Paper:

Extraction of Urban Information for Seismic Hazard and Risk Assessment in Lima, Peru Using Satellite Imagery

Masashi Matsuoka*, **Hiroyuki Miura****, **Saburoh Midorikawa***, **and Miguel Estrada*****

*Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology
Nagatsuta 4259-G3, Midori-ku, Yokohama 226-8502, Japan
E-mail: matsuoka.m.ab@m.titech.ac.jp, smidorik@enveng.titech.ac.jp
**Graduate School of Engineering, Hiroshima University
1-4-1 Kagamiyama, Higashi-Hiroshima 739-8527, Japan
E-mail: hmiura@hiroshima-u.ac.jp
***Japan-Peru Center for Earthquake Engineering and Disaster Mitigation (CISMID), National University of Engineering
Av. Túpac Amaru 1150, Lima 25, Peru
E-mail: estrada@uni.edu.pe

[Received November 26, 2012; accepted December 14, 2012]

Lima City, Peru, is, like Japan, on the verge of a strike by a massive earthquake. Building inventory data for the city need to be created for earthquake damage estimation, so the city was subjected to the extraction of spatial distribution of building age from Landsat satellite time-series images and an assessing building height from ALOS/PRISM images. Interband calculation of Landsat time-series images gives various indices relevant to land covering. The transition of indices was evaluated to clarify urban sprawl taking place in the northern, southern, and eastern parts of Lima City. Built-up area data were created for buildings by age. The height of large-scale mid-to-high-rise buildings was extracted by applying spatial filtering for a DSM (Digital Surface Model) generated from stereovision PRISM images. As a result, buildings with a small square measure, color similar to that of their surroundings, or complicated shapes turned out to be difficult to detect.

Keywords: building inventory data, earthquake damage estimation, urban sprawl, building height, landsat image, ALOS/PRISM, Lima City

1. Introduction

In the assessing building damage caused by an earthquake, building inventory data for the target area need to be provided in addition to the evaluation of possible earthquake ground motion and seismic building performance. Building inventory data are usually created through field surveys or visual interpretation using aerial photographs. Such work requires great effort and cost. In Japan, detailed building inventory data are usually provided and utilized for earthquake damage estimation by central and local governments.

In developing countries, however, rapid urban sprawl has been taking place and building inventories thus need to be provided. Building inventory data suitable to the current situation are not sufficient, however, discouraging proper earthquake damage evaluation. Peru, one of the developing countries of interest here, is situated in an area where an ocean floor plate subducts a trench near land, similar to Japan, and a massive earthquake is predicted as highly likely to occur. Although earthquake damage evaluation is an urgent need in effective disaster response and disaster mitigation, it is, at present, based on insufficient building inventory data.

The number and distribution of buildings damaged by a possible earthquake are estimated by modeling the ground condition and earthquake source, calculating earthquake ground motion, calculating the damage rate and distribution of earthquake ground motion calculated based on building seismic performance, and multiplying the damage rate by building inventory data information from the target area. Accordingly, building inventory data are important for conducting damage estimation, with information such as structure type, age, and the number of building stories as attributes. Information on age is required because older buildings are likely to have been constructed under older seismic standards that may not require sufficient seismic capability. Even with the same structure type, low-rise buildings and mid-to-high-rise buildings have different deformation performance and seismic capability, and thus information on the number of building stories is important.

Regarding building inventory data, building information on age and the story number are gathered efficiently by using remote sensing data from satellites. In recent years, satellite image performance has improved enough to allow images with a ground resolution of 1 meter or less to be acquired, thereby enabling assessment on a building-by-building basis. With the intention of creating building inventory data in Metro Manila, for example, Miura et al. (2006) have proposed a method in which mid-to-high-rise buildings are automatically detected using building shadows appearing in IKONOS satellite images and ex-

Table 1. Characteristics of Landsat images used in this study.

Satellite	Sensor	Date	Num. of Bands	Resolution
Landsat-5	Thematic Mapper (TM)	1987/3/5	7 bands	30 m
Landsat-5	Thematic Mapper (TM)	1998/5/6	7 bands	30 m
Landsat-5	Thematic Mapper (TM)	2002/1/17	7 bands	30 m
Landsat-5	Thematic Mapper (TM)	2006/5/12	7 bands	30 m

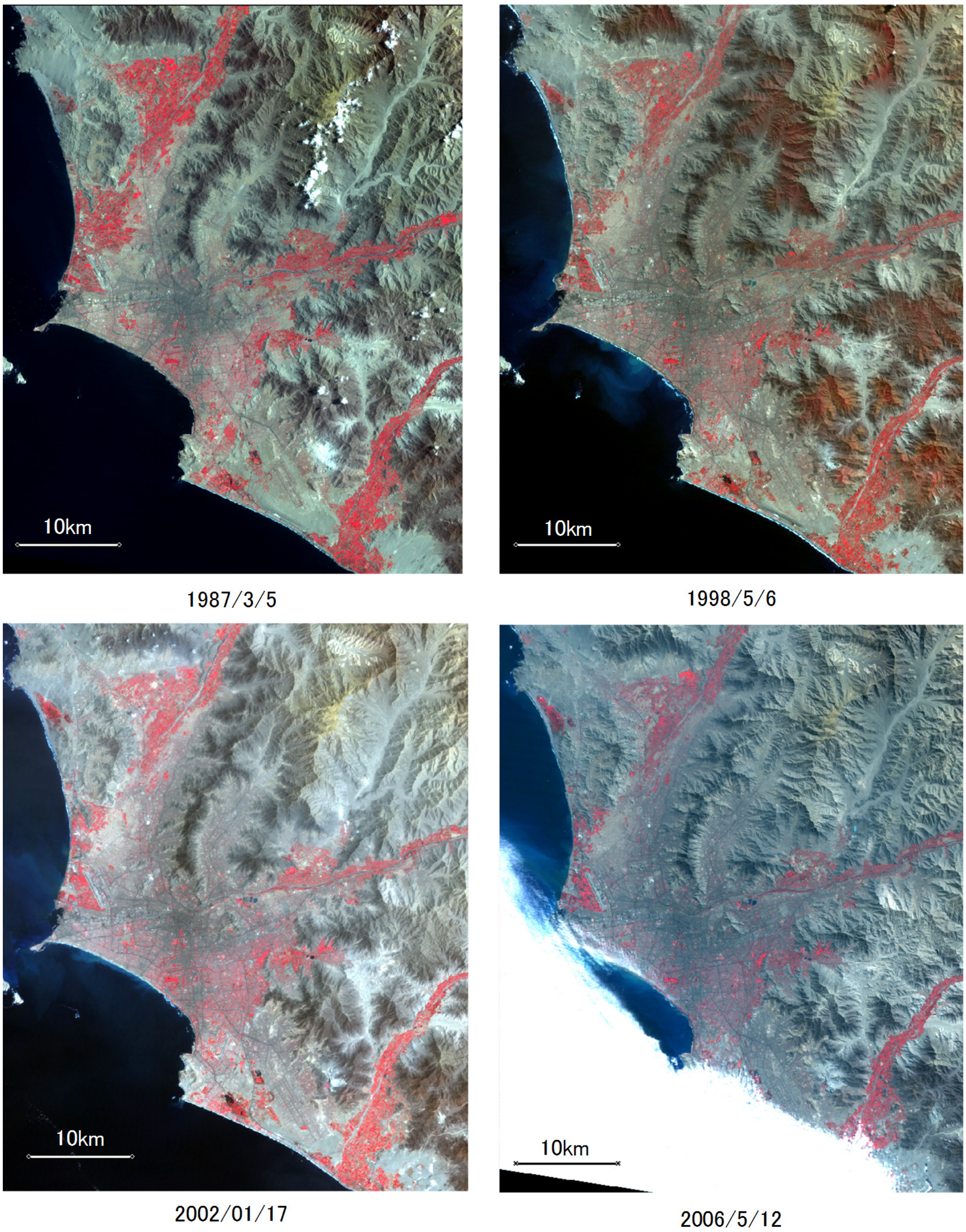

Fig. 1. Landsat images observed in 1987, 1998, 2002 and 2006.

isting building inventory data are updated [1]. Another attempt in which urban sprawl transition in Metro Manila is assessed was based on the land cover classification of time-series Landsat imagery and urban area distribution by age was assessed overall [2]. An assessment of building height has been conducted by Shaker et al. (2011) to stereo-pair images from the IKONOS satellite where high-density residential areas in Cairo, Egypt, were observed [3].

As seen above, various remote sensing images have been used to create building inventory data. None of them have been applied, however, to the situation in Peru. Our study in this paper, targeting Lima City, Peru, attempts as a first step to create building inventory data necessary for earthquake damage estimation, extracting the spatial distribution of building age from Landsat imagery and an assessing building height from stereovision of ALOS/PRISM imagery.

2. Data

This study assesses built-up area distribution by age using satellite Landsat-TM (Thematic Mapper) image data, which is a rich archive of past image data. We searched first for good cloudless image data on Lima and found oldest to have been observed in 1987. In order to get data by age, we then searched for data on images shot 10 and 20 years after that and found data on cloudless images shot in 1998 and 2006. We decided to use images observed in those three different years for this study (**Table 1** and **Fig. 1**). As presented in **Fig. 1**, however, the 2006 image include clouds covering the western and southern coasts of Lima, so we used an auxiliary image observed in 2002 for the area covered with clouds. The image data includes seven bands in the visible to near-infrared or thermal-infrared range. The spatial resolution of images is 30 m, except for thermal infrared image, which is 120 m.

