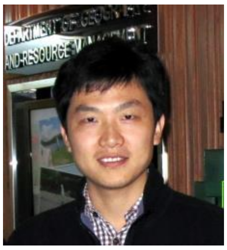
Fusion of World-view2 stereo and TerraSAR-X image for 3D building extraction in urban areas



Xu Yong[1], Ma Peifeng[2], Lin Hui[2], Ng Edward[3]
[1] *The Institute of Future Cities, The Chinese University of Hong Kong, Hong Kong, China, xuyong@cuhk.edu.hk*
[2] *The Institute of Space and Earth Information Science and Shenzhen Research Institute, The Chinese University of Hong Kong, Hong Kong, China, mapeifeng@cuhk.edu.hk; huilin@cuhk.edu.hk*
[3] *The School of Architecture, the Institute of Environment, Energy and Sustainability (IEES), and the Institute of Future Cities(IOFC), The Chinese University of Hong Kong, Hong Kong, China, edwardng@cuhk.edu.hk*

15 June 2015

1. Abstract

In this paper, we investigated the joint use of both a pair of high resolution WorldView-2 optical satellite images and the TerraSAR-X synthetic aperture radar satellite images, to extract 3D building information (including 2D building footprints and the third dimension building height) in high density urban areas. The main idea of the proposed fusion model is to take full advantage of both datasets for building extraction. Compared with SAR images, optical satellite images are more suitable for extracting building footprints, and can be applied to retrieve the heights of low buildings with higher accuracy. Whereas SAR images perform much better in retrieving the heights for tall buildings. But due to positioning errors and mutual interference of surrounding buildings, SAR images cannot be applied to retrieve the heights of low buildings efficiently. Therefore, in this study, both dataset are combined to generate 3D building product for high density urban area, where a large number of both tall and low buildings are included.

The proposed approach includes two main stages. Firstly, building footprints extraction. Optical satellite images are utilized to extract building footprints by using object-based analysis method, in which both building spectral, texture, contextual and elevation information are used. Secondly, building height retrieval. Initial heights of buildings are retrieved from both Stereo images and SAR images using photogrammetry and SAR interferometry techniques separately, then both initial results are combined with a novel object based fusion method, in which heights of points for the same building footprint are retrieved and integrated. The proposed approach is especially suitable for building extraction in high-density urban area where single satellite data has certain limitations. Experimental tests on Mong Kok area of Hong Kong city showed that the proposed approach using both stereo images and SAR images can achieve high overall building extraction accuracy.

2. Introduction

3D building information are important GIS data for urban planning and management. Though conventional field surveys can be conducted to obtain the 3D building information for urban area, they are labor intensive and time-consuming. Given the comparatively large area coverage and near real-time acquisition mode, satellite data have widely been applied to extract building information in urban areas [1-3]. According to the available satellite data, there are three categories of method for building height retrieval, stereo photogrammetry technology with pairs of optical images (stereo images), synthetic aperture radar (SAR) technology, and light detection and ranging data (LiDAR) technology.

Stereo images have been widely used for building retrieval since they became available [1]. This method usually processed the building footprint extraction and height retrieval separately. For building footprint extraction, various advanced image classification/segmentation approaches are available [1]. For building height retrieval, it usually has four main stages, data preprocessing, stereoscopic matching, digital surface model (DSM) generation, and above-ground building height extraction. Studies using stereo images for building height extraction have mainly focused on parts of the above-mentioned stages. Some studies have emphasized on the development of advanced stereoscopic matching methods to improve the building height retrieval accuracy [2], while others have investigated the use of multiple stereo images to identify common targets for height retrieval [3-4]. In addition, some studies have made use of ancillary data, such as approximate digital elevation modeling or prior building footprints [5]. In addition to stereo images, monocular satellite image can be used to retrieve the height of buildings in areas where the shadows of the buildings can be measured [6]. However, the accuracy of this method is affected by the quality of the measured shadows and the surrounding environments, which limits its practical use.

