

A Hybrid Approach for Delineation of Building Footprints From Space-borne Stereo Images

Gholam Reza Dini, Karsten Jacobsen, Franz Rottensteiner, Mehdi Ravanbakhsh, Paolo Gamba and Christian Heipke

I. INTRODUCTION

The automatic extraction of building footprints from remotely sensed images has been used for updating geospatial databases in urban areas [1]. The launch of High Resolution Spaceborne Stereo (HRSS) sensors (e.g. GeoEye, WorldView, QuickBird) started a new era by providing the possibility to obtain stereo images and 3D maps from space [2]. Indeed, building identification, reconstruction, and change detection have been carried out using stereo image matching, as well as 3D edge matching techniques [3,5-6]. As stated in [3], 3D edge matching based on stereo images delivers promising results, but only if the buildings are large enough with respect to the spatial resolution of the data, have a simple rectangular shape, and a good radiometric contrast compared to surrounding objects. As a matter of fact, although 3D edge matching using very high resolution aerial images can reconstruct building footprints in detail [7], using space-borne images the same approach may encounter issues, particularly where building outlines are not clearly detected in both epipolar images. Additionally, although image matching delivers a DSM representing buildings heights, building size and shapes extracted from this DSM are usually overestimated, so that auxiliary information is required.

In this work we focus on the use of stereo imagery and their fusion with information from a single radiometric data set. Specifically, we propose a hybrid approach, which integrates height and image information, and efficiently deals with complex building footprints. The proposed framework integrates the radiometric property (2D image segmentation) with the geometric property (DSM from dense image matching).

II. REGION-BASED SEGMENTATION USING LEVEL SET

For the delineation of building rooftops, rooftop segmentation based on active contours is used. In this method, it is of crucial importance to handle the changes of topology which may be caused by the curve evolution. Level sets based on the edge gradient function are usually unable to capture the

smooth boundaries that frequently characterize the outlines of buildings using coarse ground resolution of satellite images.



Fig. 1. Extraction of building footprint using (left): height threshold on nDSM derived from image matching, (right): segmentation of building rooftops over a GeoEye image using the level set approach

To solve this problem, [8] proposed a new method for the initialization of active contour model based on curve evolution and level sets [9]. This model basically assumes that buildings and background are characterized by rather homogeneous but different distributions of grey levels. Fig.1 shows the segmentation results for building rooftops in a GeoEye image initialized by the associated matching-based nDSM. As shown, due to the smoothing effect of image matching, enlarged building footprints are extracted from the DSM (See Fig. 1). This makes the extraction more complicated for the buildings with poor contrast with respect to the background. In such cases, not only the building size is overestimated in the DSM, but the building shape may be changed due to the potential mismatch. Hence, in the following image radiometric information is utilized to overcome the drawbacks by this smoothing effect, due in turn to image matching.

III. A HYBRID APPROACH FOR ROOFTOPS SEGMENTATION USING IMAGE AND HEIGHT INFORMATION

A. Data Sources and Pre-processing

The input for our method consists of a pair of high-resolution stereo satellite images having a ground sampling distance (GSD) of 1 m or smaller.

To transform the original images into epipolar geometry, stereo pairs are rotated around the viewing direction, so that the x -axes of both image coordinate systems are parallel to the base. For DSM generation, semi-global matching (SGM) [11] is employed. For the subsequent normalized digital surface model (nDSM) generation, first a digital elevation model (DEM) is generated by filtering the DSM, then the nDSM is computed as the point-by-point difference between DSM and DEM.

B. Building Rooftops Segmentation

In this section, a hybrid approach is introduced to delineate

G.R. Dini is currently with the Integrated Geospatial Innovations (IGI), Germany. This work was done while he was a PhD researcher at the IPI Hannover and then a visiting researcher at the University of Pavia, Italy (e-mail: gr.dini@igi-systems.com).

K. Jacobsen, F. Rottensteiner and C. Heipke are with the Institute for Photogrammetry and Geoinformation, Leibniz Universität Hannover, Germany (e-mail: {jacobsen, rottensteiner, heipke}@ipi.uni-hannover).

Mehdi Ravanbakhsh is with the University of Western Australia (e-mail: mehdi.ravanbakhsh@uwa.edu.au).

P. Gamba is with the Department of Electrical, Computer and Biomedical Engineering of the University of Pavia, Italy (e-mail: paolo.gamba@unipv.it).

building outlines using spectral and geometric (height) information. With respect to the problem discussed in the previous section, neither matching-based DSM nor original image information are individually able to match the requirements for building footprint extraction [12].