Building height was assessed using data from "Daichi," an ALOS satellite launched by Japan on January 24, 2006. ALOS has three sensors assigned to terrestrial observation as their main mission. One is the PRISM (Panchromatic Remote-sensing Instrument for Stereo Mapping) that, as the name suggests, creates and updates maps equivalent to 1/25,000 and ground resolution of 2.5 m. For the assessing building height discussed in Section 4, we used a PRISM image of Lima taken on October 15, 2008 (**Fig. 2**). The selected image was shot under the best observation conditions but there is still a cloud near the coastline and the situation of the surface under the cloud cannot be acquired.

As an existing basic geospatial map of Lima, we used a land-use map with a scale of 1/25,000 created in 2004. This map presents built-up areas in orange and vegetation areas in green (**Fig. 3**). Built-up area distribution presented in the land-use map is used for assessing built-up area distribution by age. For building height assessment, we also attempted to use contour lines and elevation values presented on the map in parts with bare land, roads, and so on.

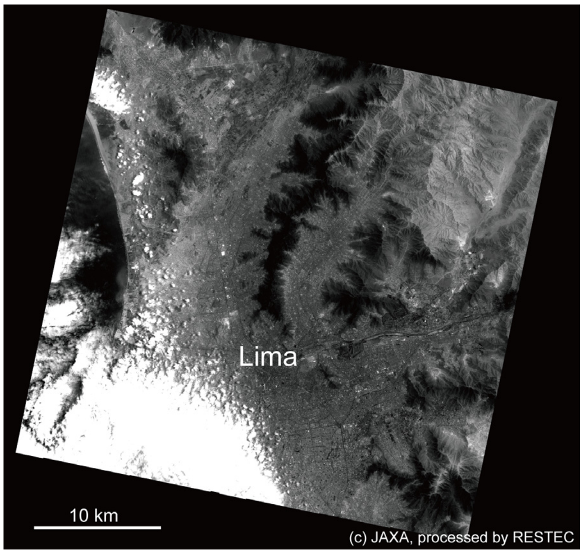

Fig. 2. ALOS/PRISM orthorectified image observed on 2008/10/15.

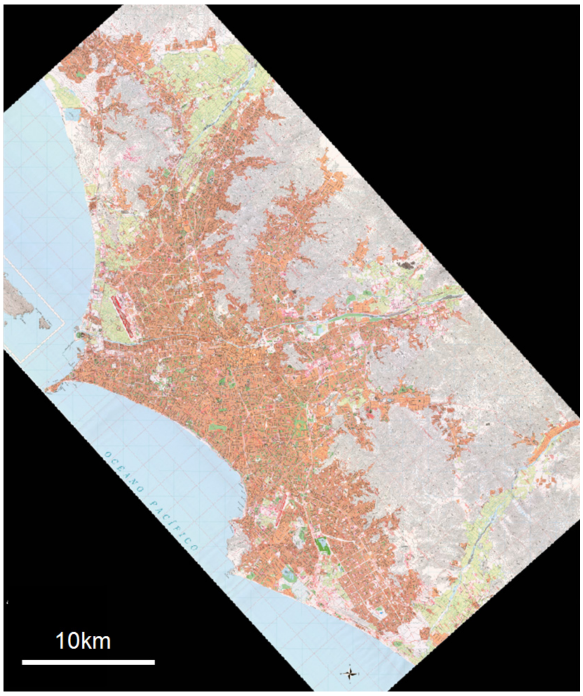

Fig. 3. Landuse map in Lima (2004).

3. Assessing Built-Up Area Distribution by Age Using LANDSAT Imagery

3.1. Outline of Assessment Method

In advance preparation, the land-use map (**Fig. 3**) was digitized at a resolution of 300 dpi (dot per inch) and converted to image data by giving position information. The surface resolution of imagery at that time was about 3 m. Next, pixel size was converted to that of an image with a resolution of 30 m, which is the same as that of the Landsat imagery, and each pixel was classified into one of three types, i.e., built-up area, vegetation, or others (bare ground, roads, etc.). The most typical land use in each pixel the size of 30 m was classified as the land use for each pixel. Built-up area distribution on the land-use map is presented in **Fig. 4**. Pixels in pink represent built-up area, in green vegetation, and in grey other areas.

Based on built-up area distribution of the land-use map presented in **Fig. 4**, assessing the distribution of built-up

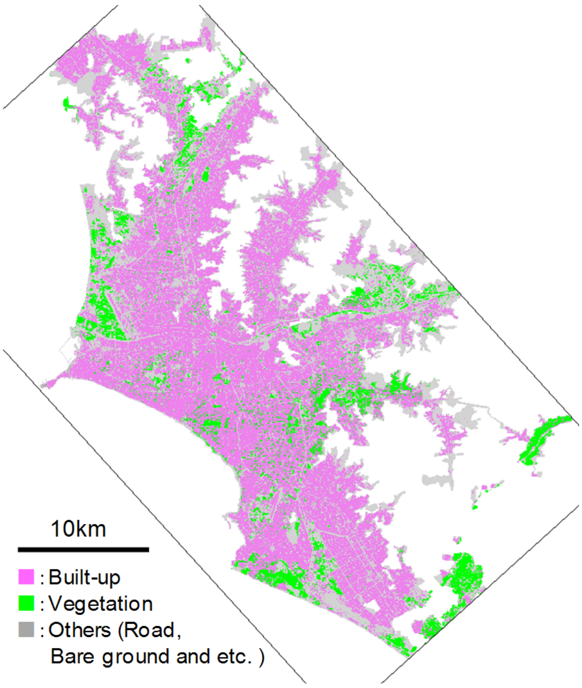

Fig. 4. Digitized landuse image.

area by age was conducted by assessing areas developed as new built-up areas through image analysis using Landsat imagery of different ages. Comparing past Landsat imagery from 1987 and 1998 with imagery from 2006 from Landsat imagery presented in **Fig. 1** as well as current built-up area distribution presented in **Fig. 4**, there are many areas that were vegetation areas or bare ground in past imagery but are currently built-up areas, indicating urban sprawl. For this reason, built-up areas are assessed by age in this study by assessing areas that were formerly vegetation or bare ground but are currently built-up areas.

Specifically, indices representing vegetation areas and bare ground are calculated from each Landsat image and differences in indices between images by age are calculated. Areas with a difference greater than a threshold are assessed as areas that have changed from vegetation areas or bare ground to built-up areas.

3.2. Calculation Indices

3.2.1. Assessing Vegetation Area

To assess vegetation areas or the bare ground in Landsat imagery, it is effective to calculate interband images to calculate indices that represent the possible existence of each type of feature. Indices typically representing the activation level of vegetation at pixels in an image include a normalized vegetation index, the NDVI (Normalized Difference Vegetation Index), which is expressed by the following equation.

$$NDVI = \frac{B4 - B3}{B4 + B3} \quad (-1 \leq NDVI \leq 1) \quad . \ . \ . \quad (1)$$

B3 and B4 indicate pixel values of band 3 and band 4 images, respectively. NDVI, widely used to assess vegetation based on remote sensing imagery, indicates that the greater the value, the more likely a vegetation area is included in an area that corresponds to a pixel. NDVI distribution calculated from the four images used is presented in **Fig. 5**. Area covered with clouds in the 2006 image is covered with a white frame. An enlarged view of NDVI distribution calculated from the 2006 Landsat image and the 2004 land-use map is presented in **Fig. 6**. As shown by the arrow in the figure, area corresponding to green space on the land-use map exhibits a high NDVI in the image, allowing us to confirm that the NDVI is effective in assessing vegetation area.

3.2.2. Assessing Bare Ground

A method used to assess bare ground from remote sensing imagery is a soil index, the NDSI (Normalized Difference Soil Index), which is expressed by the following equation [4]:

$$NDSI = \frac{B5 - B4}{B5 + B4} \quad (-1 \leq NDSI \leq 1) \quad . \ . \ . \quad (2)$$

B5 indicates the pixel value of band 5. Similar to the NDVI, the NDSI is an index calculated from interband calculation, indicating that a pixel with a greater value is more likely to represent bare ground. Bare ground and built-up area are relatively close in terms of spectral characteristics of imagery, however, and accordingly, both may be difficult to discriminate when using the NDSI [5] so, when assessing bare ground, we also discuss the following indices:

$$NBI = \frac{B3 \times B5}{B4} \quad . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \ . \quad (3)$$

$$NUI = \overline{B3} + \overline{B6} - \overline{B4} - \overline{B5} \quad . \ . \ . \ . \ . \ . \ . \ . \ . \quad (4)$$

where

$$\overline{Bn} = \frac{Bn - Ave_{Bn}}{SD_{Bn}} \times 50 + 100 \quad (n = 3,4,5,6) \qquad (5)$$

Here, NBI is an acronym for the New Built-up Index [5], which uses the ratio of pixel values between band 4 and band 5 to express the characteristic in which a pixel value of band 3 becomes larger for bare ground than for an built-up area. NUI is an acronym for the Normalized Urban Index [6], which has the characteristic in which pixel values of band 3 and band 6 become large in built-up areas and bare ground whereas pixel values of band 4 and band 5 are relatively small. The pixel value of each band is calculated using Eq. (5) from a value normalized using average value Ave_{Bn} and standard deviation SD_{Bn} of imagery. The NUI is characterized by values that become larger in built-up areas and bare ground than in green space.

