

AUTOMATIC ROOFTOP EXTRACTION IN STEREO IMAGERY USING DISTANCE AND BUILDING SHAPE REGULARIZED LEVEL SET EVOLUTION

J. Tian^{*}, T. Krauß, P. d'Angelo

German Aerospace Center (DLR), Remote Sensing Technology Institute, 82234 Wessling, Germany
(Jiaojiao.Tian, Thomas.Krauss, Pablo.Angelo)@dlr.de

Commission WG I/4, III/4

KEY WORDS: Rooftop extraction, Level set, DSM, DTM

ABSTRACT:

Automatic rooftop extraction is one of the most challenging problems in remote sensing image analysis. Classical 2D image processing techniques are expensive due to the high amount of features required to locate buildings. This problem can be avoided when 3D information is available. In this paper, we show how to fuse the spectral and height information of stereo imagery to achieve an efficient and robust rooftop extraction. In the first step, the digital terrain model (DTM) and in turn the normalized digital surface model (nDSM) is generated by using a newly step-edge approach. In the second step, the initial building locations and rooftop boundaries are derived by removing the low-level pixels and high-level pixels with higher probability to be trees and shadows. This boundary is then served as the initial level set function, which is further refined to fit the best possible boundaries through distance regularized level-set curve evolution. During the fitting procedure, the edge-based active contour model is adopted and implemented by using the edges indicators extracted from panchromatic image. The performance of the proposed approach is tested by using the WorldView-2 satellite data captured over Munich.

1. INTRODUCTION

Automatic extraction of building rooftops from satellite images is one of the fundamental tasks in remote sensing image understanding. And there are a number of approaches available regarding this topic (Weidner and Förstner, 1995; Sohn and Dowman, 2007; Cote and Saeedi, 2013; Wang et al., 2015; Liasis and Stavrou, 2016).

Classical 2D image processing techniques are expensive due to the high amount of features required to locate buildings. This problem can be avoided when 3D information is available. Besides LiDAR, 3D information can be derived from satellite stereo imagery, which has a larger field of view and at a lower cost with respect to aerial imagery (Qin et al., 2016). The improvement of image resolution and the additional 3D information lead to several new challenges. On one hand, with higher resolution, sharpen building boundaries are visible, but many unexpected details from other objects may influence the rooftop detection procedure. On the other hand, higher resolution brings also large data size, thus the improvement on robust and efficiency of the rooftop extraction approaches are required, especially for complex shaped rooftops. However, many existing approaches were only suitable for rooftops with simple shape.

Level set, as an efficient edge detection and segmentation method is presented in recent years and has achieved good performance. The key idea of level set method was firstly published by Dervieux and Thomasset (1980, 1981). This idea was further improved by Osher and Sethian (1988) and adopted for capturing dynamic interfaces and shapes. Caselles et al. (1993) introduced the level set method to image segmentation in the context of active contour model. Level set is advantaged in the ability of handling topological changes, such as splitting and merging in an efficient way, and has achieved good

performance in image segmentation and boundary detection in computer vision. Level set has been also involved in building extraction using remote sensing data (Ahmadi et al., 2010; Kim and Shan, 2011; Li et al., 2014). Cote and Saeedi (2013) used corner points detected from multispectral aerial images as building corner candidates and further refined through level set curve evolution. Liasis and Stavrou (2016) proposed to use the building masks from morphological filtering as the initial building boundaries for level set evolution.

In this paper, a more efficient approach is proposed by combining the distance regularized level set evolution (DRLSE) model proposed by Li et al. (2010) and the height information from the stereo imagery. This paper is organized as follows. Section 2 describes the proposed building extraction methodology. In Section 3, the proposed method has been tested on buildings with various rooftop shapes. Conclusion and perspectives will be found in Section 4.