Fig. 2 demonstrates the general framework of the proposed approach. As shown, first building blobs from nDSM are used to initialize the segmentation, limiting the image analysis to built-up areas and impeding the segmentation of non-building objects. Instead, the region-based segmentation proposed by [9] for building footprint extraction is implemented.

The segmentation is performed on epipolar images using nDSM blobs as starting built-up regions. In this study 3D segmentation refers to level set image segmentation which is performed on DSM blobs. Once building footprints on epipolar images demonstrate an overlap more than 75%, that footprint is confirmed. However, only one of them which is close to the nadir point is selected for footprint extraction. Initial ROIs in the nDSM contain the pixels with the height values larger than θ_{DSM} and without any overlap with the removal masks introduced by [3]. This delivers a rough approximation of buildings blobs so the segmentation gets an accurate initialization (i.e. building blobs). The curve within a building rooftop then evolves from the initial region of interest outward if it finds the building outline.

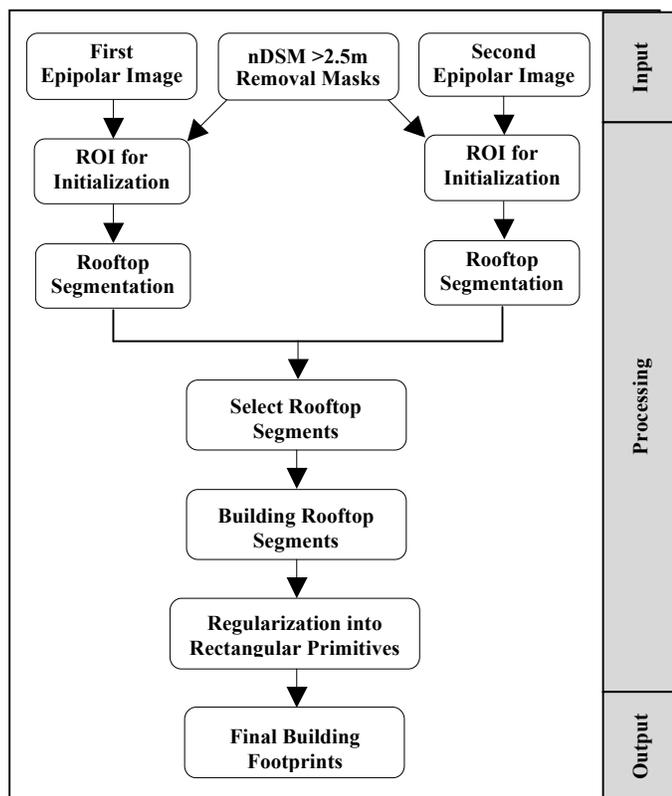


Fig. 2. A framework for the extraction of building footprints using 3D segmentation.

The segmentation starts from initial ROIs. It is iteratively expanded over homogenous building rooftops. In case the curve evolution reaches a building outline with a sufficient

sharp edge, the segment is not expanded further.

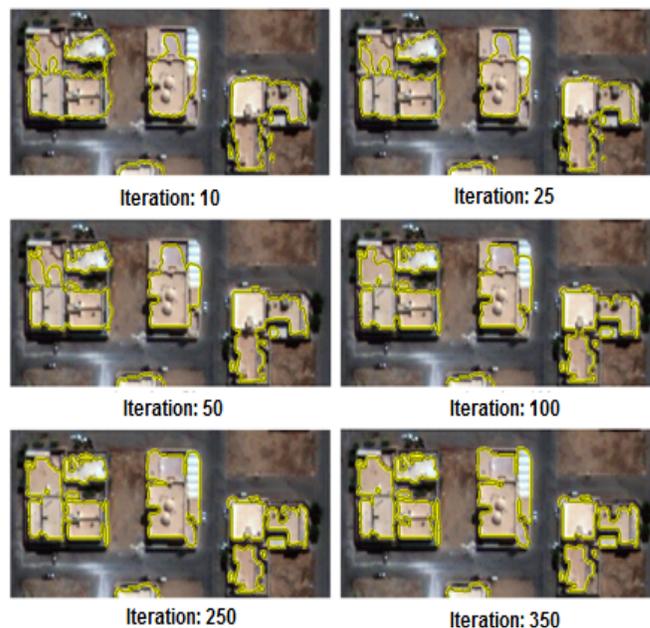


Fig. 3. Level set based segmentation of building rooftops demonstrating different number of iterations

A key point in rooftops segmentation is the estimation of an optimized number of iterations representing the best balance between segmentation accuracy and run time. The number of iterations depends on the building size as well as the complexity of building structure. In the next sections, the optimum number of iteration is investigated in order to delineate the building outline within the study area. Note that the number of iterations are most likely different for one area to another, and depend on the structure of building the image resolution and the scene complexity.

