the class "concave". The levelling algorithm is rewritten as a decision rule based on the greatest value of the membership function. The segmented image of the characteristic is a tessellation of three different labels like "convex", "concave" and "flat". For this segmented image, pixels where the lower levelling is strictly lower than the original image are labelled "convex", and pixels where the upper levelling is strictly greater than the grey function are labelled as "concave". Pixels labelled "flat" have maintained the same value of the grey function in both the upper and lower levelling. Therefore, those pixels have been indifferent to the erosion/dilation-reconstruction process with a given structuring element. In experiments with satellite images, the proposed method demonstrated good performance in terms of shape description and retained small but significant regions in images.

3 Experimental Results

In our experiment the algorithms described in section 2 were used for the enhancement of a DSM generated from a stereopair acquired by Cartosat-1 (C1) sensor using a WorldView-1 (WV1) scene. In the next paragraphs the data and detailed processing will be presented and discussed.

3.1 *Data Description*

The Working Group 4 of Commission I on "Geometric and Radiometric Modelling of Optical Spaceborne Sensors" is providing on its website several stereo datasets from high and very high resolution spaceborne stereo sensors on three areas in Catalonia, Spain, covering urban, rural, and forest areas in flat and medium undulated terrain as well as steep mountainous terrain. In addition to these datasets, a lidar DSM generated by the Institut Cartogràfic de Catalunya (ICC) is provided as a reference for comparison (Reinartz et al. 2010). Among the three available regions (Terrassa, La Mora, Vacarisses), the dataset located over Terrassa was chosen for our experiments.

The dataset consists of:

- a subset of a stereopair acquired by WV1 on August 28, 2008, composed by a nadir (-1.3º) and off-nadir (33.9º) scene, with ground sample distances of 0.50 m and 0.76 m, respectively; both scenes are panchromatic and have size 10,000 x 10,000 pixels; rational polynomial coefficients (RPC) are available; for our experiment only the nadir scene is used;
- a stereopair acquired by C1 in February 2008, composed of a backward (-5º) and forward (33.9º) looking scene, with ground sample distance of 2.5 m; both scenes are panchromatic; RPCs are available;
- ground coordinates of points;

- a dense 3D point cloud acquired on November 27, 2007 with airborne lidar.

The area is characterized by urban, rural and forest cover on flat and hilly terrain (Reinartz et al. 2010). The terrain elevation ranges from 200 m to 430 m.

As explained in the introduction, in this paper we propose an approach to refine a coarse resolution DSM using a single VHR scene, as generally VHR stereopairs are not available to describe pre-event scenarios. Therefore, even if a VHR stereopair is available, from the Terrassa dataset we used only the C1 stereopair to generate a DSM and the nadir WV1 scene to refine it. The processing workflow includes the following steps:

- orientation of C1 stereo scenes,
- DSM generation at 2.5 m grid spacing (DSM_C1),
- orientation of WV1 nadir scene,
- orthorectification of WV1 scene using DSM_C1 at ground sample distance of 2.5 m,
- segmentation of the WV1 orthoimage,
- enhancement of DSM_C1 with generation of a new DSM at 2.5 m grid spacing.

3.2 *Orthorectification and DSM Generation*

For the photogrammetric processing of C1 and WV1 scenes the commercial software SATPP (SATellite image Precision Processing) by 4DiXplorer AG was used (Zhang 2005).

The C1 stereopair was oriented by estimating the correction of the available RPC through an affine transformation using nine ground control points (GCPs). Then a DSM was generated using the advanced approach in SAT-PP based on a coarse-to-fine hierarchical solution with an effective combination of several image matching methods and automatic quality indication (Zhang & Grün 2004). To improve the conditions for feature extraction and matching, the images were preprocessed using Wallis filter. Few seed points were measured in the epipolar images in stereo mode to provide the initial approximation of the surface. The DSM was generated with a regular grid space of 2.5 m (DSM_C1, Fig. 3). No further editing was applied on the surface model. In the DSM it is possible to distinguish urban areas, but buildings are not well defined from adjacent buildings or roads. In open areas, the surface model follows the terrain shape without outliers.