Index distribution calculated from the 2006 Landsat image and the 2004 land-use map is present in **Fig. 7**. White area indicated by an arrow on the land-use map represents bare ground, such as mountains. The NDSI distribution value tends to be slightly lower in the area of bare ground and relatively high in urban areas. There are many pixels that indicate low values, however, and the

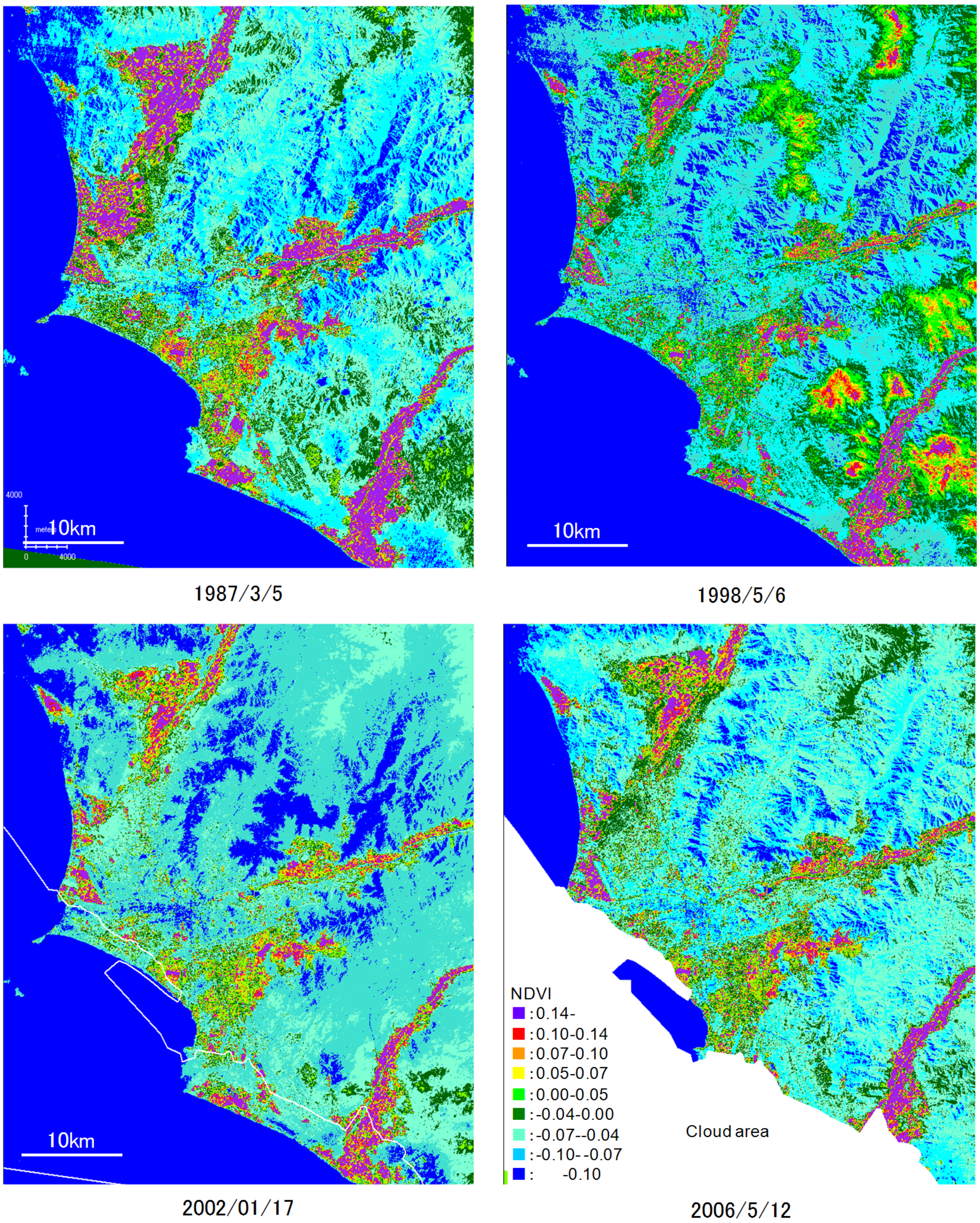

Fig. 5. NDVI images in 1987, 1998, 2002 and 2006.

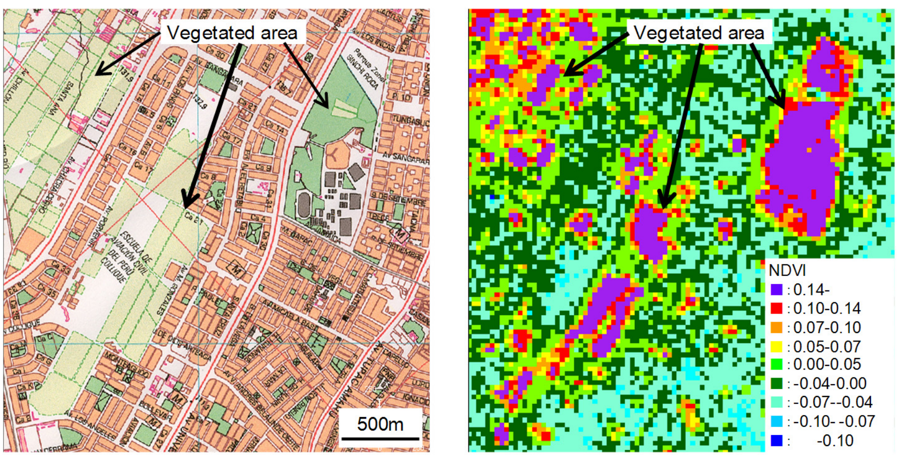

Fig. 6. Comparison of 2004 landuse map and 2006 NDVI image.

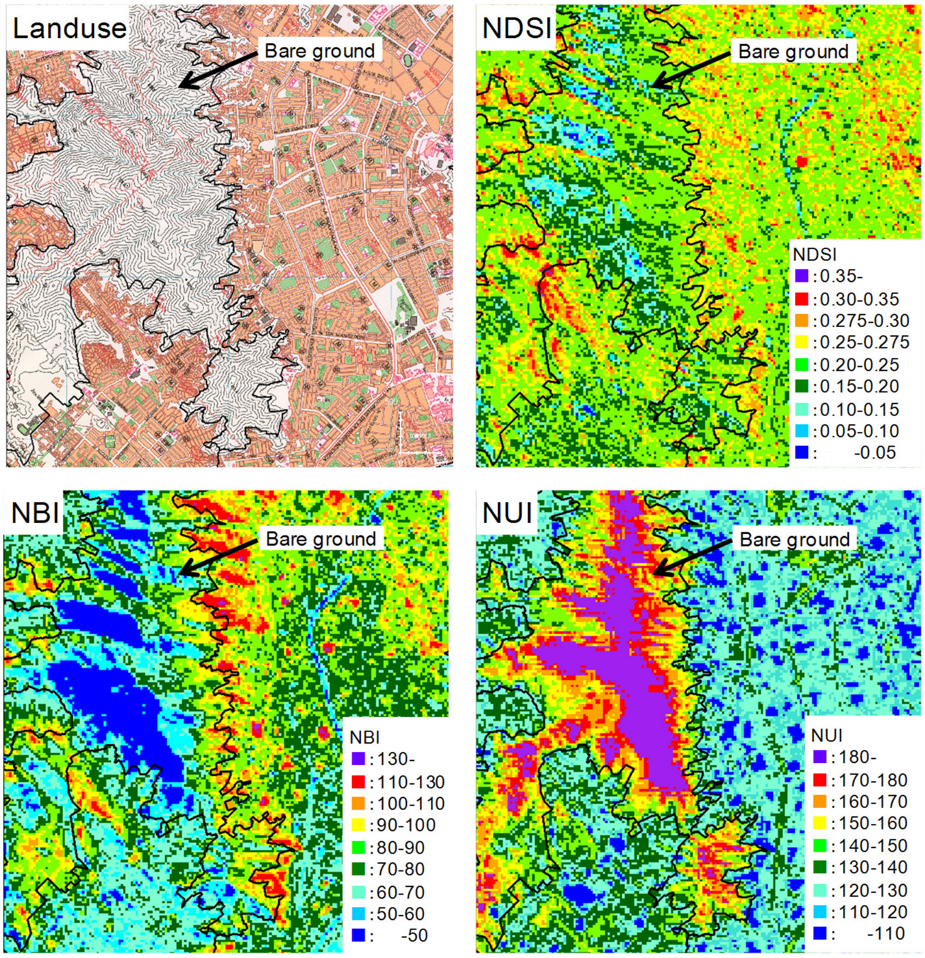

Fig. 7. Comparison of 2004 landuse map, NDSI, NBI and NUI images in 2006.

boundary between bare ground and built-up area is not clear. Similar to NDSI distribution, NBI distribution exhibits a low value for bare ground and a high value for an built-up area, indicating a boundary relatively clearer between bare ground and built-up area than for the NDSI. There are also areas with a small value, however, even in built-up areas. Unlike other indices, NUI distribution exhibits a large value for bare ground and a small value for an built-up area, exhibiting a relatively clear boundary. This is likely because, compared with other indices, the NUI uses more pieces of information of band 6 (thermal infrared region) and surface temperature of bare ground in the target area tends to be higher than that in urban area and thus the NUI is effective in extracting bare ground. As a result of discussion, this study uses the NUI for assessing bare ground. NUI distribution calculated from Landsat imagery is presented in **Fig. 8**.

3.3. Assessing Built-Up Area by Age

3.3.1. Discussion on Threshold Values of Indices

Areas that were vegetation in past imagery are high on the NDVI and become low after developing into built-up areas. Areas that were bare ground in past imagery are high on the NUI and become low after having changed to built-up areas. For this reason, urban sprawl is assessed from imagery by using NDVI and NUI difference values calculated from the images.