2. METHODOLOGY

Three most relevant steps for automatic building rooftop extraction are building location detection, building boundary generation and building modelling. The proposed approach is dedicated to stereo imagery. As a pre-processing step, the Digital Surface Model (DSM) is generated and the panchromatic and multispectral images are orthorectified by using the generated DSM. As shown in Figure 1, the proposed approach includes three steps: Digital Terrain Model (DTM) generation, initial building mask generation and rooftop boundary detection.

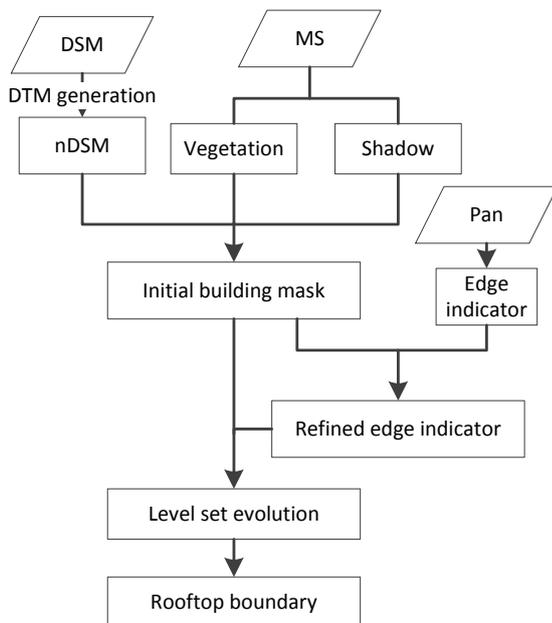


Figure 1 Building rooftop extraction method flowchart.

2.1 DTM generation

The DTM will be derived by a newly developed approach based only on height steps and not as usual relying on horizontal distances. The method called “DSMtoDTMstep” works also on unfilled DSMs which are often the result of dense stereo matching algorithms due to many voids emerging from occlusions or mismatches.

The proposed method detects in contrary to most other methods not “ground pixels” but removes “high pixels” from the DSM. For this the DSM is traversed in eight directions (left to right, right to left, top to bottom, bottom to top and the four diagonal directions). First a mask of the size of the DSM is initialized to “unknown”. Starting with a valid height value h the next valid height value h_n (omitting voids) in the scanning direction is investigated. If $h_n - h > T_{up}$ (1) the actual value in the mask is set to “high”. Following the scanning direction all further pixels are also marked as “high” until the actual height value h and the previous valid height value h_p satisfy

$h_p - h > T_{dn}$ (2). All following mask pixels are not changed until Equation (1) meets again.

Afterwards the mask represents all detected “high” regions. For deriving the DTM all values in the DSM with a high-mask value are set to void and this masked DSM is interpolated and filled to the final DTM. The two threshold parameters are set to $T_{up} = 2$ m and $T_{dn} = 1$ m.

2.2 Initial building mask generation

Building location detection is not easy especially for urban area with buildings in high density. However this problem can be solved easily when height information is available. Height information can separate higher objects such as buildings and trees from other objects. Additionally regions with vegetation covers can be removed through the analysis of the normalized difference vegetation index (NDVI).

As the first step of building location detection the normalized DSM is calculated by subtracting the digital terrain model

generated from the DSM ($nDSM = DSM - DTM$). In $nDSM$ only the objects above ground are preserved. High-level object mask ($Mask_h$) is generated after set a threshold value T to the prepared $nDSM$.

$$Mask_h = \begin{cases} 1, & \text{if } nDSM \geq T; \\ 0, & \text{if } nDSM < T. \end{cases} \quad (3)$$

We suppose that pixels with a higher probability to be shadows and vegetation have a lower probability to be part of the buildings. The details of the vegetation /shadow probability map calculation procedure are described in Tian et al (2014). The initial building mask is thus generated after removing the vegetation and shadow regions from $Mask_h$.

