Building height detection from high-resolution TerraSAR-X imagery and GIS data

Wen LIU and Fumio YAMAZAKI
Department of Urban Environment Systems
Chiba University
Chiba, Japan
Email: wen_liu@graduate.chiba-u.jp; fumio.yamazaki@faculty.chiba-u.jp

Abstract— **Due to the remarkable improvements that have been made in radar sensors, high-resolution SAR images are now available, thus providing detailed ground surface information. In this study, a new method is developed to detect building heights automatically using 2D GIS data and a single high-resolution TerraSAR-X (TSX) intensity image. A building in a SAR image shows a layover from the actual position in the direction of the sensor, due to the side-looking nature of SAR. Since the length of the layover is proportional to the height of the building, it can be used to estimate the building height. To do this, we shift the building shape obtained from 2D GIS data in the direction of the sensor. The proposed method was tested on a TSX image of San Francisco, U.S.A. in HighSpot mode, using a resolution of about 1 m. Comparing the result with that obtained using Lidar data, the RMS error was found to be less than 3 m, which is about the height of one building story.**

I. INTRODUCTION

In the last few years, data from the new high-resolution (HR) SAR sensors, e.g., Cosmo-SkyMed and TerraSAR-X (TSX), have provided us with more detailed surface information. The features of an individual urban structure, especially its height, have become available from these SAR images. The development of methods to automatically derive cartographic information from urban areas is a new and highly interesting area. Several methods related to height detection from HR SAR images have been proposed, and they can be divided into three categories.

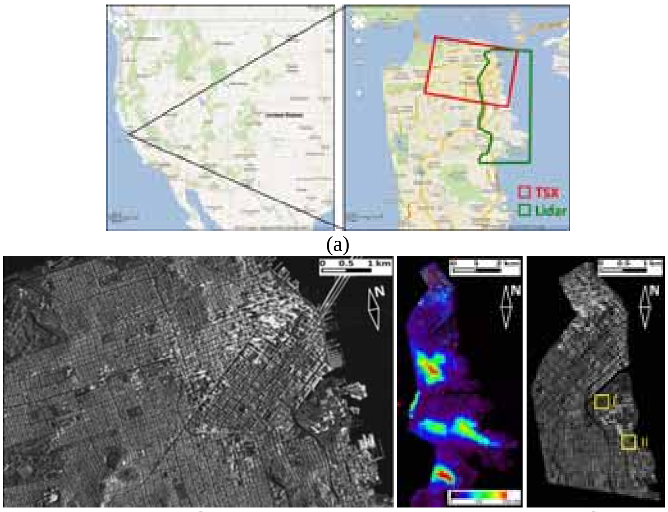
The first category is interferometric (InSAR) analysis. Bolter and Leberl [1] extracted buildings and their heights from multiple-view InSAR datasets. Thiel *et al.* [2] used two approaches to detect buildings of different sizes from multi-aspect HR InSAR datasets. Building features were extracted independently for each direction from the amplitude and phase information in the InSAR data. These methods, however, require at least two SAR images in different flight angles. The second category is based on a direct electromagnetic backscattering model. Franceschetti *et al.* [3] proposed a method for extracting the building height information based on a radiometric analysis of the double-bounce contribution. They tested this method upon an airborne SAR sensor image [4], and Liu *et al.* [5] applied the method to a TSX image of Shanghai, China. This method requires prior knowledge of the material and surface roughness properties of the background, and it cannot be used in different areas. The third category is based on the geometrical characteristics of the layover and radar-shadow areas. Xu and Jin [6] proposed a method for the automatic reconstruction of building objects from multi-aspect very HR SAR data. Brunner [7] estimated building heights by matching simulated SAR images with a real image containing buildings of various heights. However, these methods were complicated and time consuming.