In addition to stereo images, SAR images have been widely used for building extraction in recent decades, and various SAR techniques have been explored. According to a data source on the usage of SAR, there are several different processing strategies, such as interferometry with inference SAR images [7-10], radargrammetry with stereo SAR images [11], monoscopic SAR images [12-14] and multi-aspect SAR images [15]. However, due to the intrinsic characteristics of SAR images (e.g., layover effect) and the effect of the mutual inference of the surrounding environment, studies using SAR data to derive building information are still limited to simple scenes of isolated buildings [5][10][12][13][16].

The advent of LiDAR has opened up a new phase of building detection research. LiDAR provides huge point clouds that significantly improve the accuracy of building detection [17][18]. Although the results obtained using LiDAR data are promising, it is still expensive and restrictions on flight plans in some countries limits its application in large urban areas.

To achieve a more promising result, scholars have investigated the integration of different data sources for 3D building retrieval, such us SAR and optical image fusion [16][20], and LiDAR and optical image fusion [17]. However, most of the fusion approaches take advantage of optical images for building footprint extraction rather than building height retrieval. Therefore, we investigated the joint use of stereo optical images and SAR images for both building footprint and height extraction in urban areas, where single data have limitations.

The remainder of the paper is organized as follows. In Section III, a novel building extraction approach using both stereo and SAR images is introduced. The results and the accuracy analysis are given in Section IV. Conclusions are presented in Section V.

3. Method

This section presents the proposed 3D building extraction approach using both stereo and SAR images. As shown in Figure 1, the approach is implemented using stereo images, multi-temporal SAR images as the inputs, and it includes two main stages. First, building footprints are extracted from high-resolution satellite image using object-oriented classification technology, and then building heights are estimated using an object-based height fusion approach [21], with which initial height estimates from both stereo images and multi-temporal SAR images are combined.

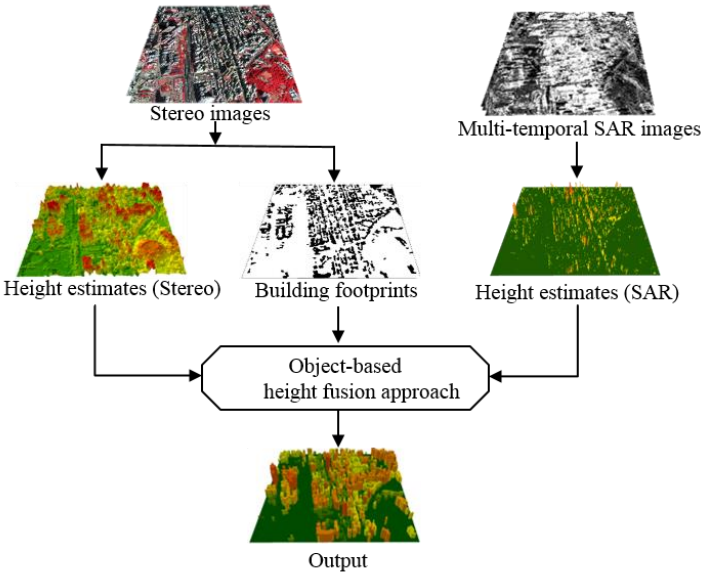


Figure 1. Procedure of the proposed approach for 3D building extraction

3.1 Building footprint extraction

In this stage, object-oriented classification technology was utilized to extract building footprints, and it mainly includes four main sub-steps.

A. Multi-scale object-oriented segmentation approach was used to obtain image objects, in which some small objects are merged with their surrounding objects according to their spectral difference. In this study, the object-oriented parameters were set as: scale parameter = 30, shape parameter = 0.3 and compact parameter = 0.5.

B. Non-building classes (e.g., vegetation and shadow) were removed from the obtained objects according the vegetation index NDVI and shadow index.

C. Non-building classes (e.g., roads) with different shape index were removed from the remained objects by using the shape index and object width/height ratio.