2.3 Building rooftop extraction

2.3.1 Level set

The basic idea of level set function (LSF) is to model the temporal evolution of a curve $C(t)$ using a family of embedding functions ϕ (Caselles et al., 1993; Li et al., 2010; Yang et al., 2014)

$$C(t) = \{(x, y) | \phi(x, y) = 0\} \quad (4)$$

The evolution of the LSF (LSE) can be written as (Osher and Sethian, 1988), with using the definition of the normal $n = \nabla\phi / |\nabla\phi|$

$$\frac{\partial\phi}{\partial t} = -F|\nabla\phi| \quad (5)$$

where F indicates the speed of the evaluation, and ∇ is the gradient operator, t is a temporal variable.

A number of data terms and regularizing terms have been proposed to obtain accurate and robust results (Cremers et al., 2007). As an extension of the edge-based LSEs (Caselles et al., 1993), the geodesic active contour was proposed by Caselles et al (1997). Its level set formula was given by

$$\begin{cases} \frac{\partial\phi}{\partial t} = g(I)|\nabla\phi| \left(\text{div} \left(\frac{\nabla\phi}{|\nabla\phi|} \right) + \vartheta \right) + \nabla\phi \cdot \nabla g(I) \\ g(I) = \frac{1}{1 + |\nabla G_\sigma * I|^2} \end{cases} \quad (6)$$

Where ϕ is the embedding LSF, I denotes the original image. $g(I)$ is the defined edge indicator function respect to the image gradient by using G_σ as a Gaussian kernel with a standard deviation σ . $*$ is the convolution operator. $\text{div}(\nabla\phi/|\nabla\phi|)$ is the mean curvature of the zero level curve and ϑ is a constant value. $\nabla\phi \cdot \nabla g(I)$ pushes the zero level curve to the desired object boundary according to the high gradient variations.

2.3.2 Distance regularized level set curve evolution (DRLSE)

To solve the revitalization step of the LSF, Li et al (2010) proposed the distance regularized level set curve evaluation (DRLSE) which does not need to reinitialize LSF repeatedly and enables a more general and efficient initialization of LSF. Moreover, DRLSE is able to maintain a desired shape of LSF. The evolution equation of DRLSE is defined by:

$$\frac{\partial\phi}{\partial t} = \mu \text{div}(d_p(|\nabla\phi|)\nabla\phi) - \frac{\partial\epsilon_{ext}}{\partial\phi} \quad (7)$$

Where

$$d_p(s) = \begin{cases} \frac{\sin(2\pi s)}{2\pi s}, & \text{if } s \leq 1 \\ 1 - \frac{1}{s}, & \text{if } s \geq 1 \end{cases}, \text{ and } \operatorname{div}(d_p(|\nabla\phi|)\nabla\phi) \text{ is the}$$

proposed distance regularization term. $\partial\varepsilon_{ext}/\partial\phi$ is the Gâteaux derivation of the external energy function ε_{ext} with respect to the zero level curve.

The *matlab* code from Li et al. (2010) is adopted for the application. The building boundary of the prepared building mask is served as the initial level set function. As a building shape regularization approach the edges indicators ($g(I)$) is further refined, which means only the values along the building mask boundary regions are used in the LSF evolution procedure.

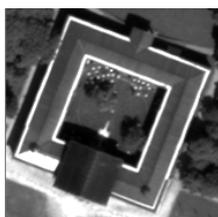
3. EXPERIMENT

In this section, we test the proposed building rooftop extraction methods on satellite stereo imagery.

3.1 Data sets

The date set consists of multi-spectral and panchromatic images from WorldView-2 data captured over Munich, Germany. The test site features buildings with complex shapes and high density. Two overlapping WorldView-2 stereo-pairs acquired on 12th of June 2010 were available for central Munich. After RPC block adjustment, a DSM with ground sampling distance of 0.5 m was generated from the panchromatic images using Semi-Global-Matching (Hirschmüller, 2008; d'Angelo and Reinartz 2011). All image combinations were matched and the resulting pairwise DSMs were fused into the final DSM using a pixel-wise median filter. The four way overlap minimizes occluded areas and provides dense data with fewer outliers than a single stereo pair.

