Urban monitoring and change detection of central Tokyo using TerraSAR-X images

Wen Liu¹) and Fumio Yamazaki¹)

¹) Department of Urban Environment Systems, Chiba University, Japan

ABSTRACT
Monitoring urban growth and change is an important issue for urban planning and disaster management. In this study, two different approaches are conducted to monitor the urban areas. One is that a TerraSAR-X image is compared with an ALOS/PALSAR image to investigate the characteristics of urban land cover in different radar wavelengths. Also, a very high resolution optical QuickBird image is used as the ground truth. The other is that two temporal TerraSAR-X images are used to detect urban changes. Difference and correlation coefficient between two images are calculated with a sliding window. A new factor that composites difference and correlation coefficient is proposed to detect the changed areas. An aerial photograph is introduced to verify the accuracy of detection results.

INTRODUCTION
Urban grows and changes rapidly all over the world. Monitoring urban areas has become an increasingly important topic. Regular and up-to-date information on urban change is required for strategic city planning purposes, environmental impact assessment, and for the appropriate allocation of services and infrastructure within towns and cities [1-3]. Current approaches to urban monitoring generally involve ground surveys and interpretation of aerial photographs. The improvement of remote sensing technology makes satellite image become an efficient method to collect urban information. Compared to optical sensors, synthetic aperture radar (SAR) does not suffer from the limitations of weather condition. The high resolution of images (ALOS/PALSAR and TerraSAR-X) is well suited for urban applications, e.g. urban modeling as well as for damage mapping after disasters.

Urban areas are generally characterized by high SAR backscatter intensities due to the predominance of single- and double-bounce backscattering [4]. Previous studies of using multi-temporal SAR data for change detection have achieved some promising results. Rignot et al. [5] used the backscattering intensity ratio and the coherence value from repeat pass ERS-1 SAR data to identify scene changes. Grey et al. [6] used a multi-temporal sequence of ERS interferometric coherence data acquired between 1993 and 1999 for automated mapping of urban changes in South Wales, UK. Liao et al. [7] detected land-cover changes in urban area based on backscattering intensity and long-term coherence from ERS-1/2 InSAR data. These studies focused on using ERS interferometric coherence data to detect the land cover changes in urban area. However, TerraSAR-X image has higher resolution (1.25 m) to detect more details, e.g. a single building
change. Matsuoka et al. [8] developed an automated method to detect hard-hit areas in the 1995 Kobe Earthquake, using the backscattering coefficient difference and correlation between pre- and post-event ERS images. This method can also be used to detect urban changes in case of TerraSAR-X image.

In this study, an L-band ALOS/PALSAR image and an X-band TerraSAR-X image of central Tokyo are used to investigate the characteristics of urban areas in different radar wavelengths. Then two temporal TerraSAR-X images are used to detect the urban changes by an improvement method proposed by Matsuoka et al. [8]. The result of change detection will be compared with an optical QuicBird (QB) image and an aerial photograph to verify the accuracy.

STUDY AREA AND DATA

The study area focuses on central Tokyo, Japan, shown in Fig. 1, since there are many new constructions every year in this area. Two TerraSAR-X images taken in different years were used to detect urban changes, shown in Fig. 2 (a, b). The first image was taken on May 23\textsuperscript{th}, 2008 (UTC) with 42.82\degree incidence angle, and the second one was taken on November 23\textsuperscript{th}, 2009 with 42.81\degree incidence angle at the center. Both images were taken by HH polarization, and in the descending path. Since they were taken in StripMap Mode [9], the azimuth resolution was about 3.3 m and the ground range resolution was about 1.2 m. After the enhanced ellipsoid correction (EEC), two images were transformed into 1.25 m/pixel. Since a SAR image has only one band, it is difficult to collect ground details just using one single polarization SAR image. A QB image was introduced in this study as the ground truth data, shown in Fig. 2 (c). The QB image was taken on Match 20\textsuperscript{th}, 2007, one year before the first TerraSAR-X image. After pansharpening, a 0.6 m resolution 4-band (B, G, R, NIR) image was obtained.