Verification of the effectiveness of this method requires past and current high-resolution image data other than Landsat imagery. For the target area, there are aerial photographs observed in 1984. For this study, as presented in **Fig. 9**, we acquired aerial photographs on new residential areas in the suburb of northern Lima. This aerial photograph is used as past validation data because it was observed at a time close to when 1987 Landsat imagery was observed. This monochrome image has a spatial resolution of about 20 cm. High-resolution satellite images have become available relatively easily. This study uses a satellite WorldView-2 (WV2) images observed in 2010 as current validation data in correspondence to the 2006 Landsat image. The acquired WV2 images have the area presented in **Fig. 9**. This color image has a spatial resolution of 50 cm.

Difference values between NDVI and NUI calculation from 2006 and 1987 Landsat images were calculated for each pixel (2006-1987). The negatively larger the difference value, the more likely the corresponding area has changed from vegetation or bare ground to built-up area. In order to analyze areas that are currently built-up area,

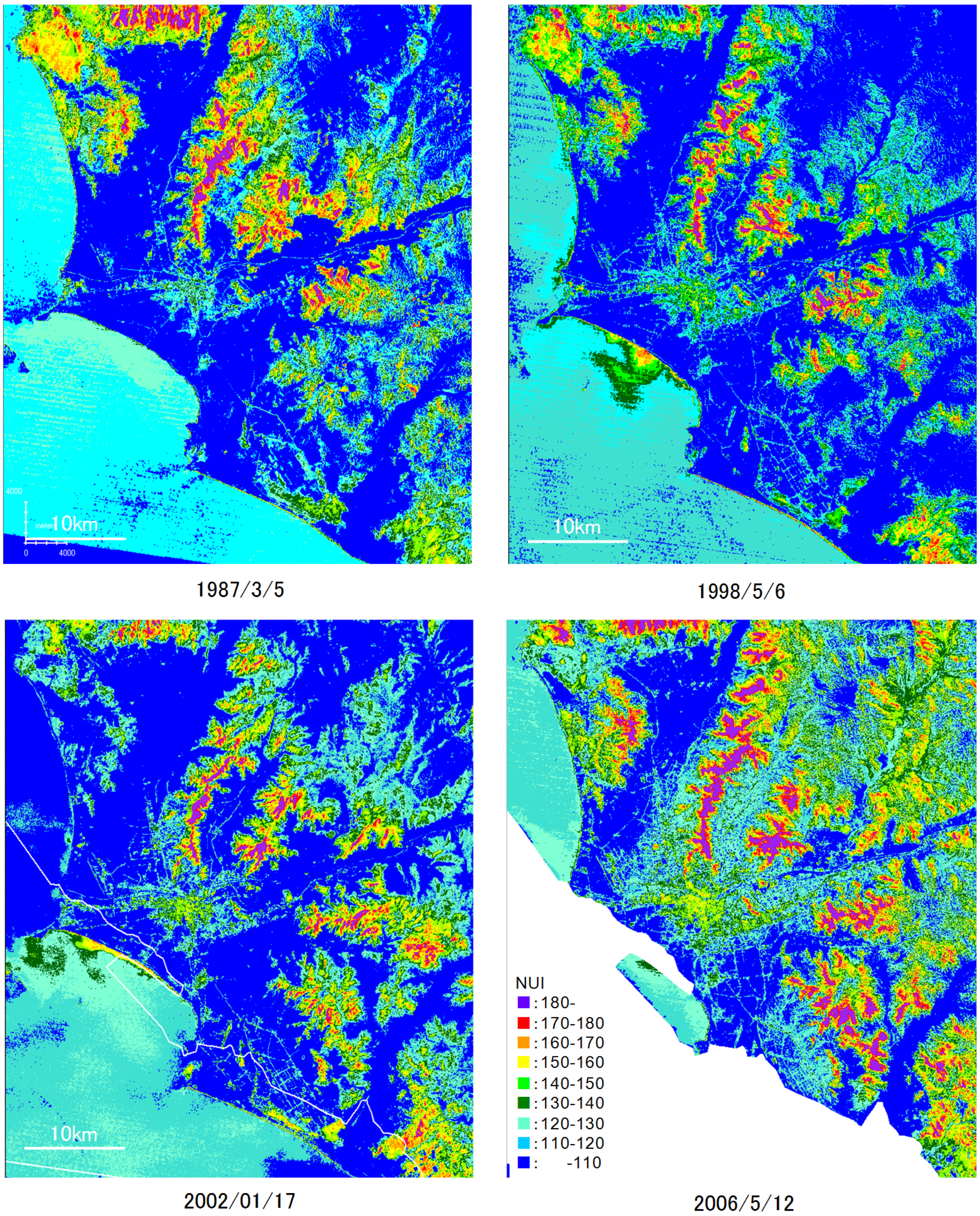

Fig. 8. NUI images in 1987, 1998, 2002 and 2006.

we focus on the current built-up area pixels presented in **Fig. 4**.

Figure 10 presents a comparison of the 1984 aerial photograph, the 2010 WV2 image and distributions of the NDVI and NUI difference. In the figure of the NDVI difference and the NUI difference, the closer to purple the color, the negatively greater the difference value. By visually judging the aerial photograph and the WV2 image, we extracted, as training areas, areas that have changed from vegetation to built-up area, from bare ground to built-up area, and that have been built-up at any period in time. In **Fig. 10**, the area indicated by the green frame represents areas that have changed from vegetation to built-up area, the area indicated by the orange frame represents areas that have changed from bare ground to built-up area, and the area indicated by the purple frame represents areas that have been built-up at any period in time. The distribution of the NDVI difference indicates that there are many red and purple pixels in areas that have changed from vegetation to built-up area and the difference value is negatively high. Similarly, the distribution of the NUI difference indicates that there are many red and purple pixels in areas that have changed from bare ground to built-up area and the difference value is negatively high.

According to **Fig. 10**, we extracted pixels in polygons of areas that have changed from vegetation to built-up area (green frame), from bare ground to built-up area (orange frame), and built-up at any period in time (purple frame), and created histograms of the NDVI and NUI difference values. Histograms are presented in **Figs. 11(a)**

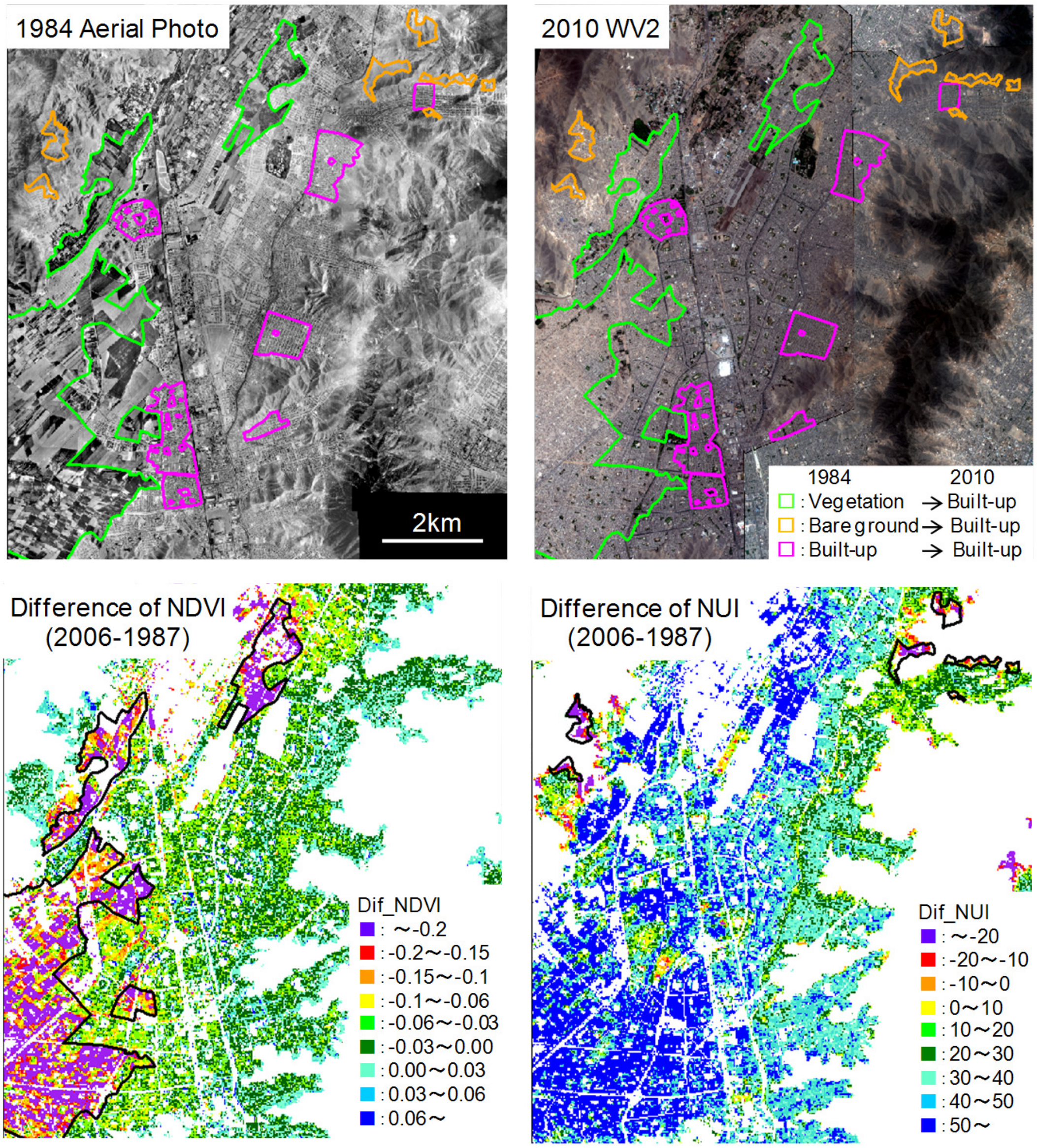

Fig. 10. Comparison of 1984 aerial photo, 2010 WV2 image, difference of NDVI and difference of NUI images.