In this study, a simple method is proposed for detecting the heights of individual buildings by estimating their layover lengths from a single HR SAR image. The outline of a building taken from 2D GIS data is needed to create a template on a SAR image. The template is then shifted in the direction of the sensor step by step. In each step, the inside area of the template is calculated to distinguish the boundaries of the layover. This method was applied to a TSX image of San Francisco, California, U.S.A. In addition, Lidar data was introduced as a reference to verify the accuracy of the proposed method.

II. STUDY AREA AND IMAGE DATA

The study area focuses on the bay coast of San Francisco, CA, USA, as shown in Fig. 1(a). A TSX image taken on October 2, 2011, which is shown in Fig. 1(b), was used to detect building heights. The image was taken by the HH polarization in the descending path. Since the image was taken in the HighSpot Mode (HS), the azimuth resolution was about 1.10 m and the ground range resolution was about 0.92 m. The image was recorded as a single-look slant-range complex (SSC) product. A digital surface model (DSM) taken by a Lidar flight in June 2010, which is shown in Fig. 1(c), was also used as a reference for the building heights. The DSM data was resampled as 0.5 m/pixel.

Several pre-processing steps were carried out on the images prior to the detection of building heights. First, the TSX image was transformed into a geocoded ground range amplitude image using *SARscape* software. Since the digital elevation model (DEM) extracted from the Lidar data was used for geocoding, the processed TSX image was resampled as 0.5 m/pixel, the same pixel size as that of the DSM image obtained from the Lidar data. Next, the image was transformed into a Sigma Naught (σ^0) value, which represents the radar reflectivity per unit area in the ground range, following the radiometric calibration process. An enhanced Lee filter [9], one



(a)

(b) (c) (d)

Figure 1. Study area on the bay coast of San Francisco, USA (a); the TSX image taken on Oct. 2, 2011 (b); the DSM taken by Lidar in June 2010 (c); the pre-processed Sigma Naught TSX image (d).

of the most commonly used adaptive filters, was used to reduce the speckle noise in the original SAR image, which makes the radiometric and textural aspects less efficient. To prevent losing the information included in the intensity images, the window size of the filter was set as 3 × 3 pixels (about 1.5 m × 1.5 m). The TSX image after all the pre-processing steps is shown in Fig. 1(d).

III. METHOD OF HEIGHT DETECTION

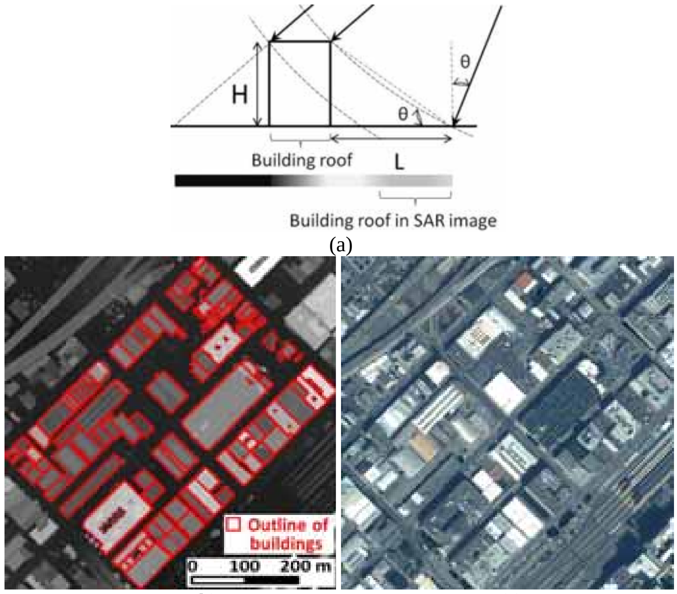
Due to the side-looking nature of SAR, a building in a TSX image shows a layover from the actual position, in the direction of the sensor, as shown in Fig. 2(a). The layover is proportional to the building height, as expressed in (1).

$$L = H / \tan\theta \tag{1}$$

where θ is the incident angle of the radar for the TSX image.