The nadir WV1 scene was georeferenced, with the estimation of affine correction parameters of the given RPCs by means of GCPs. This was followed by the orthorectification, using the DSM_C1 as elevation model. As a result an orthoimage at 2.5 m GSD was produced.

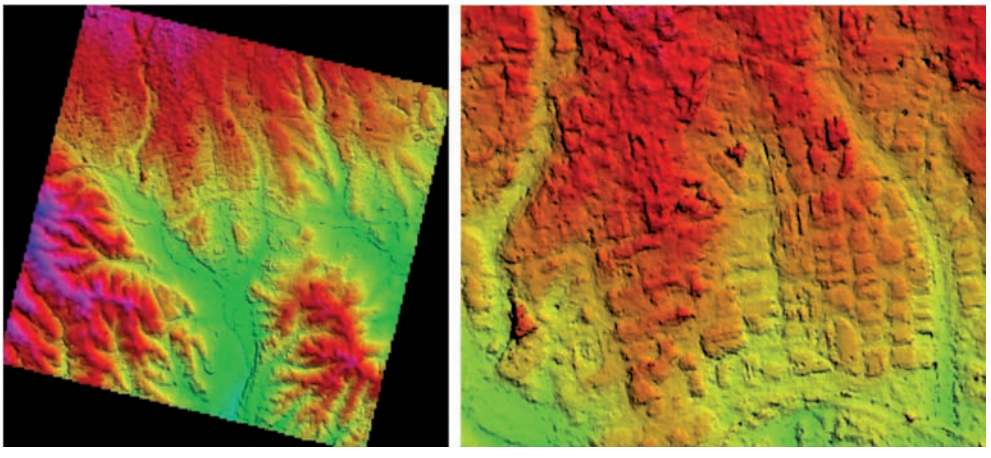
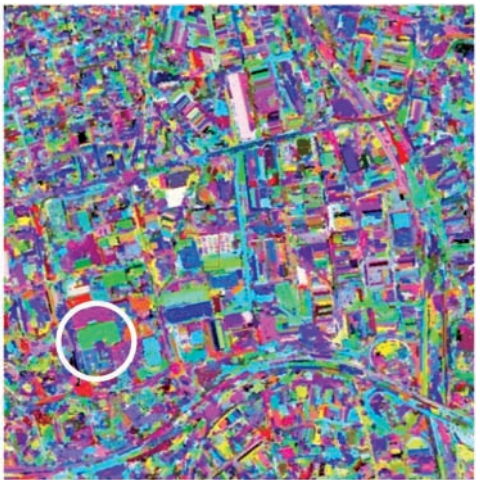


Fig. 3: Colour-shaded visualization of DSM generated from Cartosat-1 stereo scenes over Terrassa area. Overview of the full model (left) and zoom on urban and open areas (right).