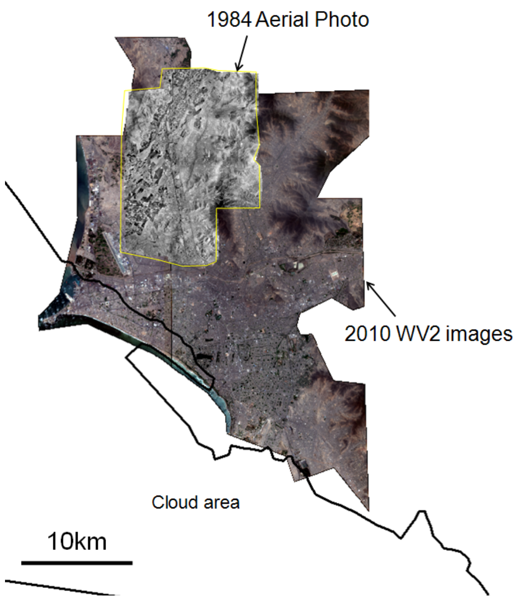

Fig. 9. Coverage of 1984 aerial photos and 2010 WV2 images.

and **(b)**. For the sake of histogram distribution clarity, values normalized by the maximum at each histogram are represented on the vertical axis. The red line represents the histogram of areas changed from vegetation to built-up, the blue line areas have changed from bare ground to built-up, and the black line represents areas that have been built-up at any period in time.

The distribution of the NDVI difference value of **Fig. 11(a)** indicates that areas that have changed from bare ground to built-up area and that have been built-up at any period in time each exhibit a difference value of approximately -0.1 or greater. Areas that have changed from vegetation to built-up exhibit a wide distribution from -0.6 to 0. This study sets a threshold value of -0.1 and extracts pixels with an NDVI difference value of -0.1 or less as areas that have changed from vegetation to built-up.

The distribution of the NUI difference value in

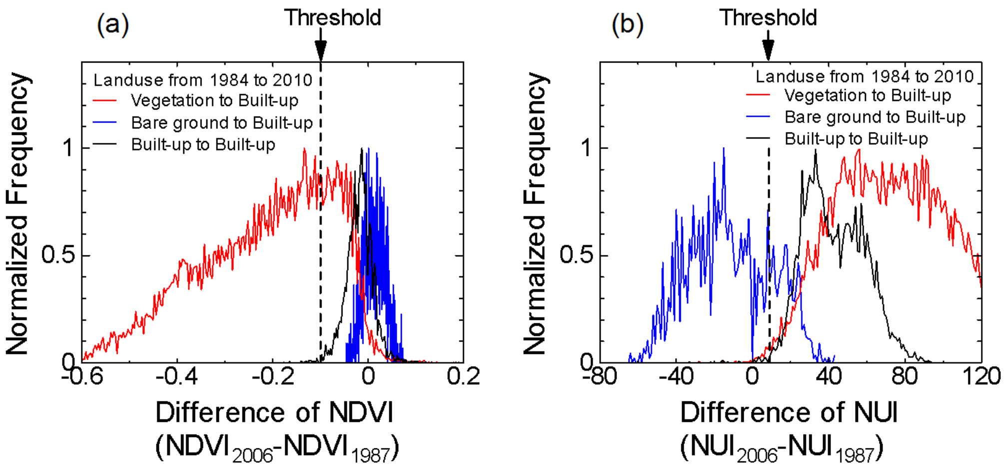

Fig. 11. Histograms of difference of NDVI and difference of NUI.

Table 2. Classification accuracy of threshold by difference of NDVI and NUI.

Landuse change from 1984 to 2010	Num. of Pixels	Threshold by Difference of NDVI			Threshold by Difference of NUI		
		Correctly classified (%)	Falsely classified (%)	Total (%)	Correctly classified (%)	Falsely classified (%)	Total (%)
Vegetation to Built-up	486,540	75.2	24.8	100.0	99.3	0.7	100.0
Bare ground to Built-up	30,105	100.0	0.0	100.0	80.5	19.5	100.0
Built-up to Built-up	149,859	99.5	0.5	100.0	99.3	0.7	100.0

Fig. 11(b) indicates that areas that have changed from vegetation to built-up and that have been the built-up at any period in time each exhibit a difference value of 10 or greater. Areas that have changed from bare ground to built-up area exhibit a wide distribution from −60 to 40. This study sets a threshold of 10 and extracts pixels with a NUI difference value of 10 or less as areas that have changed from bare ground to built-up.

Table 2 presents results of classification using each threshold value, i.e., the number of pixels included in each polygon and the correct classification rate and false classification rate of threshold value processing for NDVI and NUI difference values. The threshold value processing result from the NDVI difference value indicates that correct classification rates of areas changed from bare ground to built-up and areas that have been urban at any period in time are each almost 100% and the correct classification rate of areas that have changed from vegetation to built-up is relatively high at 75%. Similarly, the threshold value processing result from the NUI difference value indicates that correct classification rates of areas changed from vegetation to built-up and areas that have been built-up at any period in time are each almost 100% and the correct classification rate of areas changed from bare ground to built-up is relatively high at 80%.

3.3.2. Assessing Built-Up Area by Age and Discussions

We assessed the distribution of built-up area by age using the threshold values discussed in the previous section. The assessment flow is presented in **Fig. 12**. As presented in **Fig. 4**, built-up area pixels are extracted from the 2004 land-use map. Targeting these, NDVI and NUI difference calculated from 2006 and 1998 Landsat images are used to carry out threshold value processing using the threshold values discussed in the previous section. Pixels indicating a difference value lower than the threshold are classified as new built-up area developed after 1998. Threshold value processing is carried out similarly for pixels indicating a difference value higher than the threshold using NDVI and NUI differences calculated from the 1998 and 1987 Landsat images. Pixels indicating a difference value lower than the threshold are classified as relatively newer built-up area developed between 1987 and 1998, and pixels indicating a difference value higher than the threshold are classified as older built-up area that existed before 1987. In areas covered with cloud in the 2006 Landsat image, pixels are similarly classified using the 2002 Landsat image in place of the 2006 Landsat image. These results are combined to assess the distribution of urban area by age.

Assessment results are presented in **Fig. 13**. Pixels in pink represent older built-up area that existed since before 1987, pixels in red represent built-up area developed between 1987 and 1998, and pixels in yellow represent newer built-up area developed after 1998. Although most of Lima City had been developed before 1987, there are relatively new urban areas in northern, southern, and eastern Lima City, which indicates that urban sprawl has been gradually taking place.

The number of pixels of the built-up area of each age counted from assessment results in **Fig. 13** is presented in **Table 3**. The square of built-up area for each age is also calculated from the number of pixels and pixel size (30 m × 30 m). In the target area, there are built-up areas in a

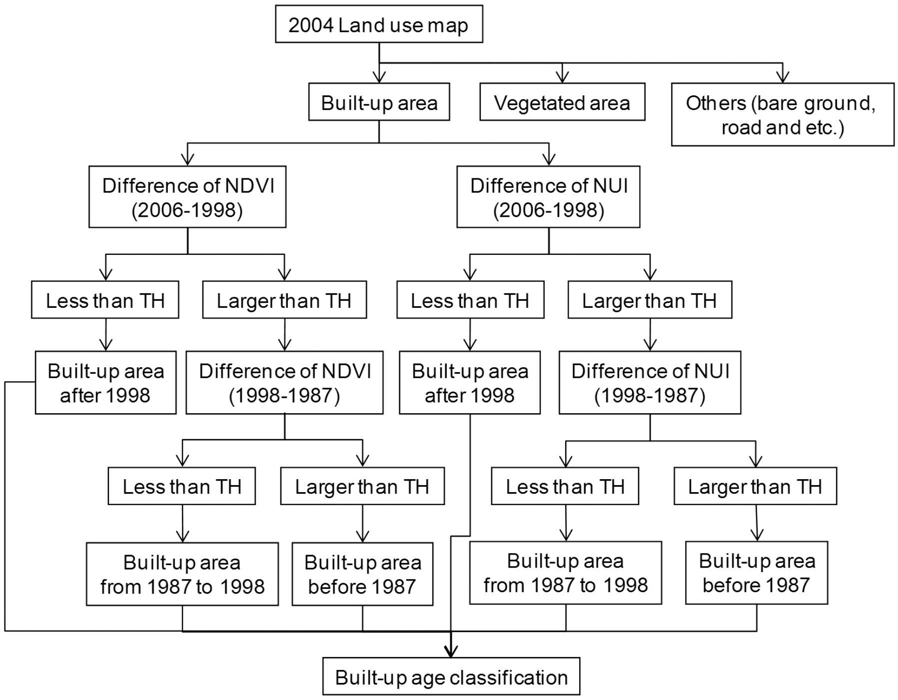

Fig. 12. Flowchart of built-up age classification using landuse map, NDVI and NUI images. (TH: Threshold value).

