

# Fusion of multi-spectral bands and DSM from Worldview-2 Stereo imagery for building extraction

Jiaojiao Tian

Remote Sensing Technology Institute  
German Aerospace Center (DLR)  
Wessling, Germany  
Jiaojiao.Tian@dlr.de

Peter Reinartz

Remote Sensing Technology Institute  
German Aerospace Center (DLR)  
Wessling, Germany  
Peter.Reinartz@dlr.de

**Abstract**—This paper describes a building extraction approach by fusion panchromatic image, multispectral image and Digital Surface Model (DSM) generated from WorldView-2 stereo imagery. First, the DSM is used to indicate a rough building location, and then Hough-transform is followed to generate 2D segments from panchromatic images. Hough lines in the two main directions of each building are kept. The probability of each segment belonging to a building is calculated from random forest estimations. An automatic training data selection strategy is designed for this supervised classification. Finally, the 2D segments and classification results are combined to get the final building boundaries. This method has been tested in the city center of Munich, Germany.

## I. INTRODUCTION

Automatic building extraction is one of the most challenging problems in remote sensing image analyses. In addition to 2D information from multi-spectral satellite images, height information from DSMs has received increasingly attention for automatic building extraction. A lot of work has been done by combining use of Airborne LiDAR data and airborne multi-spectral data. However both kinds of datasets are expensive and exhibit low temporal resolution. Therefore, the building extraction from satellite imagery is an interesting research topic.

Numbers of building extraction methods from stereo imagery have been developed in the last 20 years. A detailed workflow of DSM assisted building extraction has been promoted in [1], restricted to the DSM quality at that time only a coarse model of buildings could be generated. Ridley et al. [2] used airborne data to simulate the IKONOS stereo data for land cover extraction, from which only 73% and 86% of buildings could be interpreted correctly, while no automatic method was mentioned in that paper. After the IKONOS imagery became commercially available, more detail 3D building models were extracted, but this work was still based on manual interpretation [3]. In [4] a building detection method was promoted from multiple airborne data, however only

simple shapes were considered. In [5] DSM from Quickbird stereo imagery was mainly used for Orthophoto generation, building extraction result was still from manual digitalization.

Even though many methods have been promoted and tested on stereo imagery, they either require a lot of manual work or are only suitable for well separated and convex shape buildings. Therefore, more work is required for automatic building extraction from space borne stereo imagery, especially in dense high building areas with complex building shapes.

The automatic building extraction method developed for airborne multi-spectral imagery and Lidar DSM cannot be directly adopted for e.g. Worldview-2 stereo imagery. Because the airborne data have better resolution than space borne satellite data and the LiDAR DSM has higher quality. Figure 1 shows the quality of the DSMs generated from Worldview-2 stereo imagery.

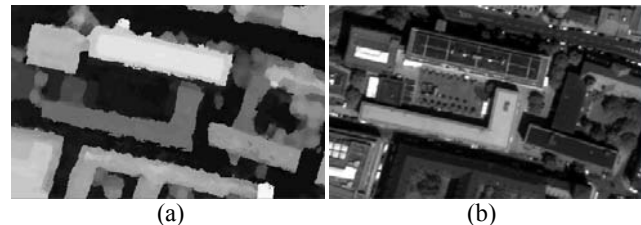


Figure 1. DSM and Panchromatic images (a) DSM generated from Worldview-2 stereo imagery; (b) Corresponding orthorectified panchromatic image).

Therefore, this paper has proposed a novel efficient building extraction method based on Worldview-2 stereo imagery. It does not need any manual interpretation or any other data source. It combines height from stereo view, line segments from the panchromatic channel together with building class probability from random forest field classification based on the multispectral channels. We introduce also a knowledge based training data extraction method for the supervised classification.

## II. METHODOLOGY

The three most relevant steps for automatic building extraction are building detection, building boundary generation and building modeling. We use different data input for each step. As shown in Figure 2, building location is detected from the DSM. Building boundaries are generated by line detection based on panchromatic images. In the last step, buildings are generated from the automatic classification results.

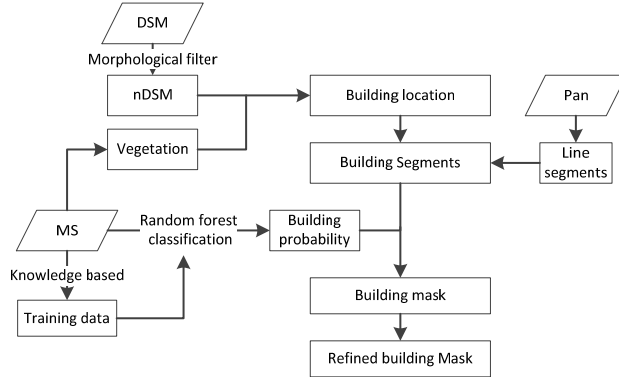


Figure 2. Flowchart of the proposed method (Pan: panchromatic image; MS: Multispectral image; DSM: Digital Surface Model; nDSM: normalized DSM).