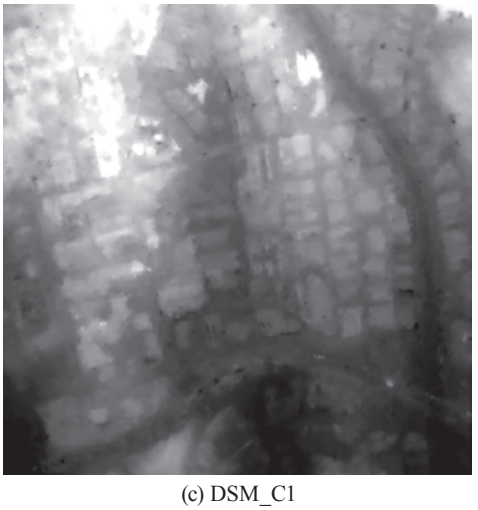
3.3 *Segmentation and DSM Enhancement*

The WV1 orthoimage was pre-processed with edge-sharpening and segmented. For the segmentation, a value of 50 intensity levels for both the alpha and omega thresholds was found to be optimal, in the sense that it enabled the formation of connected components matching the subparts of the building roofs while avoiding their merging with adjacent objects. An example is given in Fig. 4 (a) and (b) on a subset with buildings, main and secondary roads on different levels, a river with bridges, open areas with vegetation. Following the methodology previously described, four DSMs were calculated from the original DSM, using the minimum value (ENH_MIN), the maximum value (ENH_MAX), the mean value (ENH_AVG) of the heights in each segment, and planar fitting (ENH_PLA). From a visual comparison of the original DSM (Fig. 4 (c)) and the enhanced DSMs with average and planar fitting (Fig. 5 (a) and (b) respectively) on the subset mentioned above, it is evident that in the enhanced DSMs the building edg-

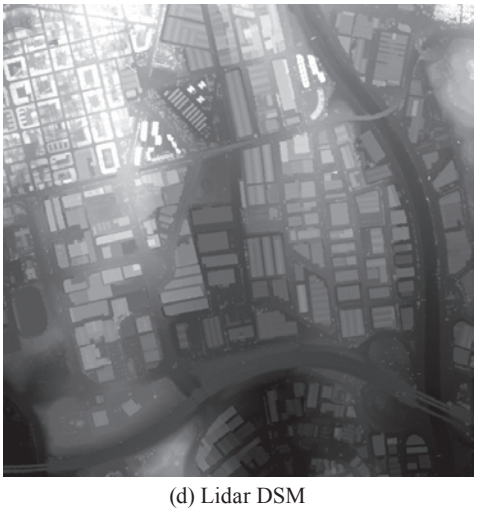
(a) orthophoto



(b) segmented image



(c) DSM_C1



(d) Lidar DSM

Fig. 4: Results of DSM enhancement over a subset. (a) WV1 orthoscene at 2.5 m grid spacing, (b) segmentation, using alpha = omega = 50 over a subset, (c) original Cartosat-1 DSM, (d) reference lidar DSM used for quality analysis.

es are more delineated and objects at different heights, like rivers and bridges, are better defined than in DSM_C1. In homogeneous areas (roof faces, roads, parking areas, fields, and so on) small details due to noise are removed. For the purposes of our research, we focused the analysis on buildings. With respect to the function used to enhance the DSM, there is not a significant difference between average and planar fitting at this resolution, as in DSM_C1 itself it is not possible to recognize shape details, like roof faces, with sufficient accuracy.

The enhanced DSMs obtained using the minimum and maximum values (ENH_MIN and ENH_MAX) were compared to the reference one and to DSM_C1 to evaluate if the original errors of the DSM_C1 were increasing or decreasing with the respect to the lidar after the refinement. As expected, we notice in Fig. 6 that by applying the minimum values, the errors in correspondence of ground level decreases, while using the maximum value the enhanced DSM improves in correspondence of buildings. Therefore the criteria should be combined. The original DSM was segmented in concave, convex and flat regions and the new DSM was calculated with minimum and maximum height values in segments falling in concave and convex regions respectively, and medium value elsewhere (ENH_HYB).