D. Non-building classes with similar heights of nearby grounds (e.g., bare lands) were removed by using the height model from stereo images.

The obtained objects were further processed using mathematical morphology operations to eliminate the spurious objects, and the remaining objects reflect building footprints.

3.2 Building height extraction

In this stage, heights of buildings were estimated and assigned to each building footprint. It has two sub-steps. Firstly, two initial building height estimates were extracted from stereo and SAR images. For the stereo images, an empirical model was used to extract the building height estimate using the built-in rational polynomial coefficient.

For the multi-temporal Terra-SAR images, the TomoSAR technique was used to derive the vertical heights of scatters [8]. Secondly, both of the initial building height estimates were combined to generate a fused result for the same building footprint using an object-based fusion approach. Details of the object-based fusion approach are given in [21].

4. Case study

4.1 Study area

The study area was located at 22N, 113E in the Mong Kok district of Hong Kong, a high-density urban area with an average building height above 40 meters and where the heights of some buildings reach several hundred meters. Using this study area, we collected the actual building data (including 2D building footprints and building height information) from 2010, the 14 TerraSAR-X images from 2011, and a pair of stereo images on 2014. The actual 3D building information from 2010 were used as validation data.

4.2 Experimental results

We used both SAR images and stereo images to extract the 3D building model. Figure 2(a) shows the used high-resolution satellite data. Figure 2(b) shows the preliminary building footprints from Figure 2(a) using the approach introduced in section 3.1. Figure 2(c) shows the filtered building footprints by removing some bare lands with lower heights. Figure 2(d) shows the refined building footprint result using morphology operations to fill some holes and eliminate some spurious objects. Figure 2(e) shows the final extracted buildings with heights, in which building heights are estimated from both stereo images and SAR images using the approach proposed in [21]. In comparison, actual building information is also provided in Figure 2(f).