Since there is no available 2D GIS data, the outline of buildings was made manually in this study, as shown in Fig. 2(b), based on the DSM data. An optical image taken on October 9, 2011 from WorldView-2 (WV-2) was used as support information to divide buildings of similar height that stand in a row; the image is shown in Fig. 2(c) The outline of 116 buildings in the target area was extracted to detect their heights. Fig. 3(a) shows the outline overlapping on the TSX image. It is obvious that the high backscatter areas due to the walls were away from the building outline. If the length of the layover along the sensor direction is detected, the height of this building can be calculated using (1).

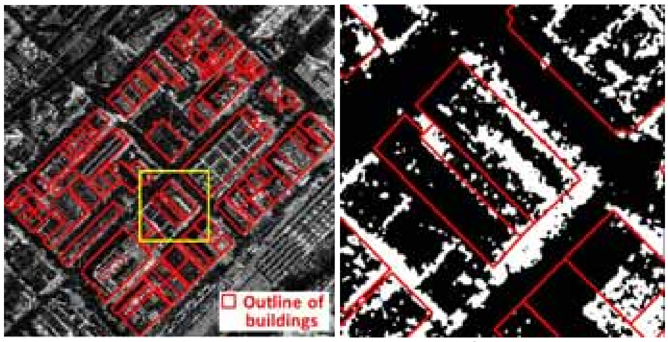
To detect the layover length, a template was created based on the building outline and was then shifted in the direction of the SAR sensor (southeast). Since a typical building is usually higher than 2 m, 2 m was set as the initial height for searching a layover area. Due to the 39.88° incidence angle and the 187.79° path angle (clockwise from the north), a 2-m building shows a 2.36-m layover to the east and a 0.41-m layover to the south. The width of the layover increases 1.18 m to the east and 0.20 m to the south, for a 1-m increase in



(a)

(b) (c)

Figure 2. Simulation of the location of a building in a SAR image (a); the outline of overlapping building shapes from the DSM data (b); the WorldView-2 image taken on Oct. 9, 2011(c) in area I of Fig. 1(d).

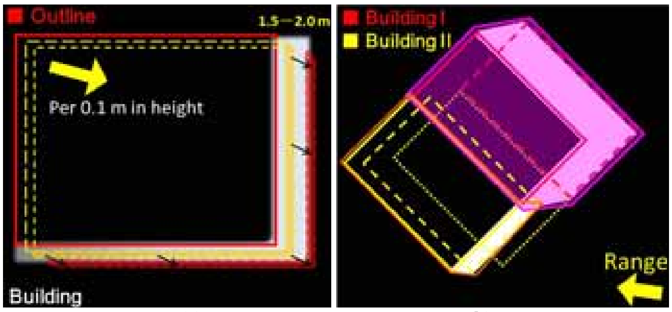


(a) (b)

Figure 3. Outline of overlapping buildings in the TSX image (a), and a close-up of the yellow frame of the 2-binary image (b).

building height. Considering the resolution and pixel size of the TSX image, the width of the template was set as a 0.5-m building height. Thus, the initial template for the layover of a wall with a height between 2.0 m and 2.5 m was created, as shown in the yellow block in Fig. 4(a). Since most of the layover occurs on a street area, the area within the building outline was removed from the template. Thus, the shape of the template depends on the size of the building and on the surrounding conditions.

A building of 5 × 5 m^2 would have a template of about 3.5 m^2. However, in the case of buildings standing in a row, the layover area is seen only in one direction and is relatively small. To confirm that there is enough information within a template, a single template should be larger than 1.0 m^2. If there is not a large enough initial template, as in the case that it is enclosed by other buildings, layover searching will not begin. Next, the created template is shifted in the direction of the sensor step by step, by increasing the building height at intervals of 0.1 m. To improve the accuracy of the height detection, using cubic convolution, the area in the TSX image surrounding a building was resampled to 0.1 m, which is 1/5 of the initial pixel size.