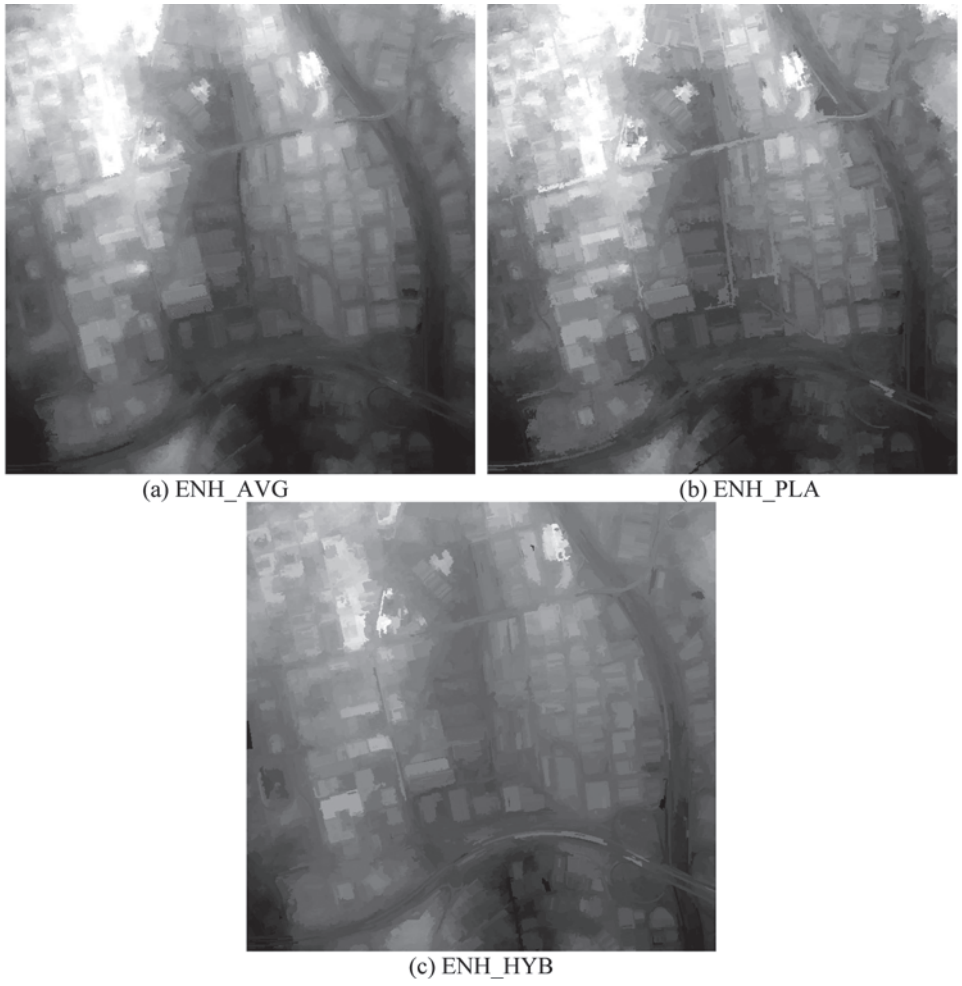


Fig. 5: Results of DSM enhancement over a subset using (a) average values, (b) planar fitting, (c) a combination of minimum, maximum and median values.

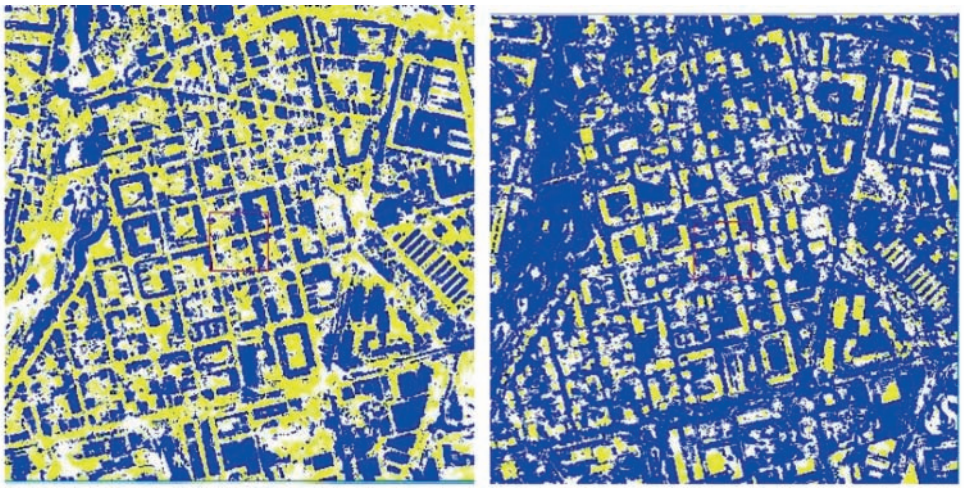


Fig. 6: Zoom on height gradients of ENH_MIN (left) and ENH_MAX (right): yellow = error decreases, blue = error increases, white = error does not change.