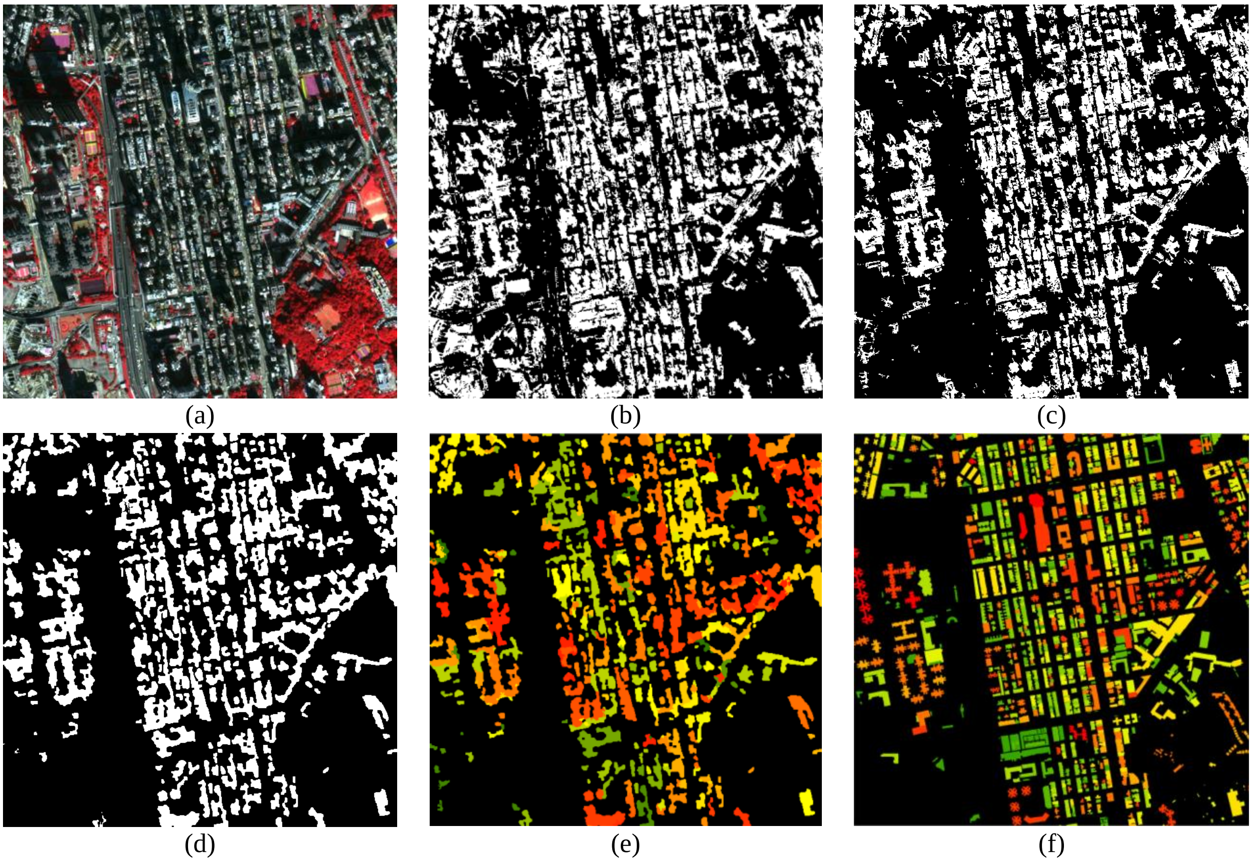


Figure 2. 3D building extraction with the proposed approach. (a) High-resolution satellite image. (b) Preliminary extracted objects using the object-oriented classification method. (c) Filtered result by removing the objects with low heights. (d) Extracted building footprints. (e) The extracted building footprints and building heights with the proposed approach. (f) Actual 3D building information for comparison.

4.3 Accuracy validation

The overall accuracy was used to assess the performance of the proposed approach in detecting building footprints. As the last row of Table 1 shows, the overall accuracy is 78 percent. The result in table 1 also indicated that the building detection result has certain limitations in high density urban areas. In particular, some features, like some bare lands and roads, having similar spatial and spectral characteristics of buildings, were wrongly detected as buildings. Meanwhile, considerable part of the actual buildings were failed to be detected with the proposed approach, part of the reasons was because of the shadow of some tall buildings, which block some low buildings.

The average absolute difference (AAD) was used to assess the performance of the proposed approach in detecting heights of buildings. The AAD index was calculated by comparing the predicted height with the actual building height. The lower of the AAD value is, the better quality of the predicted result. The AAD value using both datasets is 9.5 m in this study. Moreover, we also tested the quality of the extracted building information in generating two urban morphological parameters including building coverage ratio and building volume density. Statistical results showed that the accuracy for both indices are up to 80 percent when compared with the actual indices at spatial resolutions of several hundred meters.

Table 1. Accuracy statistics for the extracted building footprints using the proposed approach

	Building	Others	Producer's (%)
Building	48573	36485	57.1
Others	47307	227635	82.8
User's (%)	50.7	86.2	-
Over accuracy (%): **76.7** Kappa: **0.38**			

5. Result and Discussion

In this study, both stereo images and SAR images were jointly used to extract 3D building information in a high-density urban area. The benefit of the proposed approach is that the advantages of both stereo and SAR images are well combined, as SAR images are good at predicting the height of tall buildings, while stereo images can provide the complementary height information of buildings for which SAR is missing. Experimental Results show that the overall building footprint extraction accuracy can reach up to 78 percent, and the average absolute difference for building height estimation is of 9.5 meters using the proposed approach. Furthermore, high quality urban morphological parameters like building coverage ratio and building volume density can be obtained with the generated 3D building model, having a prediction accuracy up to 80 percent. These findings indicate that the generated building information using satellite data has a great potential for urban studies.

Acknowledgment

This study is supported by an internal research grant from the social science panel of the Chinese University of Hong Kong.