To prove the effectiveness of the methodology, in Fig.4 the results at multiple iterations for a subset of the test scene is provided. In Fig. 4 the accuracy of the extraction at each iteration is represented against the available ground truth. It is noted that, from the 1st to the 100th iteration, the overlap between the segmentation and ground truth sharply increases. After the 250th iteration, the overlap rate reaches its maximum value (91.15%).

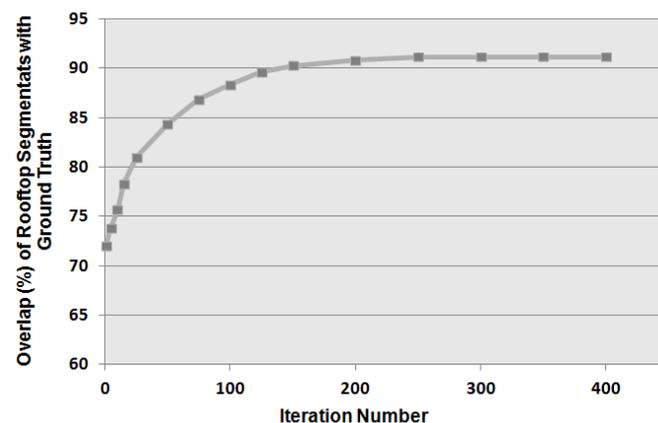


Fig. 4. Overlap (%) between rooftop segments and ground truth under the different number of iterations

C. Regularization of segments into rectangular primitives

In case that building rooftops have a simple rectangular shape, the simple object-oriented bounding box fits a rectangle over them representing building footprint. However, for the extraction of complex building shapes (e.g. L shape or U shape buildings) a more sophisticated approach is required.

To extract the minimal bounding box of a given object, a box aligned with the axes of the coordinate system and surrounding the object by fully enclosing it is searched. If the major axes of the box are parallel with the Cartesian axes, the dimensions of the bounding box are obtained as the maximum and minimum values of the object points along each of the x and y coordinates. However, since the bounding box may be arbitrarily oriented with respect to the coordinate system, the x and y coordinate system is rotated first to match the major axes of the segment. To this aim, the image skeleton technique is used. Indeed, in shape analysis the image skeleton shows the general shape of an object as a thin version of that shape. It detects the geometrical and topological properties of the shape, e.g. connectivity, topology, and length, as well as the main directions which are important for us here.

As shown by Fig. 5, In case of a more complex shape than a rectangle, first a bounding box is fitted to the whole U or L shape of the building, and then the segment derived from 3D segmentation is subtracted from the associated object-oriented bounding box. Subsequently, a morphological filtering is applied to eliminate the salt-and-pepper noise caused by unsmooth edges. Finally, the blob obtained from morphological filtering is represented by its own bounding box, which is subtracted from the initial whole-object bounding box to obtain the refined and final building footprint (see Fig. 5 f).

(a): Original rooftop segment	(b): Fit object-oriented bounding box	(c): Subtract segment from bounding box
(d): Morphological opening on remaining blob	(e): Subtract object-oriented bounding box from (4)	(f): Final shape of building footprint

Fig. 5. Schematic procedure to extract L and U shapes of building footprints by regularization of segments into rectangular primitives

IV. RESULTS AND ANALYSIS

Figure 6 shows five samples in our study area of building footprint extraction and regularization based on 3D segmentation. The 3D segmentation successfully delineates the building outlines for all samples except for the sample in

the last row which has very low contrast compared to the background. The last column shows the building footprints superimposed on to the associated DSM. Because of smoothing effects, the height constraints are unable to suppress the unsuccessful result (as in the last row) to fit an appropriate rectangle. This is the main drawback of height estimation based on image matching [11].