3.4 *Quality Analysis*

For the quality analysis of the enhanced DSM, the lidar DSM was used as reference (Fig. 4 (d)). We evaluated the height differences through the surface profiles in dense urban and industrial areas. Some examples are reported in Fig. 7.

With respect to the original DSM, the enhanced one produces step profiles in correspondence of buildings external edges, as in profile 2, and in some cases the algorithm detected buildings that were missing in the original DSM (profiles 1 and 5). The enhancement is not sensitive to the terrain slope (profile 4). Obviously the accuracy of the enhanced DSM depends on the accuracy of the initial DSM. Indeed small corridors between buildings are well defined in the lidar DSMs, but not in DSM_C1; as a consequence the ground height in the enhanced DSM is similar to the one in the original DSM, but the profiles of the buildings are improved (profiles 1, 2 and 5).

The presented method shows a limitation in case of complex buildings with homogeneous texture. With reference to profile 2, the building portions at different heights cannot be distinguished in the panchromatic channel (white circle in Fig. 4 (a)), resulting in a single segment after partitioning (white circle in Fig. 4 (b)), and a single height value for that segment in the enhanced DSM. A refining of the segmentation with additional information, like multispectral channels, could improve the identification of sub-structures of the roofs at different heights.

The statistics on the height differences between the original and enhanced DSMs with respect to the lidar one are summarized in Tab. 1. The mean error, the mean absolute

Tab. 1: Statistics of height differences between the original and refined DSM with respect to lidar along five transects (μ = mean value, MAD= mean absolute deviation, RMSE = root-mean-square error) (m).

	Profile 1			Profile 2			Profile 3			Profile 4			Profile 5		
	μ	MAD	RMSE	μ	MAD	RMSE	μ	MAD	RMSE	μ	MAD	RMSE	μ	MAD	RMSE
Original	-1.64	2.91	3.78	-2.00	3.37	4.15	0.70	2.41	3.25	-1.22	2.59	3.21	-4.21	4.86	5.88
Refined	-1.03	2.90	3.89	-1.18	2.08	2.77	0.40	1.90	1.87	-0.51	2.50	3.35	-2.87	4.06	4.92
# points		37				70			140			209			74