To investigate the characteristics of urban area in SAR images, a comparison of L- and X-band images was carried out. An ALOS/PALSAR image obtained from GEO Grid [10] was used, shown in Fig. 2 (d). The PALSAR image was taken on November 27\textsuperscript{th}, 2009, with L-band. Since it was taken in the same month of the second TerraSAR-X image, urban ground condition in two images can be considered as the same. The incident angle of PALSAR image was 21.2\degree, also in the descending path. The resolutions to the azimuth and range directions were both 12.5 m, and the product level was Multi-look Grand Range Geocode (MGG). Different from the TerraSAR-X image, the PALSAR image was already transformed into dB unit.

IMAGE PREPROCESSING

Since the four images used in this study were from three different types of sensors, several image pre-processing steps were needed. The first step is co-registration. Considering the resolution and product level, the TerraSAR-X images were used as the base image. Eight points chosen manually from each image were used to match the PALSAR image and the QB image to the TerraSAR-X images. The resolution of QB image which was higher than TerraSAR-X image was deteriorated to 1.25 m after co-registration. Then all
images can be matched with the spatial errors less than 1 pixel.

To compare two different SAR images, two TerraSAR-X images and the PARSAR image were transformed to Sigma Naught ($\sigma^0$), which represents the radar reflectivity per unit area in the ground range. According to Infoterra [11], TerraSAR-X images can be transformed to Sigma Naught from a digital number by Eq. (1).

$$\sigma^0 = 10 \cdot \log_{10} (k_s \cdot |DN|^2) + 10 \cdot \log(\sin \theta_{loc})$$

where $k_s$ is the calibration factor which can be found in the header file, and $\theta_{loc}$ is the local incidence angle which is derived from the Geocoded Incidence Angle file.

From the PALSAR image, the following calculation was required as radiometric calibration to transform Beta Naught ($\beta^0$) to Sigma Naught [10].

$$\sigma^0 = \beta^0 - 20.67[\text{dB}]$$

There were many speckle noises in the original SAR images, which make the radiometric and textural aspects less efficient. Hence, an adaptive speckle filter was applied. In this study, Lee filter [12], one of the most common adaptive filters, was used. Considering the building size and the image resolution, the window size
of Lee filter was set as 9×9 pixels (about 11 m×11 m), similar to a small building’s size. The filter was applied on the two TerraSAR-X images and the PALSAR image.

CHARACTERISTICS OF URBAN AREA IN SAR IMAGES

A comparison of X-band and L-band images was carried out using the TerraSAR-X image and the PALSAR image taken in November, 2009. Since the incident angles of the two images were very different, it is difficult to compare them directly. Most differences between them were occurred due to the incident angle and the resolution, not the radar wavelength. However, an unsurprised classification was applied on two images to compare their difference in the backscattering coefficient. Both the TerraSAR-X image and the PALSAR image were classified by the k-means method into 7 classes, shown in Fig. 3 (a, b). The first two classes with the lowest backscattering coefficient were considered as water, roads and shadow. The third class was considered as bare ground. The fourth and fifth classes were considered as grass and trees. The last two classes with the highest backscattering coefficient were built-up areas.

Comparing the results for the TerraSAR-X and PALSAR images, the border line of ground and water can be classified clearly for the TerraSAR-X image but difficult for the PALSAR image. Several bridges were classified into water class in the PALSAR image. Almost all the built-up areas can be classified correctly for the TerraSAR-X image, but only big scale buildings can be classified from the PALSAR image. Using these characteristics, an experiment was carried out to classify an image which composites the TerraSAR-X and PALSAR images as 2 bands. Then the result of classification was shown in Fig. 3 (c). Since this image includes the information of both X- and L-bands, the bridges and ground can be separated from water. Almost all the built-up areas can be classified correctly, and the different scale buildings can be separated into different classes. Recent researches have proved that the polarimetric information of SAR can be used to classify urban land cover with high accuracy [13-15]. According to the result of this experiment, the different
wavelength information is also useful for land cover classification.