### A. Building location detection

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

The first step of building location detection is to remove the underlying elevation by subtracting the digital terrain model (DTM) from the DSM. For this purpose we have used the top-hat morphological filter. After that only the objects above ground are preserved in the normalized DSM (nDSM).

After removing the vegetation (MaskNDVI) from nDSM by using Equation 1, the initial building areas can be generated.

$$\text{Initial Mask} = \text{nDSM} - \text{nDSM} \cap \text{MaskNDVI}$$

The initial building mask can only show the possible locations of each building. If vegetation areas are located on the roof of some buildings, only the parts of buildings without vegetation cover are shown in the initial building mask.

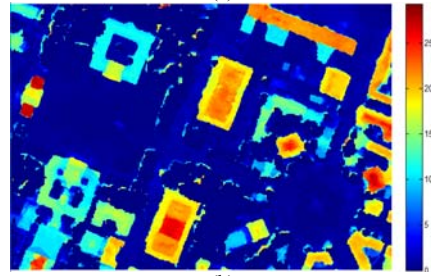
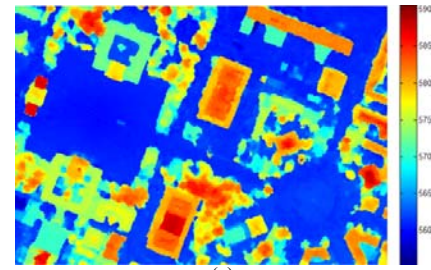


Figure 3. Building location detection ((a) original DSM; (b) nDSM after removing vegetation).

### B. Hough line based segmentation

Separating buildings into 2D polygons has been proven to be an efficient method in building extraction [6]-[7]. Here we use more robust Hough transformation in detecting building directions through line segments.

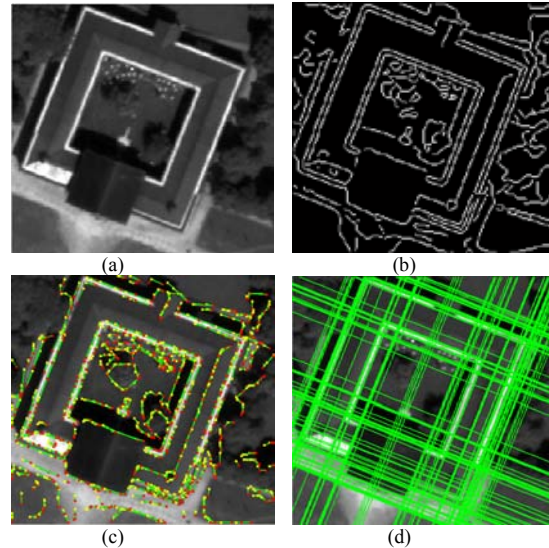


Figure 4. Building boundary detection (a) Panchromatic images (b) Canny edges (c) Hough lines (d) extended Hough lines in main directions.

The building segments detection procedure is shown in Figure 4. Single buildings can be easily separated from the initial building mask. The corresponding panchromatic image patch around each building is processed separately to reduce the influence from other buildings. An example is shown in Figure 4a. The Canny lines extracted from panchromatic image are transformed using a Hough transform [8]. Figure 4c

displays all of the detected Hough lines. At the same time, we can also get the main directions of the building. It is assumed that each building has only two main directions [7], and these two directions are perpendicular to each other. The intersection of these lines provides 2D segments as shown in figure 4d.

### C. Knowledge-based training data selection

Recently, supervised classification is attracting more attention, from the maximum likelihood to support vector machine, or random forest. Their efficiency and accuracy have been proven in several publications [9]-[10]. However, these methods cannot be directly used for automatic classification as they require training data. Training data selection is time consuming, and incorrect or insufficient number of training data may influence the finally accuracy. Therefore, we developed a knowledge-based automatic training data selection method for supervised classification.

Existing knowledge can make the classification easy and fast, if the knowledge is accurate and detailed enough [11]. However, normally only part of the knowledge can be obtained in advance, which complicates the class definition. In this paper, only the high probability features are used to define a class. Here, we have used the height from nDSM, NDVI and shadow mask from multi-spectral images. Shadow is extracted with the method introduced by Marchant and Onyango [12]. The NDVI and shadow values are projected to a sigmoid curve to get the Belonging-Certainly Value (BCV) from 0 to 1[13]. Thresholds are selected by long term experience and are fixed for all data.