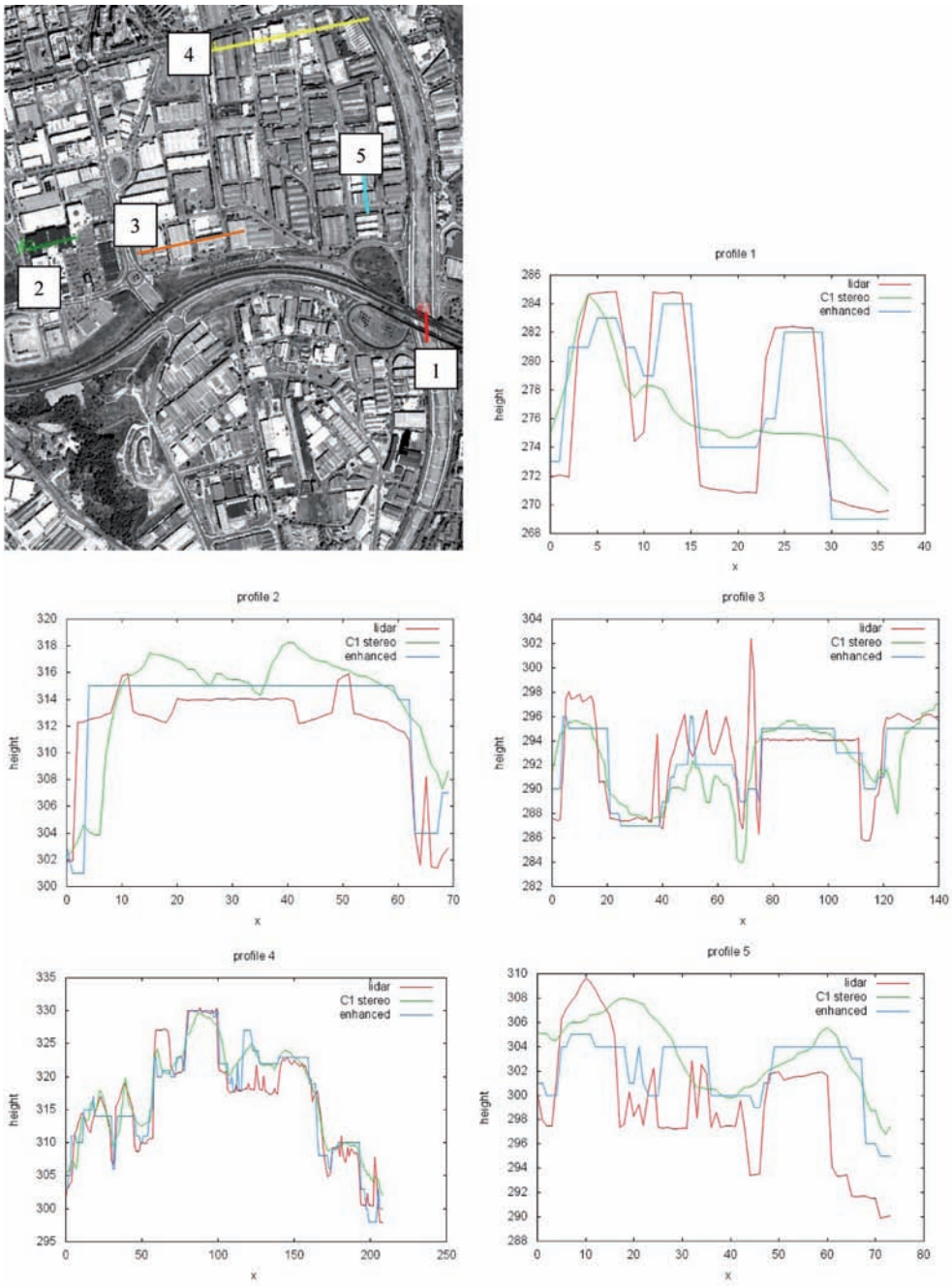


Fig. 7: Profiles of original C1 stereo DSM (green), lidar DSM (red) and ENH (blue) along the five transects shown in the upper left image.

Fig. 8: Perspective 3D visualization of original DSM (DSM_C1, left) and enhanced DSM (ENH_HYB, right), both with vertical exaggeration 2.0, texture with C1 orthoimage (2.5 m).

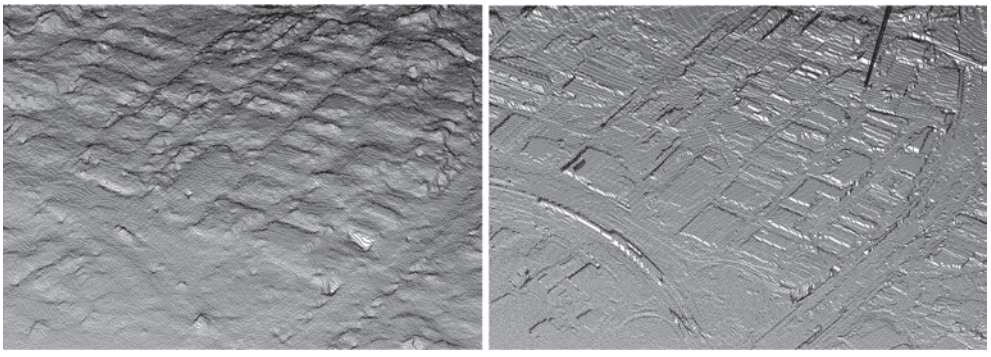


Fig. 9: Perspective 3D visualization in wireframe mode of original DSM (DSM_C1, left) and enhanced DSM (ENH_HYB, right), both with vertical exaggeration 2.0.