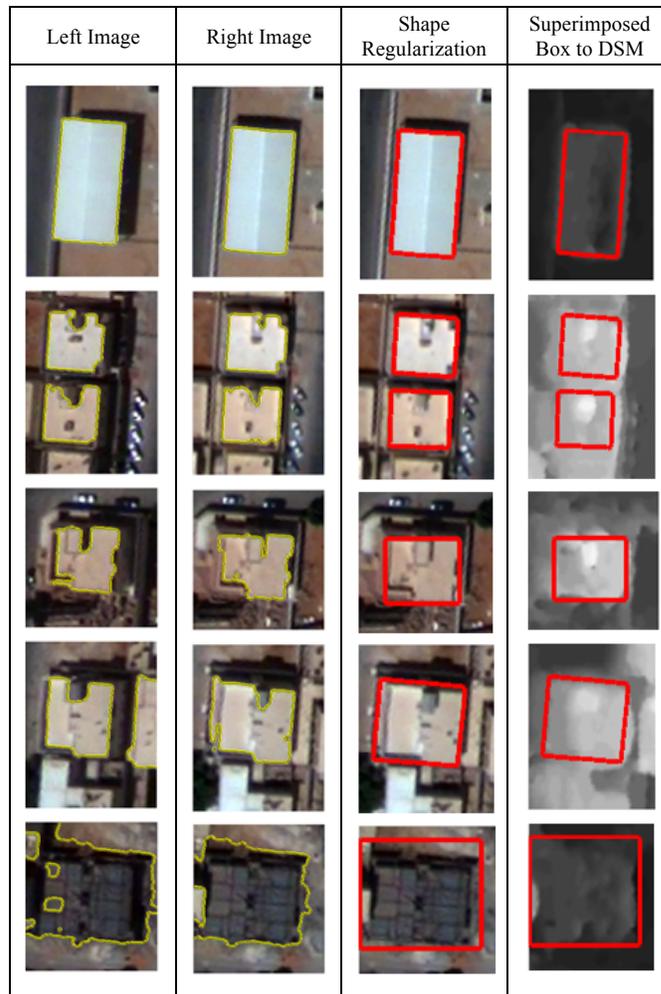


Fig. 6. Sample of 3D segmentation of building rooftop and regularization to rectangular building footprint superimposed to epipolar images and DSM.

To evaluate the quantitative performance of our approach, the footprints of all buildings in the test area were manually extracted and considered as the ground truth (GT). Table 1 compares the obtained results with the GT. The TP (True Positive) value refers to the area of buildings (in m^2) that are correctly extracted, the FP (False Positive) value indicates buildings area (in m^2) incorrectly extracted; whereas the FN (False Negative) value identifies the missed buildings' area. As seen in Table 1, we were able to automatically correctly extract around 90% of the buildings' area. Note that such a high rate of correct extraction is obtained because of the accurate segmentation initialization by height information. Indeed, the segmentation of building rooftop without prior information about building blobs delivers a disappointing rate of correct extraction.

Scenario	TP	FP	FN
no regularization	47223 m ²	9905 m ²	4948 m ²
after regularization	50362 m ²	6770 m ²	4948 m ²

Table 1. Quantitative evaluation of proposed method

Most of the false extractions (FP) are due to a low contrast between the building rooftops and the surrounding objects, while missed buildings are usually caused by small lofts over building rooftops which impede accurate rooftop extraction. To better evaluate the results, completeness and correctness are computed for two different scenarios: without the building segments regularization and after the regularization described in the previous section.

$$completeness = \frac{TP}{TP + FN}$$

$$correctness = \frac{TP}{TP + FP}$$

It is important to note that the regularization of building segments improves the detection correctness by more than 5%, from 82.7% to 88.1%.