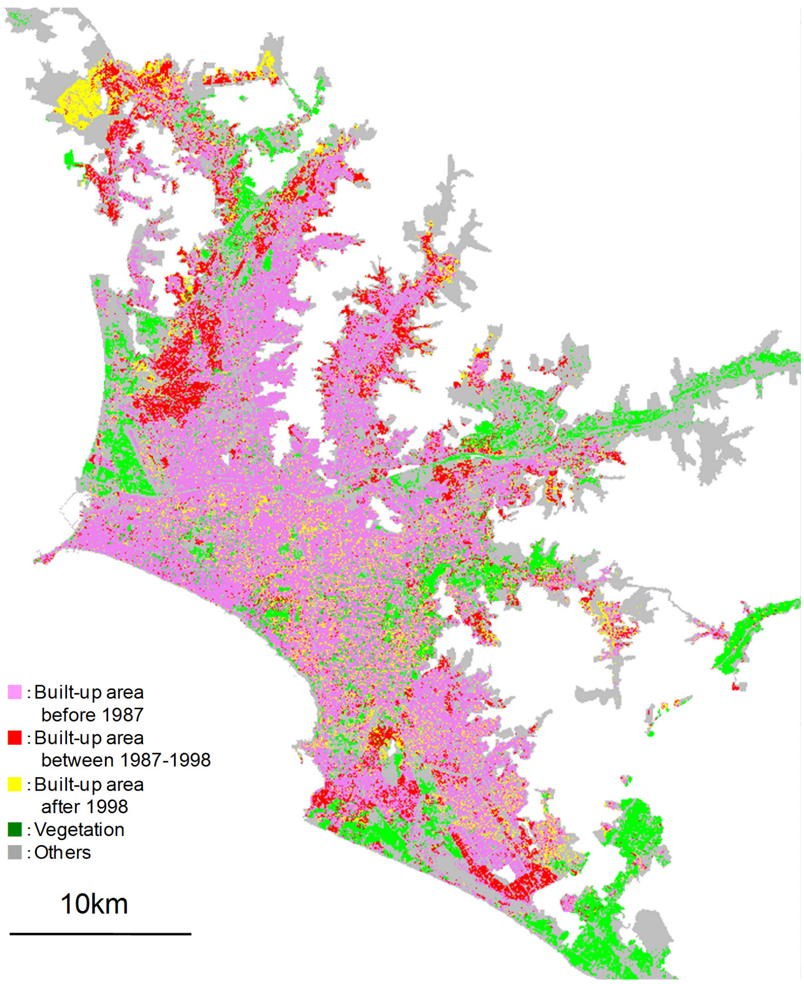

Fig. 13. Result of built-up age classification.

Table 3. Number of pixels and area of each built-up area.

Age	Num. of Pixels	Area (km^2)	Percent (%)
Built-up before 1987	1,192,567	1,073	68.1
Built-up between 1987-1998	314,830	283	18.0
Built-up after 1998	243,192	219	13.9
Total	1,750,589	1,576	100.0

total of about 1,600 km^2, about 70% of which have been urban since before 1987, about 20% of which developed as urban between 1987 and 1998, and slightly over 10% that are built-up areas developed after 1998.

Since analysis results are classification results in 30 m mesh, it is difficult to comprehend the age of individual buildings. It is possible, however, to evaluate when each built-up area was developed. In future, if building inventory data that indicates the number of buildings are created, age information necessary for earthquake damage estimation of buildings will be able to be added using these analysis results.

4. ALOS/PRISM Image-Based Estimating Building Height

4.1. Characteristics of PRISM and DSM

Information measured by stereovision using two aerial photographs and satellite imagery observed from two directions is called the DSM (Digital Surface Model). This includes not just the height of ground but also the height of features such as buildings and trees. To get a wide area of DSM, the use of a satellite with a wide range of observation is efficient, and small-size features such as buildings are usually extracted using images from high-resolution satellites such as IKONOS and QuickBird. These satellites are not, however, constantly engaged in stereo observation, so DSM is not always acquired with the expected area at the expected period of time. PRISM on the ALOS satellite has three sensors – views forward, nadir, and backward – with angles different from one another. In the orbital direction, a triplet stereo image with an acquisition time difference for each sensor of 45 seconds is obtained constantly. The DSM is generated through matching processing with the triplet stereo image.

This study attempted to extract building height from the DSM of a PRISM image that has such advantages. The DSM calculation procedure is as follows (see reference [7] for details): A stereo-matching point is searched by simultaneously moving forward and backward viewing correlation windows in the disparity direction with respect to the correlation window centered on the calculation grid (equivalent of pixel size of DSM) of the nadir image. Correlation windows are moved on an approximate straight line that meets geometric conditions of observation orbit in regard to images forward, nadir, and backward. The point at which sum of correlation coefficients and obtained from a stereo pair of forward and nadir viewing and a stereo pair of backward and nadir viewing, respectively, becomes maximum is designated as the stereo-matching point. The DSM value is uniquely determined from the thus searched-for stereo-matching point. The size of correlation windows in matching, which is one of the parameters difficult to set, is set to a variable between 9×9 pixels and 29×29 pixels and an automatically optimized size is used. In general, a smaller correlation window can track a smaller target object and can accurately extract a spatially high-frequency component, but it is difficult to apply to an area with poor texture and susceptible to noise. A larger correlation window has the advantage of better noise tolerance but it is difficult to extract a feature with high spatial frequency. An optimal

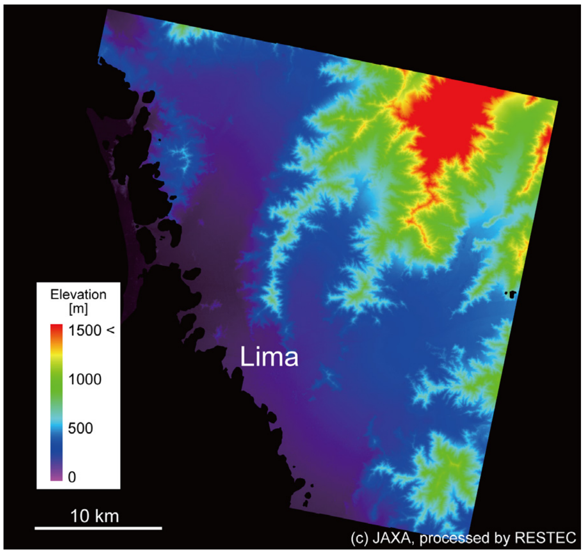

Fig. 14. Digital surface model (DSM) derived from ALOS/PRISM image.

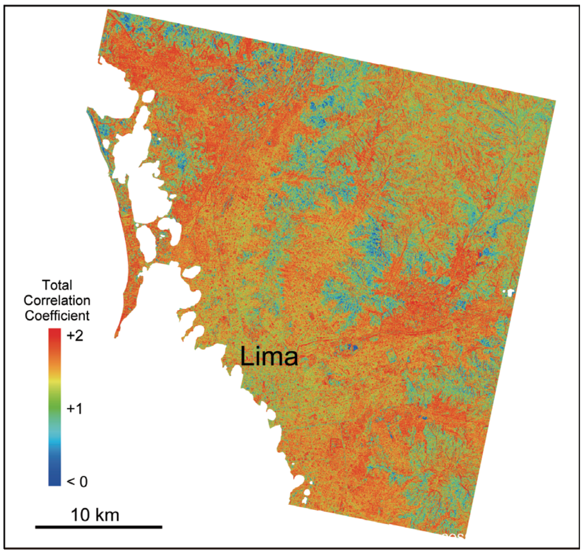

Fig. 15. Sum of correlation coefficient image of triplet stereo matching.

size that balances such a tradeoff is used to determine the stereo-matching point [7].

Figure 14 presents a DSM image of the study target area. The DSM was created with grid intervals of 5 m. Water bodies and cloud areas are masked because the DSM cannot be calculated. **Fig. 15** presents an image of the sum of correlation coefficients. According to its value distribution, the value of urbanized areas is generally higher than that of hills. Among built-up areas, the value is relatively small in central Lima, indicating the difficulty in searching for a corresponding point at the time of matching processing. As mentioned earlier, the DSM acquired as shown in **Fig. 14** includes not only ground level height but also feature height. We will now discuss how to extract the height of buildings from the DSM.

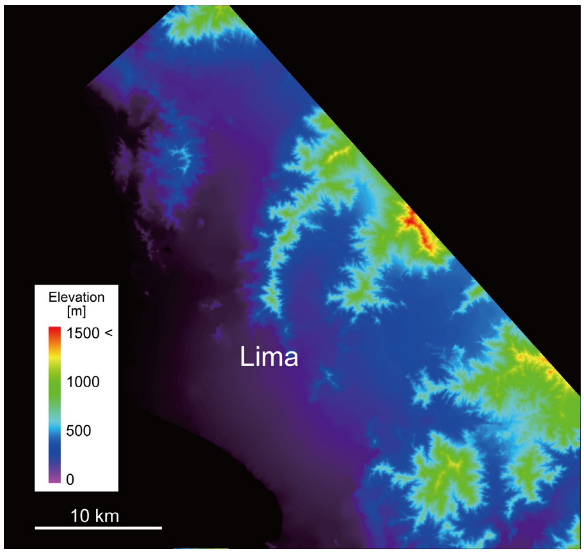

Fig. 16. Digital elevation model (DEM) generated by landuse map.

4.2. Feature Height Assessment

4.2.1. Difference Between DSM and DEM

The most common way to extract the height of features is to deduct ground level height from the DSM. Data on the digitized ground level height are called a DEM (Digital Elevation Model). Global dataset is known as GTOPO30 [8]. GTOPO30 assumes, however, a grid size of elevation points of 30 seconds (about 1 km), which gives low ground resolution. In Japan, the Geospatial Information Authority of Japan publishes Basic Geospatial Information (Digital Elevation Model) [9] based on large-scale topographic maps that approximately covers Japan completely in a grid of 10 m. There are data available in a grid of 5 m for some areas of Japan. Such a high-resolution DEM does not exist in developing countries such as Peru, however.