The panchromatic and multi-spectral scenes closest to nadir were orthorectified using the generated DSM. Prior to orthorectification, remaining occluded areas in the DSMs were filled by a ground based interpolation scheme, which only used the lower segments adjacent to each occlusion, to avoid blurring building boundaries.



(a)



(b)



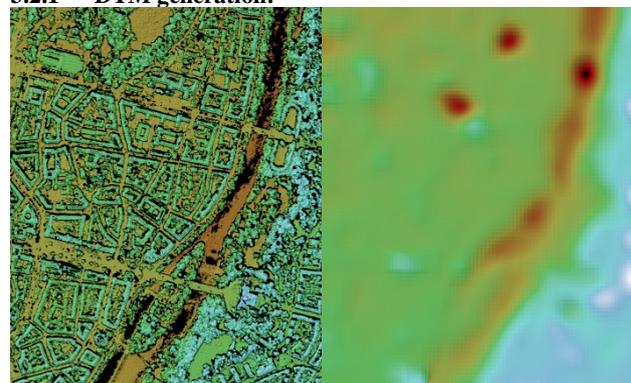
(c)

Figure. 2 Test regions, including two single building test sits (a) and (b) and one larger test site (c).

Three test regions are selected for the building rooftop extraction procedure. The first two regions which are shown in Figure 2 (a) and (b) contain one single building with complex shaped rooftops, respectively. The third test region (shown in Figure 2(c)) is a larger test site with a number of buildings with various rooftop shapes.

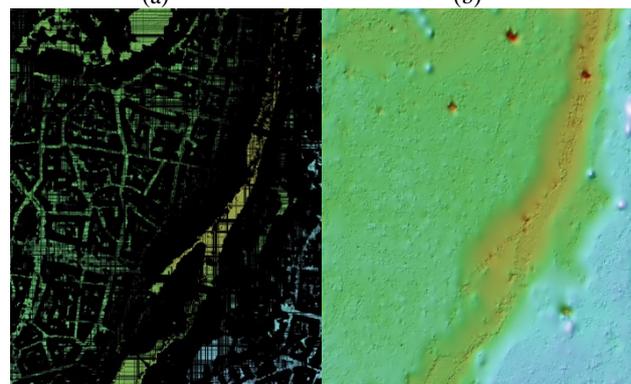
3.2 Results

3.2.1 DTM generation:



(a)

(b)



(c)

(d)

Figure. 3: DTM generation, (a) unfilled original DSM, (b) DTM derived by „classical“ morphological approach, (c) masked non-high areas in DSM created by DSMtoDTM step, (d) filled final DTM of DSMtoDTMstep

As shown in Figure. 3 the finally derived DTM represents the ground values of the DSM much better than the “classical” approaches. But there are still outliers (red “holes” in the upper part) which originate from mismatches in the DSM generation process.

3.2.2 Refined building rooftop boundary:

After DTM generation the rooftop extraction approach has been further performed on the three test sites. The obtained results are presented in Figure 4-6. The rooftop boundary improvements can be well observed from Figure 4 and 5. In both case, the rooftop boundaries are well presented in the final results.

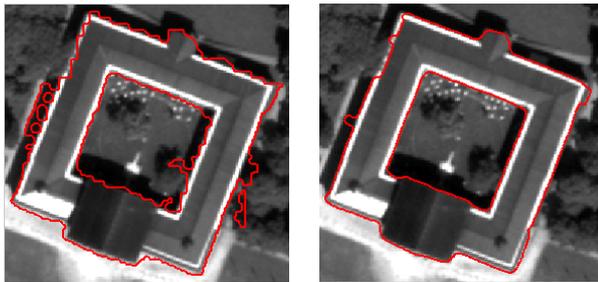
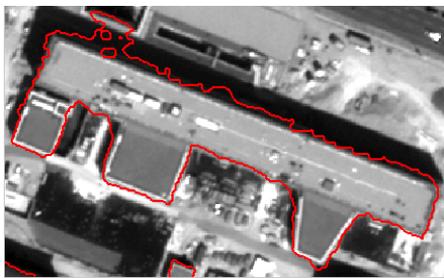
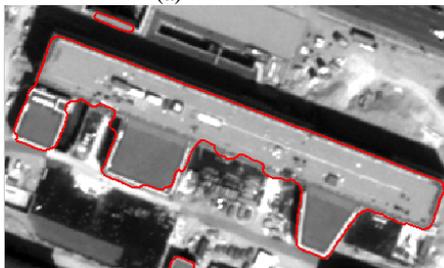