TABLE I. KNOWLEDGE BASED TRAINING DATA SELECTION RULES

Classes	Rules		
	Height (m)	BCV_veg	BCV_shadow
Trees	>10	>80%	<0.8
Grass	<2	>80%	<0.8
Building	>10	<50%	-----
Shadow	<10	-----	>80%
Ground	<2	<50%	<50%

The Training data selection rules have been listed in Table I. Fig. 5 shows one part of the generated training data. Segments with size smaller than 100 pixels have been removed from each class layer.

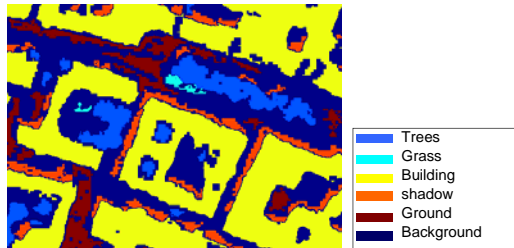


Figure 5. Generated training data (with 5 classes).

### D. Random forest classification and building extraction

Random forest is a robust and powerful machine learning classifier, and capable to process large datasets [14]. Here this classifier is used to extract the building class based on the generated training data. All of the channels from multispectral images, together with panchromatic image and DSM are used as input features. Instead using the classification result, we used probability values for each pixel belonging to the building class instead. These probability values from each pixel are averaged to each 2D segments obtained from section 2B. In the end, the final building mask is generated after thresholding.

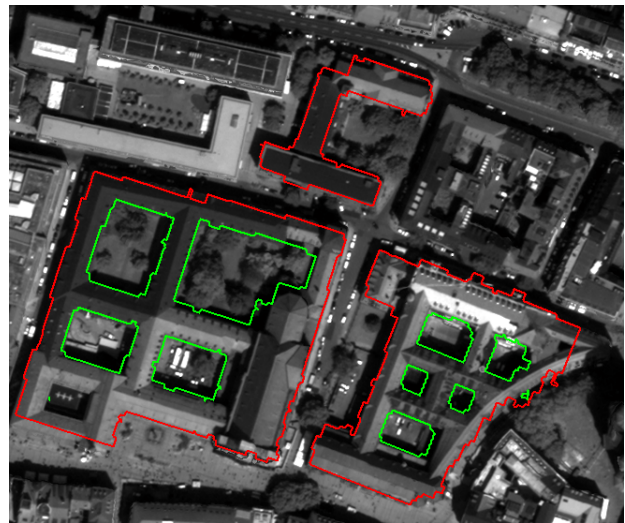
## III. DATASET AND EXPERIMENTS

### A. Dataset description

The study side is located in Munich, Germany, which features buildings with complex shapes and high density. Both Worldview-2 stereo multi-spectral and panchromatic images are available. DSM with 1m grid resolution is generated using the semi-global stereo matching algorithm [15]-[16] from the stereo panchromatic imagery. The nadir scenes of multispectral and panchromatic images are ortho-rectified by using the generated DSM. Model with 10 trees and depth of each tree set to 5 is used in the random forest classification procedure.

### B. Results

A visual inspection of the resulting building outlines is conducted to evaluate the automatic approach. Three buildings representing three difficult levels are analyzed here. In our experiments the 2D segments are detected as building, if the segments have an average probability belonging to a building of more than 80%. The outside/ inside building boundaries of the extracted building mask are displayed with red and green color respectively in Figure 6.



(a)

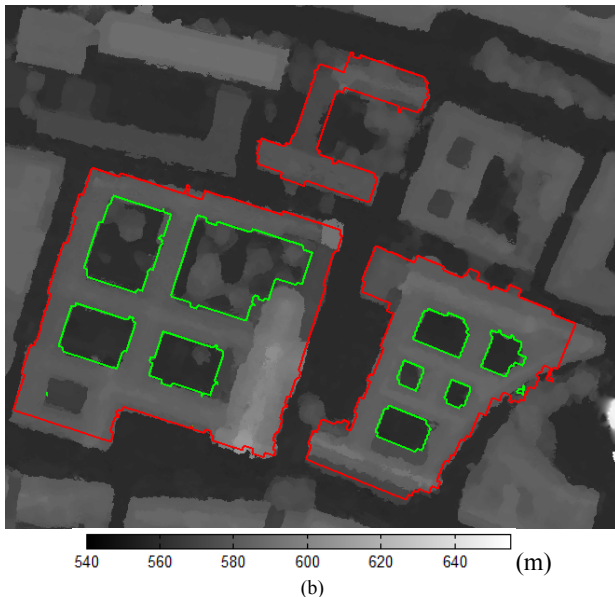


Figure 6. Extracted Buildings (a: Extracted building boundary compared to panchromatic image , b: Extracted building boundary compared to DSM)

As can be seen, the building on the top is a simple polygon shape, while the left one is composed of several polygons. One side of the right building is a curve. Most part of the building boundaries are super-imposed with these buildings, even most of the courtyards are detected correctly. And the edges are much sharper than those of the DSM (shown in Figure 6b).