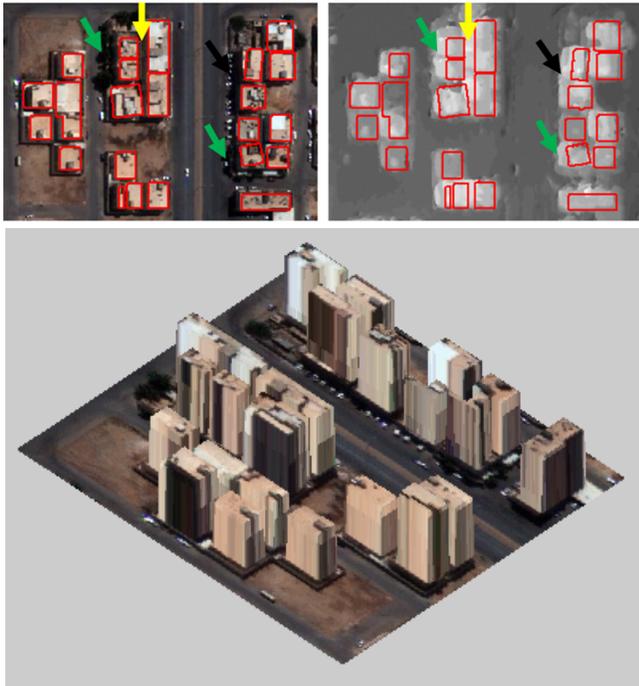


Fig. 7. (Upper): sample of 3D segmentation of building rooftop and regularization to rectangular building footprints superimposed to GeoEye image and DSM, green, yellow and black arrows demonstrate the influences of vegetation, the gap between individual buildings and shadow, respectively, (bottom): 3D visualization of buildings footprints using enhanced DSM.

V. CONCLUSION

A hybrid approach based on dense matching and image radiometric information for the extraction of building footprints is presented. In the test area of Riyadh and using GeoEye-1 stereo images, the proposed approach could extract most of building footprints, even those with relatively complex structures. Still, the proposed algorithm fails to detect building with very poor contrast with the background.

It can be concluded that when our previous investigations

[3], [5] based on the subtraction of two DSMs can be used as an alarm system for building change detection, it is not an efficient tool for the delineation of building outlines. In contrast, 3D segmentation can delineate building outlines provided that they have a good radiometric contrast with an acceptable size and shape. Note that the results of proposed method is highly dependent on building blobs created by height information. On the negative side, 3D segmentation followed by regularization may change the actual alignment of a building (i.e. the orientation of its main axes) so few minor misalignments are visible in Fig 7 (e.g., buildings that should be parallel to the street but show small deviations).

VI. BIBLIOGRAPHY

- [1] C. Heipke, P. A. Woodsford and M. Gerke, Updating Geospatial Databases from Images, Advances in Photogrammetry, Remote Sensing and Spatial Information Sciences. London: Taylor & Francis, S.355-362, 2008.
- [2] K. Jacobsen, "Characteristics of very High Resolution Optical Satellites for Topographic Mapping," in: IntArchPhRS vol. XXXVIII-4/W19, Hannover, 2011.
- [3] G. R. Dini, K. Jacobsen and C. Heipke, "Delineation of Building Footprints from High Resolution Satellite Stereo Imagery using Image Matching and a GIS Database," in: IntArchPhRS vol. XL-1/W1, 2013.
- [4] Y. Hui Lu, J. Trinder and K. Kubik, "Automatic Building Detection Using the Dempster-Shafer Algorithm," Photogrammetric Engineering & Remote Sensing, pp. 395-403, 2006.
- [5] G. R. Dini, K. Jacobsen, F. Rottensteiner, M. Alrajhi and C. Heipke, "3D Building Change Detection Using High Resolution Stereo Images and a GIS Database," in: IntArchPhRS vol. XXIX-B7, S. 299-304, 2012.
- [6] G. Dini, G. Lisini, M. Harb and P. Gamba, "Comparison of Estimated Building Story Number for Exposure Mapping From High Resolution Space-Borne Images," in IGARSS, Québec, 2014.
- [7] C. Baillard, C. Schmid, A. Zisserman and A. Fitzgibbon, "Automatic Line Matching and 3D Reconstruction of Buildings from Multiple Views," in: IntArchPhRS, Vol. 32 (Part 3-2W5), 69-80, 1999.
- [8] M. Ravanbakhsh, H. C. and P. K., "Road Junction Extraction From High Resolution Aerial Imagery," The Photogrammetric Record, vol. 23, no. 124, pp. 405-423, 2008.
- [9] T. F. Chan and L. A. Vese, "Active Contours Without Edges," Transactions on Image Processing, vol. 10, no. 2, pp. 266-277, 2001.
- [10] S. Osher and R. Fedkiw, Level Set Methods and Dynamic Implicit Surfaces, New York: Springer-Verlag, 2002.
- [11] H. Hirschmueller, "Stereo Processing by Semiglobal Matching and Mutual Information," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30(2), pp. 328-341, 2008.
- [12] F. Lafarge, X. Descombes, J. Zerubia and M. Pierrot-Deseilligny, "Automatic Building Extraction from DEMs Using an Object Approach and Application to the 3DCity Modeling," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 63, pp. 365-381, 2008.