Figure 4. Building rooftop extraction result of Building1 (a) initial building mask boundary (b) improved boundary.



(a)



(b)

Figure 5. Building rooftop extraction result of Building2 (a) initial building mask boundary (b) improved boundary.

Besides single complex shaped buildings, results from a larger test site are displayed in Figure 6. Figure 6(a) is the extracted nDSM. The initial building masks are shown in Figure 6(b), in which the building locations and initial rooftop shapes are well derived. The final extracted rooftop boundaries are plotted with red polygons in Figure 6 (d). As it shows, most of the closed rooftops are well presented. However, several problems may also be identified at some complex rooftops which feature more than one hole on the rooftops, like building No.1. In the initial building mask, some parts of the building are separated from the main part. Therefore, these parts are not well preserved in the final result.

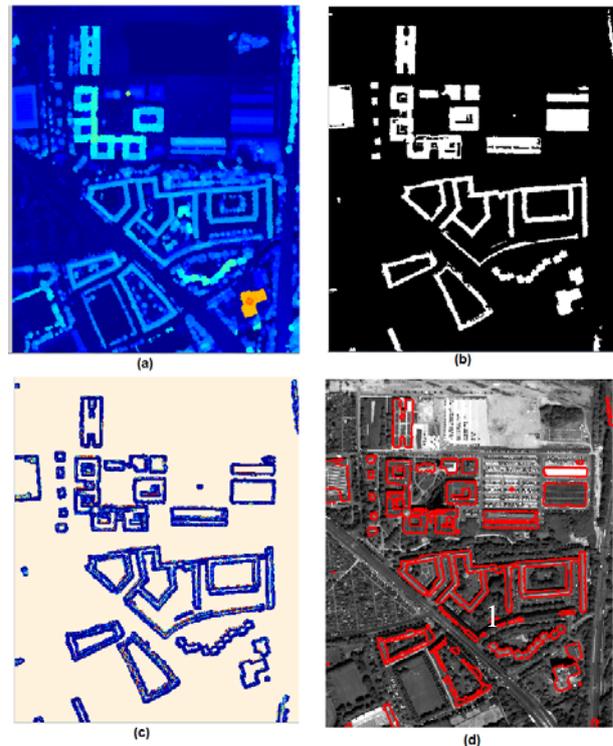


Figure 6. Building rooftop extraction result of Building2 (a) nDSM; (b) initial building mask; (c) edge detector around building masks; (d) improved building boundary

In Figure 6,

4. DISCUSSION AND CONCLUSION

This paper presents a robust and efficient automatic building rooftop extraction approach. It combines all information that can be extracted from WorldView-2 stereo imagery. The framework is able to use the third dimensional information from stereoscope data to derive the building locations and detect the initial building rooftop shapes. The adopted level set evolution can further use the edge information from the panchromatic image, thus to derive more accurate rooftop boundaries. Our first contribution is the DTM generation approach, which allows obtaining the absolute building height values. The second contribution is the building rooftop extraction workflow, which has been applied to buildings with complex shaped rooftops. The extracted result can be used to update existing city maps, refine the extracted DSM, or detect changes by comparing it with existing building footprint.

Further work will involve the investigation of other satellite / aerial stereo imagery and methods to exploit other efficient edge detectors and contour energy functions.

5. ACKNOWLEDGMENT

The Authors would like to thank European Space Imaging (EUSI) for providing the Worldview-2 data in Munich area.