**CHANGE DETECTION FROM TERRASAR-X IMAGES**

After image pre-processing, two TerraSAR-X images in Sigma Naught, taken in different years were used to detect urban changes in this period (1.5 years). Firstly, color composition was applied to see changed areas visually, shown in Fig. 4 (a). The first image taken in 2008 is plotted in Red, and the second image taken in 2009 is plotted in Blue and Green. From Fig. 4 (a), the red area represents the reduction of reflection, which can be considered as removed buildings; the cyan area represents the increase of reflection, which can be considered as new constructions. Many urban changes can be seen in Toyosu, a new re-development area near the Tokyo Station. A new bridge can also be seen clearly in cyan from the color composite image.

To detect these changes automatically, the difference of the backscattering coefficients and the correlation coefficient were derived from the two images. This method was employed by Matsuoka et.al [8] to detect changes due to natural disasters in urban areas. In our case, the same method was applied to urban changes in a normal time. The approach of change detection was shown in Fig. 5. The difference \( d \) was calculated by Eq. (3) and the correlation coefficient \( r \) was calculated by Eq. (4).

\[
d = I_{a_i} - I_{b_i}
\]

\[
r = \frac{N \sum_{i=1}^{N} I_{a_i} I_{b_i} - \sum_{i=1}^{N} I_{a_i} \sum_{i=1}^{N} I_{b_i}}{\sqrt{\left(N \sum_{i=1}^{N} I_{a_i}^2 - \left(\sum_{i=1}^{N} I_{a_i}\right)^2\right) \left(N \sum_{i=1}^{N} I_{b_i}^2 - \left(\sum_{i=1}^{N} I_{b_i}\right)^2\right)}}
\]

where \( i \) is the pixel number, \( I_{a_i} \) and \( I_{b_i} \) are the backscattering coefficient of the post- and pre-event images, \( \bar{I}_{a_i} \) and \( \bar{I}_{b_i} \) are the corresponding averaged values over the \( N (=k \times k) \) pixels window surrounding the \( i \)-th pixel.

In this study, the window size is set as 9x9 pixels, the same as speckle filter’s. In the resulted image of difference, the urban changes can be shown by either negative or positive values. The maximum change in the backscattering coefficient was about ±30 dB. In the resulted image of correlation, the low correlation value means a change. However, even there was no change, vegetation and water areas showed very low correlation since the reflection of them would change much depending on the wind condition. To remove the effect of water, a water borderline data from ESRI GIS Data Collection [16] was introduced to mask water areas from the correlation coefficient image.

Firstly, the changed areas were extracted from the difference and correlation coefficient images by the respective threshold values. According to the histogram of the difference, the mean value \( (\mu_d) \) was -0.94 and the standard deviation \( (\sigma_d) \) was 2.44. The mean value of the correlation coefficient \( (\mu_r) \) was also calculated...
from the histogram as 0.41 while the standard deviation ($\sigma_r$) was 0.37. To keep the objectivity of the detected results, the threshold values were calculated by Eq. (5), using the mean value and standard deviation.

$$v = \mu \pm 2 \cdot \sigma$$  \hspace{1cm} (5)

In this study, the threshold values for the difference were set as -5.82 and 3.94, which means the areas with the difference less than -5.82 or larger than 3.94 were considered as a change. From the correlation, the areas with the value less than 0.33 were considered as changes.

To remove noises from the detected result, a filter was applied. Considering the size of buildings, the extracted areas which are smaller than $8 \times 8$ pixels ($10 \text{ m} \times 10 \text{ m}$) are removed as noises. Since the difference

---

Fig. 4 Color composition of the two TerraSAR-X images (a), and the extracted change areas by the difference and correlation superposed on the QuickBird image (b).