(a) (b)

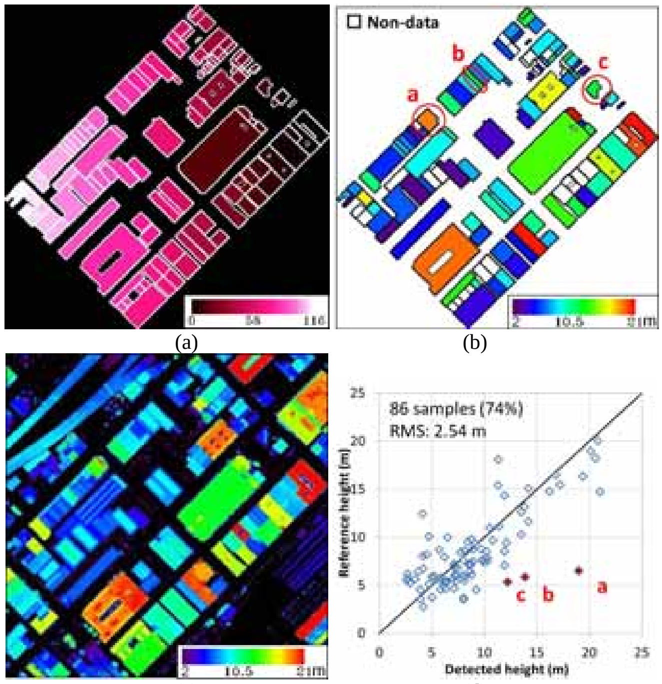
Figure 4. Simulation of the template for layover searching (a), and boundary definition for two connected buildings with different heights (b).

To estimate the boundary of the layover, two methods are used to define the threshold values. The first method is used to define the threshold value for the backscatter intensity. The average value of the backscatter in a TSX image within the template is calculated. When it is lower than the threshold value, the template is classified as a street and the shifting is stopped. However, this method will show errors when there is an outstanding bright or dark object within the template. The second method is used to define the threshold value for the areal percentage. In this method, the TSX image is first divided into a binary image using one threshold value. High backscatter areas such as corner reflections and the layover of walls are transformed as 1, whereas low backscatter areas such as radar shadows and roads are 0. Next, the percentage of high backscatter areas within a template is calculated. When the percentage is lower than the threshold value, the template is classified as a street. In this study, the latter method was selected. The TSX image was divided by the average backscatter intensity of the entire image, which is -10.7 dB. A part of a 2-binary image is shown in Fig. 3(b). Next, the threshold value for the areal percentage was set at 0.4, which represents less than half of the template.

When there are two buildings of different heights with no space between them, their layover areas, which have different lengths, are also connected, as shown in Fig. 4(b). If two templates of buildings shift at the same time, the template for Building II, indicated by a yellow dotted line, will search both layover areas. Then, the height of Building II cannot be detected correctly. Thus, the order of the layover search is important. In this study, to confirm that template shifting was performed in the correct order, an ID number was assigned to each building. A smaller ID number was assigned to buildings located closer to the direction of the sensor. Next, the template was created and shifted in the order of ascending ID numbers. Once a layover area has been searched, it is masked and is not searched again. As seen in the case of Fig. 4 (b), the layover area of Building I was searched first and the final boundary is indicated by a red dashed line. Next, the area of Building I and its layover area were masked when searching for Building II. Finally, the boundary of Building II was detected correctly; this boundary is indicated by a yellow dashed line.

IV. DETECTED RESULT AND ITS VERIFICATION

The proposed method was applied to areas I and II, which are shown in Fig. 1(d). In the previous section, the outline of 116 buildings within target area I was created manually from the DSM data. They were then numbered, based on their distance from the SAR sensor, as shown in Fig. 5(a). The layover areas were searched in the order of their ID numbers, and their lengths were used to calculate the building heights. The maximum height from the layover search was set as 24 m, which is equal to the height of an eight-story building (the height of one story is counted as 3 m). The heights of 89 buildings were detected successfully, and the results are shown in a rainbow array of colors in Fig. 5(b). Some buildings could not be detected because of the surrounding conditions. Most of these buildings were located behind other buildings, so there was no space for the creation of a template.