REFERENCES

Ahmadi, S., Zoj, M.V., Ebadi, H., Moghaddam, H.A. and Mohammadzadeh, A., 2010. Automatic urban building boundary extraction from high resolution aerial images using an innovative model of active contours. *International Journal of*

- Applied Earth Observation and Geoinformation*, 12(3), pp.150-157.
- Caselles, V., Catté, F., Coll, T. and Dibos, F., 1993. A geometric model for active contours in image processing. *Numerische mathematik*, 66(1), pp.1-31.
- Caselles, V., Kimmel, R. and Sapiro, G., 1997. Geodesic active contours. *International journal of computer vision*, 22(1), pp.61-79.
- Cote, M. and Saeedi, P., 2013. Automatic rooftop extraction in nadir aerial imagery of suburban regions using corners and variational level set evolution. *IEEE transactions on geoscience and remote sensing*, 51(1), pp.313-328.
- Cremers, D., Rousson, M. and Deriche, R., 2007. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *International journal of computer vision*, 72(2), pp.195-215.
- d'Angelo, P and Reinartz, P., 2011. *Semiglobal Matching Results on the ISPRS Stereo Matching Benchmark*. High-Resolution Earth Imaging for Geospatial Information, Hannover, Germany. ISSN 1682-1777
- Dervieux, A. and Thomasset, F., 1980. A finite element method for the simulation of a Rayleigh-Taylor instability. In *Approximation methods for Navier-Stokes problems* (pp. 145-158). Springer Berlin Heidelberg.
- Dervieux, A. and Thomasset, F., 1981. Multifluid incompressible flows by a finite element method. In *Seventh International Conference on Numerical Methods in Fluid Dynamics* (pp. 158-163). Springer Berlin/Heidelberg
- Hirschmüller, H., 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence*, 30(2), pp.328-341
- Kim, K. and Shan, J., 2011. Building roof modeling from airborne laser scanning data based on level set approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4), pp.484-497.
- Li, C., Xu, C., Gui, C. and Fox, M.D., 2010. Distance regularized level set evolution and its application to image segmentation. *IEEE Transactions on image processing*, 19(12), pp.3243-3254.
- Li, Z., Liu, Z. and Shi, W., 2014. A fast level set algorithm for building roof recognition from high spatial resolution panchromatic images. *IEEE Geoscience and Remote Sensing Letters*, 11(4), pp.743-747.
- Liasis, G. and Stavrou, S., 2016. Building extraction in satellite images using active contours and colour features. *International Journal of Remote Sensing*, 37(5), pp.1127-1153.
- Osher, S. and Sethian, J.A., 1988. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *Journal of computational physics*, 79(1), pp.12-49.
- Qin, R., Tian, J. and Reinartz, P., 2016. Spatiotemporal inferences for use in building detection using series of very-high-resolution space-borne stereo images. *International Journal of Remote Sensing*, 37(15), pp.3455-3476.
- Sohn, G. and Dowman, I., 2007. Data fusion of high-resolution satellite imagery and LiDAR data for automatic building extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(1), pp.43-63.
- Tian, J., Cui, S. and Reinartz, P., 2014. Building Change Detection Based on Satellite Stereo Imagery and Digital Surface Models. *IEEE Trans. Geoscience and Remote Sensing*, 52(1), pp.406-417.
- Wang, J., Yang, X., Qin, X., Ye, X. and Qin, Q., 2015. An efficient approach for automatic rectangular building extraction from very high resolution optical satellite imagery. *IEEE Geoscience and Remote Sensing Letters*, 12(3), pp.487-491.
- Weidner, U. and Förstner, W., 1995. Towards automatic building extraction from high-resolution digital elevation models. *ISPRS journal of Photogrammetry and Remote Sensing*, 50(4), pp.38-49.
- Yang, X., Gao, X., Tao, D. and Li, X., 2014. Improving level set method for fast auroral oval segmentation. *IEEE transactions on Image Processing*, 23(7), pp.2854-2865.