References

[1] Avery T.E., and Berlin G.L., 1992: Fundamentals of Remote Sensing and Airphoto Interpretation. New York, NY: *Macmillan Publishing Company*.
[2] Hirschmuller. H., 2008: Stereo Processing by Semiglobal Matching and Mutual Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp.328-341.
[3] Zhang L., and Gruen A., 2006: Multi-image matching for DSM generation from IKONOS imagery. *ISPRS Journal of Photogrammetry & Remote Sensing*, vol. 60, pp.195-211.
[4] Licciardi G.A., Villa A., Mura M. D., Bruzzone L., Chanussot J., and Benediktsson J. A., 2012: Retrieval of the Height of Buildings From WorldView-2 Multi-Angular Imagery Using Attribute Filters and Geometric Invariant Moments. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, vol. 5, no. 1, pp. 71-79.
[5] Zeng C. Q., Wang J. F., Zhan W. F., Shi P. J., and Gambles A., 2014. An elevation difference model for building height extraction from stereo-image-derived DSMs. *International Journal of Remote Sensing*, vol. 35, no. 22, pp.7614-7630.
[6] Izadi M., and Saeedi P., 2012: Three-Dimensional Polygonal Building Model Estimation from Single Satellite Images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no.6, pp. 2254-2272.
[7] Zhu X. X., and Bamler R., 2010: Very High Resolution Spaceborne SAR Tomography in Urban Environment. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no.12, pp. 4296-4308.
[8] Ma P. F., Lin H., Lan H. X., and Chen F. L., 2015: On the Performance of Reweighted L_1 Minimization for Tomographic SAR Imaging. *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no.4, pp. 895-899.
[9] Guillaso S., Famil L. F., Reigber A., and Pottier E., 2005: Building Characterization Using L-Band Polarimetric Interferometric SAR Data. *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no.3, pp. 2386-2395.
[10] Colin-Koeniguer E., and Trouve N., 2014: Performance of Building Height Estimation Using High-Resolution PolInSAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5870-5879.
[11] Simonetto E., Oriot H., and Garello R., 2005: Rectangular building extraction from stereoscopic airborne radar images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 10, pp. 2386–2395.
[12] Guida R., Iodice A., Riccio D., 2010: Height retrieval of isolated buildings from single high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, vol.48, no.7, pp. 2967-2979.
[13] Brunner D., Lemoine G., Bruzzone L., and Greidanus H., 2010: Building Height Retrieval From VHR SAR Imagery Based on an Iterative Simulation and Matching Technique. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp.1487-1504.
[14] Auer S., Hinz S., and Bamer R., 2010: Ray-tracing simulation techniques for understanding high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1445-1456.
[15] Xu F., and Jin Y. Q., 2007: Automatic Reconstruction of Building Objects From Multiaspect Meter-Resolution SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no.7, pp.2336-2353.
[16] Sportouche H., Tupin F., Denise L., 2011: Extraction and Three-Dimensional Reconstruction of Isolated Buildings in Urban Scenes From High-Resolution Optical and SAR Spaceborne Images. *IEEE Transactions on Geoscience and Remote*

Sensing, vol. 49, no.10, pp. 3932-3946.
[17] Zhou G. Q. and Zhou X., 2014: Seamless Fusion of LiDAR and Aerial Imagery for Building Extraction. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no.11, pp. 7393-7407.
[18] Sampath A., and Shan J., 2010: Segmentation and Reconstruction of Polyhedral Building Roofs from Aerial Lidar Point Clouds. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no.3, pp. 1554-1567.
[19] Eckert S., and Hollands T., 2010: Comparsion of Automatic DSM Generation Modules by Processing IKONOS Stereo Data of an Urban Area. *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, vol. 3, no.2, pp. 162-167.
[20] Wegner J. D., Ziehn J. R., and Soergel U., 2014: Combing High-resolution Optical and InSAR Features for Height Estimation of Buildings with Flat Roofs. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5840-5854.
[21] Xu Y., Ma P. F., Ng E., and Hui L., 2015: Fusion of World-view2 stereo and TerraSAR-X images for building height extraction in urban areas. *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 8, pp.1-5.