(a) (b) (c) (d)

Figure 5. Buildings were assigned ID numbers according to their distance from the SAR sensor (a); detected building height (b) and the building heights obtained from Lidar data, which were used as reference (c); comparison of the heights in the results vs. the reference heights (d).

The DSM and DEM data obtained by Lidar were used as a reference. The building heights can be obtained by subtracting the DEM data from the DSM data; the results of this are shown in Fig. 5(c). Comparing Fig. 5(b) and (c), we see that most of the buildings are shown in a similar color, which means they have similar heights. Since the layover length is proportional to the wall height in the direction of the SAR sensor, the height of the boundary in the south-east direction, instead of the height of the entire building, was used as a reference. The average height of the boundary within a 2-pixel width of the building outline was calculated.

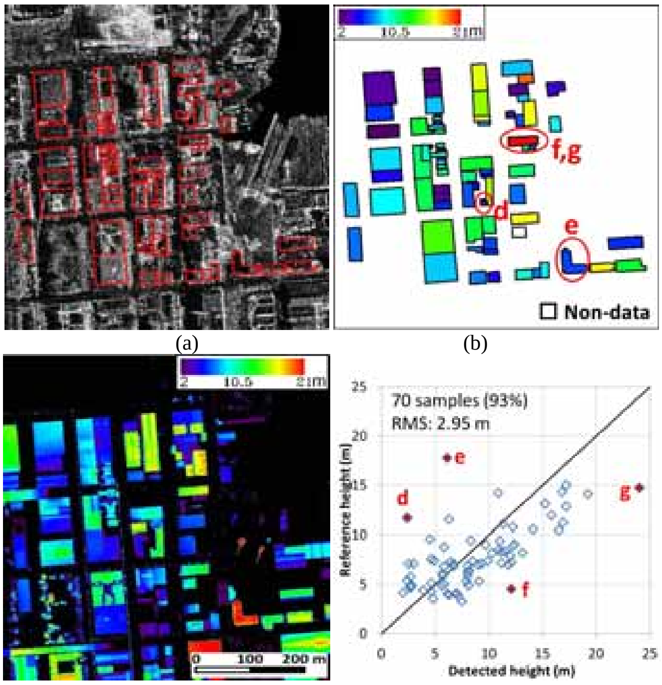
A comparison of the detected heights and the reference heights is shown in Fig. 5(d). In the area shown, 78% of the buildings were less than 10 m. Most of the points are around the isometric line, except for three points that are shown in red. The locations of these three buildings are shown in Fig. 5(b). The street view mode of Google Earth was used to estimate the reason for these errors. From the street view around building *a*, it could be seen that a high billboard is mounted on the target building *a*. Since the billboard faces in the direction of the sensor, the layover areas included the reflections from the wall and billboard, which caused our result to be higher than the reference value. The error for building *c* was caused for the same reason, as can be confirmed in the ground photo. The error for building *b* was caused by a low wall in front of the building. Since the low wall was located in the same direction as the layover, the reflection of the low wall was counted in the layover searching step. Thus, the detected result was higher than the reference value. However, the root mean square (RMS) error for the remaining 86 buildings was 2.54 m, which is less than the average height of one story (3 m).

The building outlines in area II were also created manually from the DSM data and the optical image. Since it was confirmed in area I that the height of a building that stands behind others could not be detected, the outlines of these buildings were not created. Fig. 6(a) shows the outlines of 75 buildings that are overlapping in the TSX image. The heights of 74 buildings were detected by shifting the template, and the result is shown in Fig. 6(b). A building could not be detected because of the presence of a building that was built after the Lidar data was produced. The building heights obtained from the Lidar data are shown in Fig. 6(c).