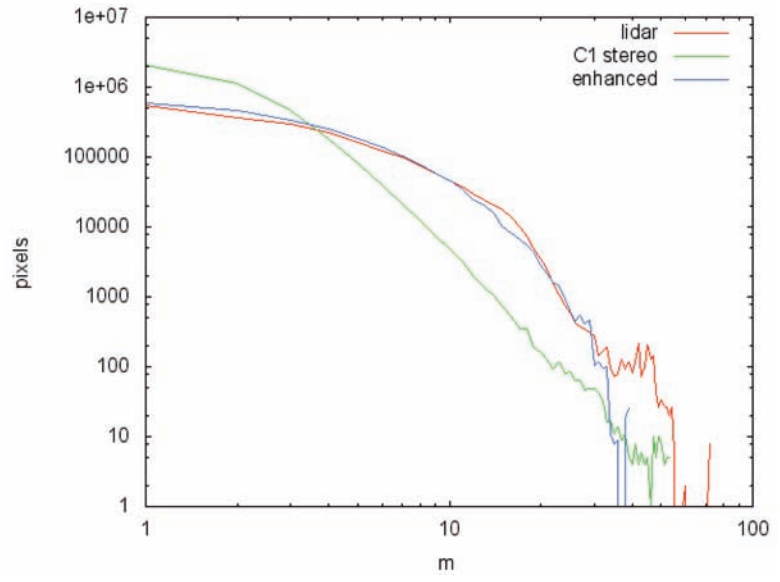


Fig. 10: Histogram of the local slope of lidar, original C1 stereo and enhanced DSM (logarithmic scale).

deviations and RMSE are decreasing for the refined DSMs, except for non-significant increase of the RMSE of profile 1.

The 3D visualizations of the original and enhanced DSMs in Figs. 8 and 9 show that after the refinement, small blobs at the ground level are removed and the building edges are sharper and more realistic than in the original DSM. This improvement is confirmed by the analysis of the distribution of the local slope, calculated as the difference between the maximum and minimum height within a 3 by 3 neighbourhood. This analysis is carried out over the lidar, C1 stereo, and enhanced DSMs, see Fig. 10. Indeed, the distribution of the slope of the original DSM gets substantially closer to that of the lidar after enhancement, apart from a few outliers at the tail of the distribution.

4 Conclusions

In this paper we presented an approach for the enhancement of digital surface models using very high resolution satellite images. After orthorectification, the images are segmented through an advanced hierarchical image partitioning, then the segments are projected on the existing DSM. The new height values are calculated for each segment with planar fitting or constant (average) value of the original DSM heights contained in that segment. The method was applied on a dataset over Catalonia, Spain, provided by ISPRS WG I/4 through the project "Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data". During this experiment a DSM was generated by a Cartosat-1 stereopair and a WorldView-1 (WV1) quasi-nadir scene was used for the DSM enhancement. The original and final DSMs were compared to the lidar DSM on the same area for quality analysis. The analysis of the profiles (with statistics), and local slopes and the 3D visualization shows that in the enhanced DSM buildings contours are sharper and more realistic than in the original one, which is favourable for automatic building detection. Further investigations will also include the use of multispectral information to improve the image segmentation and practical experiments of damage assessment using automatic building detection in the original and enhanced DSMs.

Acknowledgements

Special thanks are given to the data providers for the provision of the stereo datasets, namely: Euromap for the Cartosat-1 data, Digital Globe for the WorldView-1 data and ICC Catalunya for the reference data.

References

Arefi, H., 2009: From LIDAR Point Clouds to 3D Building Models. – PhD thesis, Bundeswehr University Munich.

Baltsavias, E.P., Mason, S. & Stallmann, D., 1995: Use of DTM/DSMs and Orthoimages to Support Building Extraction. – Gruen, A., Kuebler, O. & Agouris, P. (eds.): Automatic Extraction of Man-Made Objects from Aerial and Space Images. – Birkhäuser Verlag, Basel, 199–210.