This study attempted to create a DEM for Lima based on information on leveling, contour lines included in existing topographic maps, and so on. The land-use map presented in **Fig. 3** includes elevation values in some areas of bare ground and roads, i.e., area without buildings. Contour lines are drawn over hills. We carried out spatial interpolation using the IDW (Inverse Distance Weighted) method so that discrete elevation data are on a grid of 5 m. **Fig. 16** presents a DEM generated from a land-use map that exhibits good overall correspondence compared with the PRISM DSM in **Fig. 14**. Both include offset resulting from differences in geodetic data and thus its influence needs to be eliminated. We then extracted approximately 10,000 pixels at random for the entire image from areas that were classified as bare ground, roads, etc., from the land-use map. We then calculated offsets between the DSM and DEM of pixels. An image of difference in the DSM and the DEM based on offset is presented in **Fig. 17**. The figure indicates that the difference value is 0 or slightly larger in a wide area of low land in particular. Since in general, the DSM including the height of features has a larger value, this tendency is valid. There are areas with difference values, however, that are extremely small, i.e., the DEM value is larger, and those that are extremely large, i.e., the DEM value is smaller locally.

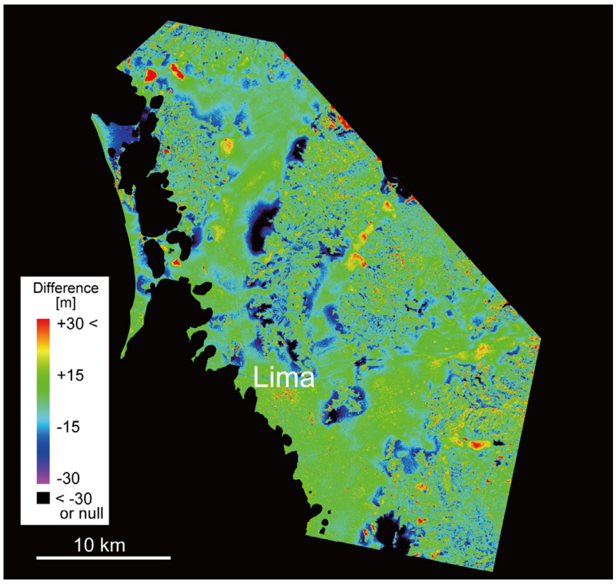

Fig. 17. Difference image between PRISM DSM and landuse DEM.

Figure 18 presents a difference image of central Lima together with an orthorectified image. The footprint of mid-to-high-rise buildings and the shape of urban areas have been confirmed and difference values reflect building height. **Fig. 19** presents an example of areas with small difference values. According to the figure, the value is extremely small in low land near a hill, i.e., the value of the DEM is extremely larger than that of the DSM. **Fig. 20** presents the same area of land-use map as that used to generate the DEM. Unlike with hills, which have a contour line, elevation values are written at a small number of points in low land, i.e., elevation values do not exist with sufficient density. DEM values are large in low land because of the influence of elevation values of hills in interpolation due to the small number of elevation values

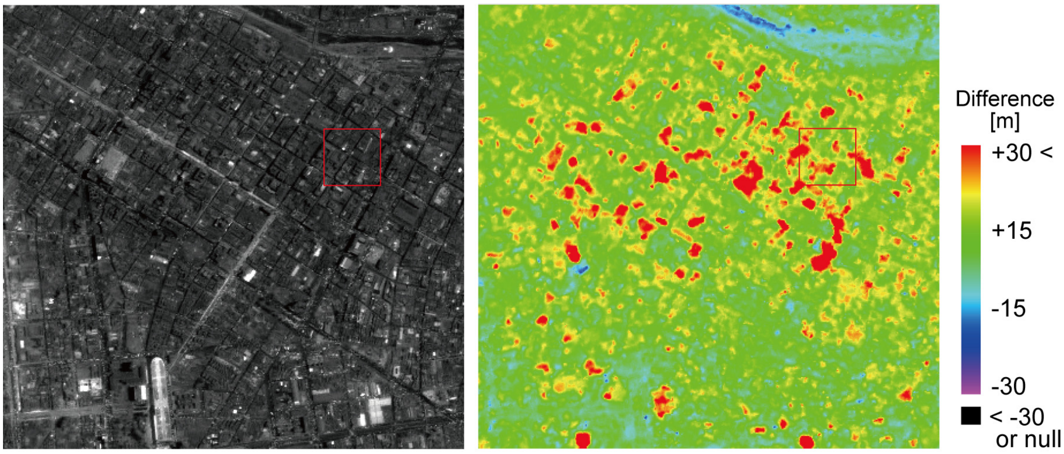

Fig. 18. Zoom-in images (Area A) of PRISM image and difference between PRISM DSM and landuse DEM.

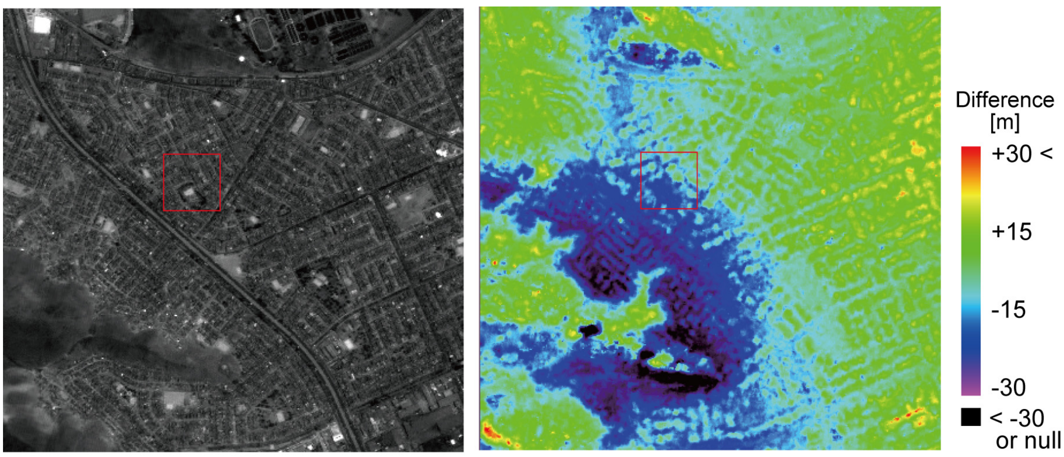

Fig. 19. Zoom-in images (Area B) of PRISM image and difference between PRISM DSM and landuse DEM.

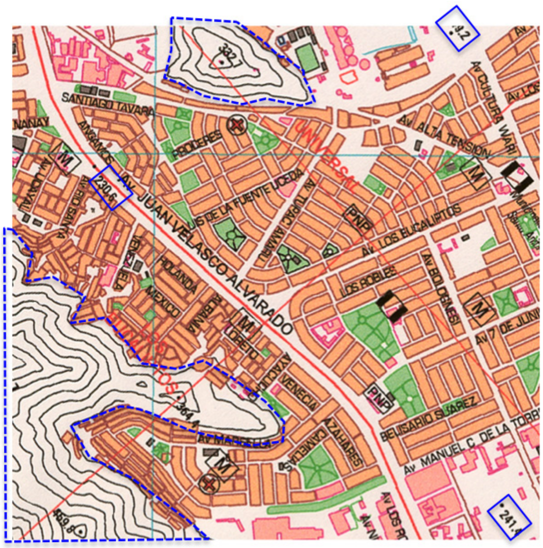

Fig. 20. Landuse map of Area B. Blue rectangle indicates the point of elevation measurement, blue dot-polygon indicates contour of elevation.

that can be referenced.

As seen above, a highly accurate DEM cannot be acquired from existing maps, and it is difficult to estimate building height using the difference from the DSM. We will now attempt to assess building height from the DSM alone.

4.2.2. DSM Filtering Process

Once the DSM value of the ground level around a building is given, the height of the building can be assessed by calculating the difference from the DSM of the building. In data on building footprint, the outer circumference of a building footprint represents ground level, where the DSM value is thought to be smaller than that of the surroundings. There are no such data on building footprint, however, in the target area. In this study, when assessing a point that is thought to be at ground level from the surrounding of a building, we set a calculation window and assumed that the minimum value of the DSM in the window is likely to be ground level. We assessed the height of a feature at the target point by deducting the value from the DSM at the center pixel (target point) in the window. Actual processing includes the step of moving the win-

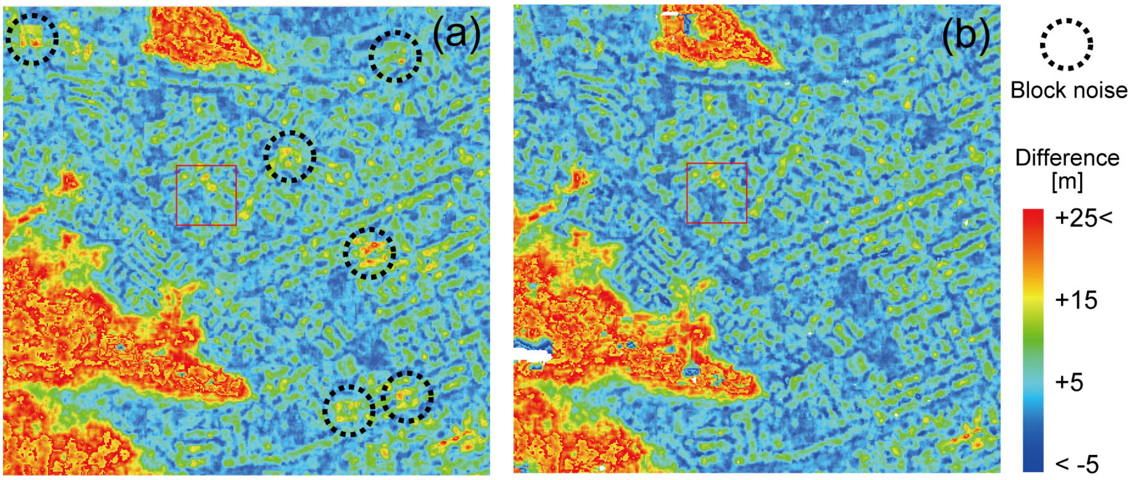

Fig. 22. Zoom-in images (Area B) of height difference generated by PRISM DSM filtering. (a) simple method, (b) modified method.