The average height of the boundary in the sensor direction was calculated for each building and was then used as a reference value. A comparison of the detected results and the reference values is shown in Fig. 6(c). The detected heights of 70 buildings were close to the isometric line. The errors for buildings *e* and *d* were caused by the complicated conditions of their surroundings, which were confirmed in the ground photos. The errors for buildings *f* and *g* were caused by construction. There were several buildings south of buildings *g* and *f* that were built after the Lidar data acquisition. Thus, the reflection of the newly built buildings was connected with that of the target buildings, thus causing excessive height detection. Except for these four buildings, the RMS error for the 70 buildings in area II was calculated as 2.95 m. Compared with the results for area I, more buildings were detected as being higher than they really were. These errors were caused by the trees in front of these buildings. Since the reflection from trees was counted as the layover of buildings in some cases, more trees caused more errors in the detected results.

V. CONCLUSIONS

In this study, a method to detect the height of an individual building from a high-resolution TerraSAR-X image was proposed. The method was tested using a 1-m resolution SAR image of San Francisco. The outlines of 191 buildings were created manually from the DSM data obtained from Lidar, and they were overlapped in the TSX image to detect the building heights. As a result, the heights of 156 buildings were detected successfully. The building heights obtained from the DSM and DEM data were used as a reference to verify the accuracy of the proposed method. The average difference between the detected results and the reference values was 0.67 m, and the RMS error was 2.73 m, which is less than the height of a single story. It was found that the accuracy of the proposed method depends on the conditions surrounding the target building.



(a) (b)

(c) (d)

Figure 6. Building outlines overlapping in the TSX image (a) in area II of Fig .1(d); building height results shown in a rainbow array of color (b); the reference building heights obtained from Lidar data (c); comparison of the detected result versus the reference values (d).

However, the results show that our method is useful for estimating the number of building stories. In the future, the method will be tested on more TSX images of different locations to verify its accuracy and to investigate its limitations.

ACKNOWLEDGMENT

The TerraSAR-X images, WorldView-2 images, and Lidar data used in this study were provided by the 2012 IEEE Geoscience and Remote Sensing Society Data Fusion Contest.

REFERENCES

[1] R. Bolter, and F. Leberl, “Detection and reconstruction of human scale features from high resolution interferometric SAR data,” In: *Pro. 15th Int. Conf. Pattern Recog.*, Vol. 4, 2000, pp. 91–294.

[2] A. Thiele, E. Cadario, K. Schulz, U. Thoennessen, and U. Soergel, “Building recognition from multi-aspect high-resolution InSAR data in urban areas,” *IEEE Trans. Geosci. Remote Sensing*, Vol. 45, No. 11, 2007, pp. 495–505.

[3] G. Franceschetti, A. Iodice, and D. Riccio, “A canonical problem in electromagnetic backscattering from buildings,” *IEEE Trans. Geosci. Remote Sens.*, Vol. 40, No. 8, 2002, pp. 1787–1801.

[4] G. Franceschetti, R. Guida, A. Iodice, D. Riccio, and G. Ruello, “Building feature extraction via a deterministic approach: Application to real high resolution SAR images,” In: *Proc. IEEE IGARSS*, 2007, pp. 2681–2684.

[5] K. Liu, T. Balz, and M. Liao, “Building height determination by TerraSAR-X backscatter analysis in dense urban areas,” In: *Proc. 2009 ACRS*, 2009.

[6] F. Xu, and Y.-Q. Jin, “Deorientation theory of polarimetric scattering targets and application to terrain surface classification,” *IEEE Trans. Geosci. Remote Sens.*, Vol. 43, No. 10, 2005, pp. 2351–2364.

[7] D. Brunner, and G. Lemoine, “Building height retrieval from VHR SAR imagery based on an iterative simulation and matching technique,” I*EEE Trans. Geosci. Remote Sens.*, Vol. 48, No. 3, 2010, pp. 1487–1504.

[8] A. Lopes, R. Touzi, and E. Nezry, "Adaptive Speckle Filters and Scene Heterogeneity", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, No. 6, 1990, pp. 992–1000.