Jaynes, C.O., Collins, R.T., Stolle, F.R., Hanson, A.R., Riseman, E.M. & Schultz, H.J., 1996: Three-Dimensional Grouping for Site Modeling from Aerial Images. – Arpa Image Understanding Workshop: 223–236.

Horowitz, S. & Pavlidis, T., 1976: Picture segmentation by a tree traversal algorithm. – Journal of the ACM **23** (2): 368–388.

Klaus, A., Sormann, M. & Karner, K., 2006: Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. – **18th** International Conference on Pattern Recognition 2006: 15–18.

Krauss, T. & Reinartz, P., 2010: Enhancement of dense urban digital surface models from VHR optical satellite stereo data by pre-segmentation and object detection. – The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences **XXXVIII** (1).

Li, Y., Zhu, L. & Shimamura, H., 2008: Integrated method of building extraction from digital surface model and imagery. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXVII** (B3b).

Lu, Y., Trinder, J. & Kubik, K., 2002: Automatic building extraction for 3D terrain reconstruction using interpretation techniques. – Techniques, School of Surveying and Spatial Information

Systems, University of New South Wales, NSW 2052 Australia.

Lu, Y., Trinder, J. & Kubik, K., 2006: Automatic Building Detection Using the Dempster-Shafer Algorithm. – Photogrammetric Engineering & Remote Sensing **72** (4): 395–403.

Maire, C., 2010: Image Information Extraction and Modeling for the Enhancement of Digital Elevation Models. – PhD Thesis, Karlsruhe Institute of Technology (KIT).

Paparoditis, N., Cord, M., Jordan, M. & Cocquerez, J.P., 1998: Building Detection and Reconstruction from Mid- and High-Resolution Aerial Imagery. – Computer Vision and Image Understanding **72** (2): 122–142.

Pesaresi, M. & Benediktsson, J.A., 2001: A new approach for the morphological segmentation of high-resolution satellite imagery. – IEEE Transactions on Geoscience and Remote Sensing **39** (2): 309–320.

Poli, D. & Soille, P., 2011: Refinement of digital surface models through constrained connectivity partitioning of optical imagery. – The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences **XXXVIII** (4/W19).

Reinartz, P., d'Angelo, P., Krauss, T., Poli, D., Jacobsen, K. & Buyuksalih, G., 2010: Benchmarking and quality analysis of DEM generated from high and very high resolution optical stereo satellite data. – The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences **XXXVIII** (1).

Soille, P., 2008: Constrained connectivity for hierarchical image partitioning and simplification. – IEEE Transactions on Pattern Analysis and Machine Intelligence **30** (7): 1132–1145.

Soille, P., 2011: Preventing chaining through transitions while favouring it within homogeneous regions. – ISMM 2011, Lecture Notes in Computer Science **6671:** 96–107.

Tao, G. & Yasuoka, Y., 2002: Combining high resolution satellite imagery and airborne laser scanning data for generating bare land DEM in urban areas. – The International Archives of Photogrammetry and Remote Sensing and Spatial Information Sciences **XXXIV** (5/W3), Kunming, China.

Zhang, L., 2005: Automatic Digital Surface Model (DSM) Generation from Linear Array Images. – PhD Dissertation, No. **88**, Institute of Geodesy and Photogrammetry, ETH Zurich, Switzerland.

Zhang, L. & Grün, A., 2004: Automatic DSM generation from linear array imagery data. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XXXV** (3): 128–133.

Address of the Authors:

Dr.-Ing. Daniela Poli, Dr.-Ing. habil. Pierre Soille, European Commission – Joint Research Centre, Institute for the Protection and Security of the Citizen, Global Security and Crisis Management Unit, Geo-Spatial Information Analysis for Global Security and Stability (ISFEREA) Action, Via E. Fermi 2749, I-21027 Ispra (VA) Italy, TP 267, Tel. +39-0332-789241, -785068, Fax +39-0332-785154, e-mail: first name.last name@jrc.ec.europa.eu

Manuskript eingereicht: September 2011
Angenommen: November 2011