#### IV. CONCLUSION AND DISCUSSION

This paper presents a robust automatic building extraction methodology. It combines all information that can be extracted from Worldview-2 stereo imagery. Our first contribution is the building extraction work flow. This work flow allows introducing various features from DSM, panchromatic image and multi-spectral image separately for location, boundary, and class detection respectively. This methodology has been applied to a high dense urban area with complicated building shapes. The experiment displays the ability of the proposed method to cope with high building density, and complicated building shapes.

It also gives possible indications of extracting buildings from just one satellite sensor and data acquired only in one orbit. This leads to the possibility of fast updates of urban area building mapping all around the world. The extracted results can be used to update existing city maps, refine the extracted DSM, or detect changes by comparing it with existing building footprint. In order to further improve the building extraction results, the boundaries refinement might help and a more quantitative analysis of the results will follow.

#### ACKNOWLEDGMENT

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

#### REFERENCES

- [1] E. Baltsavias, S. Mason and D. Stallmann, "Use of DTMs/DSMs and orthoimages to support building extraction automatic extraction of man-made objects from aerial and space images", Proc. Automatic Extraction of Man-Made Objects from Aerial and Space Images, 1995, pp.199-210.
- [2] H. M. Ridley, P. M. Atkinson, P. Aplin, J.-P. Muller, and I. Dowman, "Evaluating the potential of the forthcoming commercial U.S. high-resolution satellite sensor imagery at the ordnance survey," Photogrammetric Engineering & Remote Sensing, vol. 63, no. 2, 1997, pp. 997-1005.
- [3] C. S. Fraser, E. Baltsavias, and A. Gruen. "Processing of IKONOS imagery for submeter 3D positioning and building extraction," ISPRS Journal of Photogrammetry & Remote Sensing, vol. 56, 2002, pp. 177-194.
- [4] S. Noronha and R. Nevatia, "Detection and modeling of buildings from multiple aerial images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, 2001, pp. 501-518.
- [5] A. Tong, S. Liu and Q. Weng, "Geometric processing of QuickBird stereo imageries for urban land use mapping: a case study in Shanghai, China", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 2, no. 2, June 2009, pp. 61-66.
- [6] C. Schmid and A. Zisserman, "Automatic Line Matching Across Views", In: IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 666-671.
- [7] L. Zebedin, J. Bauer, K.F.Karner and H. Bischof, "Fusion of Feature- and Area-Based Information for Urban Buildings Modeling from Aerial Imagery". ECCV (4), 2008, pp. 873-886.
- [8] S. Cui, Q. Yan and P. Reinartz, "Complex building description and extraction based on Hough transformation and cycle detection", Remote Sensing Letters, vol. 3, no. 2, 2012, pp. 151-159.
- [9] N. M. Dempster, A. P. Laird and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," Journal of the Royal Statistical Society, Series B, vol. 39, no. 1, Jan 1977, pp. 1-38.
- [10] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for the semisupervised classification of remote-sensing images," IEEE Transactions on Geoscience and Remote Sensing, vol. 44, no. 11, 2006, pp. 3363-3373.
- [11] E. P. Baltavias, "Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems," ISPRS Journal of Photogrammetry & Remote Sensing, vol. 58, 2004, pp. 129-151.
- [12] J. A. Marchant and C.M. Onyango, "Shadow-Invariant Classification for Scenes Illuminated by Daylight," Journal of the Optical Society of America A: Optics, Image Science, and Vision, vol. 17, no. 11, 2000, pp. 1952-1961.
- [13] J. Tian, S. Cui, P. Reinartz, "Building change detection based on satellite stereo imagery and Digital Surface Models," IEEE Transactions on Geosciences and Remote Sensing, accepted, 2013.
- [14] L. Breiman, "Random Forest," Machine Learning, vol. 45, no. 1, 2001, pp. 5-32.
- [15] H. Hirschmüller, "Stereo processing by semiglobal matching and mutual information," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30, no. 2, 2008, pp.1-14.
- [16] P. d'Angelo and P. Reinartz, "Semiglobal Matching Results on the ISPRS Stereo Matching Benchmark," High-Resolution Earth Imaging for Geospatial Information, Hannover, Germany, 2011.