Fig. 5 Flowchart of change detection from two temporal TerraSAR-X images
and correlation were calculated by a moving window, the extracted areas were smaller than the original areas. Thus a buffer was introduced to enlarge the extracted areas. The pixels within 1/2 of the window size around the extracted areas were counted into the change areas. The result of detected change areas from the difference and correlation is shown in Fig. 4 (b). Comparing Fig. 4 (a) and (b), the changes of buildings could be extracted mostly. However, the results of the detection from the difference and the correlation were not completely matched at some locations. Since the algorithms of calculating difference and correlation in the same window are different, a change of a building appeared in different locations in the difference and correlation calculations. There were also a lot of errors resulting from low correlation, e.g., grasses or rough ground. Then a factor \( z \) calculated by Eq. (6), which combines the difference and the correlation coefficient was proposed to represent changes generally.

\[
    z = \frac{|d|}{\text{Max}(|d|)} - c \cdot r
\]

where \( \text{Max}(|d|) \) is the maximum absolute value in difference and \( c \) is the weight between the difference and the correlation coefficient.

Since the correlation was very sensitive about subtle changes, it shows a low value even there was no big change. On the contrary, the normalized absolute value of difference is relatively stable. Hence, in this study the weight of difference was set as 4 times of that of correlation. The result of the factor \( z \) was between 0.75 and 1.25, shown in Fig. 6 (a). The high value means the high possibility of change. Comparing with Fig. 5 (a), colored areas which means change can be seen easily by red color (high value) from Fig. 6 (a). According to the histogram of the \( z \) image, the mean value \( (\mu) \) was -0.09 while the standard variation \( (\sigma) \) was 0.12. The threshold value was also calculated by Eq. (4) to detect changed areas. In this case, the areas with value larger than 0.15 were extracted as the changed areas.

Then the extracted areas were overlapped on the result of difference, to divide it into positive and negative changes. Finally, the result of change detection was shown in Fig. 6 (b), overlapping on the QB image. There were 2.8% of the whole TerraSAR-X image that changed between 2008 and 2009, including 1.6% newly built and 1.2% removed constructions. Converting to area unit, about 6.1 km\(^2\) areas have new constructions while the removed constructions correspond to 4.6 km\(^2\). Since the reflection of buildings in SAR images are from both walls and roofs, mainly from walls in case of high buildings, the areas of change from TerraSAR-X images were not exactly the areas of change in ground level.

Toyosu area was a new developing area, with a lot of new constructions. Hence, this part was picked up, shown in Fig. 7 (a, b). The optical image in (a) is the QB image taken in March, 2007, and in (b) is an aerial photograph taken in April, 2009. Comparing (a) and (b), the extracted areas were mostly matched with the changes in these optical images. The detected changes over water areas are considered as ships. It should be noted that several areas extracted as the remove constructions were occurred due to the shadow of a new building nearby. In the future, this kind of errors would be removed by considering a new building and
removed buildings as one set when they are in neighborhood and lined in the direction of radar scattering. Since the dates of the optical images are not the same as those of TerraSAR-X images, several detected results cannot find the reference from the optical images.

CONCLUSIONS

In this study, a PALSAR image and a TerraSAR-X image were compared by unsurprised classification to
investigate the characteristics of different radar wavelength. Then the urban changes in central Tokyo were detected from two temporal TerraSAR-X intensity images, by a new factor combining the difference and correlation of the two SAR images. A pansharpened QB image and an aerial photograph were introduced to compare with the extracted results. Almost all the change buildings could be extracted correctly by the proposed method.

ACKNOWLEDGEMENT

The TerraSAR-X images used in this study were made available from SAR Application Research Committee, organized by PASCO Corporation, Tokyo, Japan. The PALSAR image was provided from GEO Grid, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan.

REFERENCE