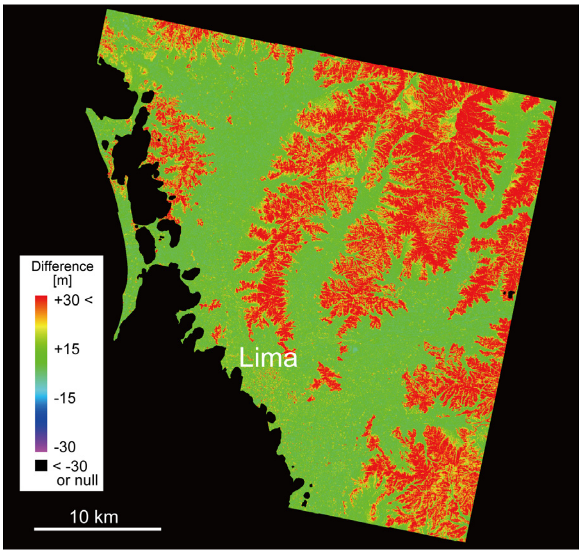

Fig. 21. Height difference image generated by PRISM DSM filtering.

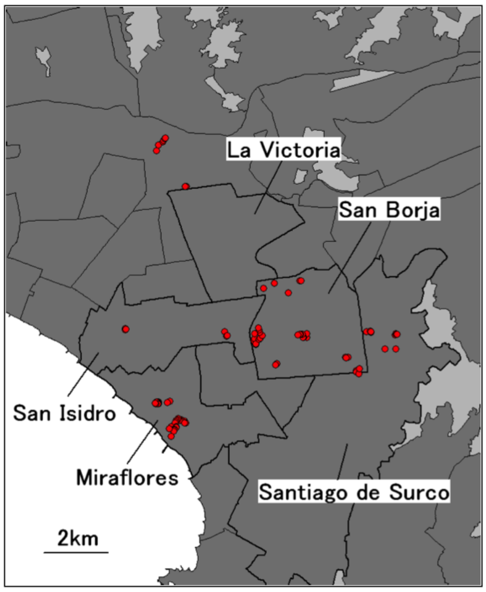

Fig. 23. Distribution of buildings carried out height measurement in field survey.

dow in units of 1 line by 1 pixel for the entire image and repeating calculation, thereby giving a distribution of feature height. The size of the calculation window was set at 17×17 pixels (about 85 m square) in view of building size.

Figure 21 presents the distribution of the estimated feature height. **Fig. 22(a)** presents an image in the same area as that of **Fig. 19**. For urban areas, height information similar to building shape seems to have been successfully extracted. The boundary between hills and low land becomes clear, but no area has been found in which the value is extremely small for low land near a hill, as shown in **Fig. 19**. Some areas, however, have rectangular block noise regardless of building distribution. Since its size is the same as the size of the calculation window, an abnormally small DSM value may have been extracted when the minimum value in the window was calculated. Such an abnormal value is caused by a matching error at the time of DSM generation and thus needs to be removed from the calculation target.

To do so, statistical values, i.e., average *av* and standard deviation *sd*, of the DSM in the calculation window are calculated in advance and points with a DSM value greater or smaller than $av \pm 5\times\ sd$ are removed from the calculation target. In addition, points for which the value of the sum of correlation coefficients at the time of DSM generation is less than 1.5 are removed from the calculation target on the assumption that these points are low in DSM value reliability. **Fig. 22(b)** presents a distribution of feature height acquired using the improved method. Although some points cannot be calculated for some hills and low land, block noise is reduced in a wide area of low land and height distribution of the feature similar to building shape.

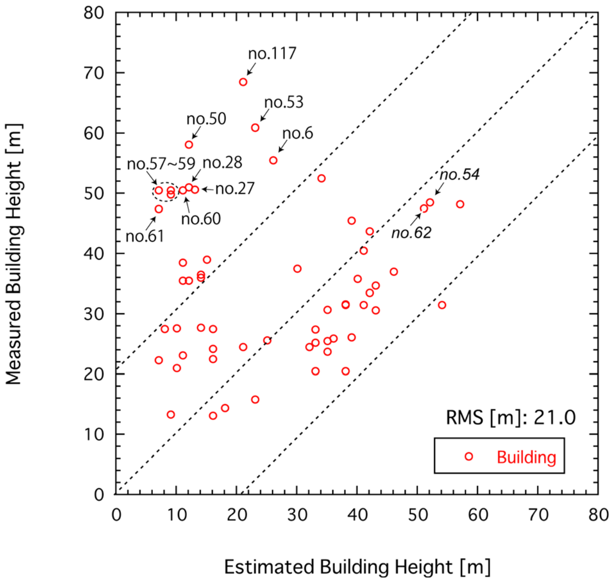

Fig. 24. Relationship between estimated building height from PRISM DSM image and measured height by field survey.

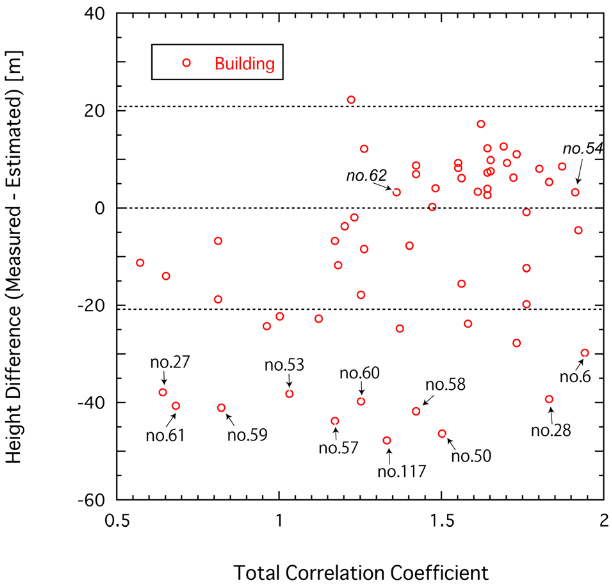

Fig. 25. Relationship between total correlation coefficient and height difference.

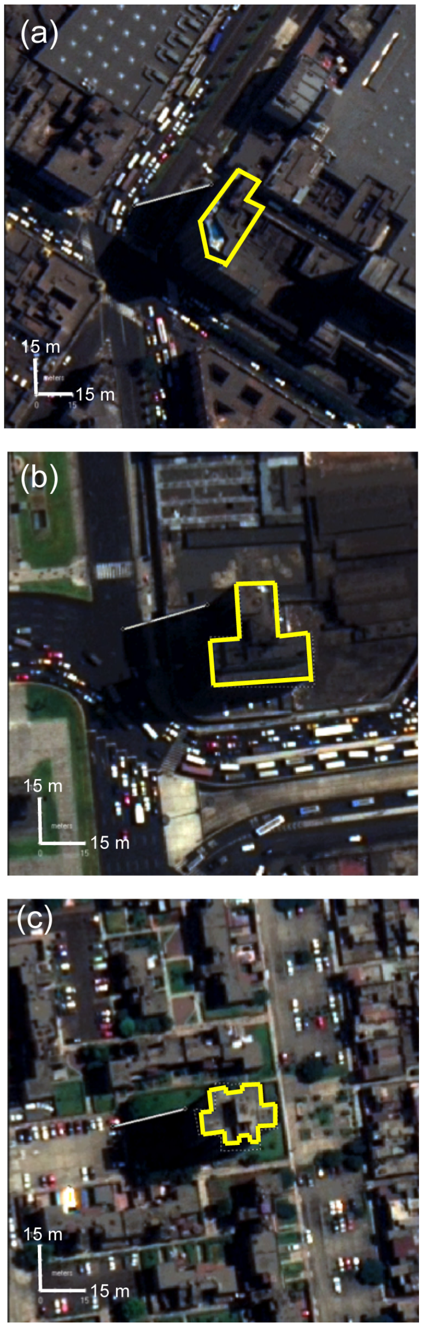

Fig. 26. Buildings carried out height measurement in WV2 image. (a) building No.117, (b) building No.50, (c) building No.59.

4.3. Building Height Estimation and Verification

We conducted a field survey in September 2011 in order to verify the assessment accuracy of feature height calculated from the DSM. We targeted mid-to-high-rise residential buildings and office buildings and measured the height of the ceiling of the top floor using a laser range finder. **Fig. 23** presents a distribution of the 119 buildings surveyed. The height range of surveyed buildings is 10 to 90 m. Among them, for 59 buildings existing in cloud-free areas, the position is confirmed from the PRISM orthorectified image and WV2 image and a building footprint is created for each, then is overlaid on feature height distribution acquired in **Fig. 21**, and the maximum value of the feature height within the building footprint is determined to be the estimated value of building height.

Figure 24 presents a result of comparison between the building height estimated from PRISM DSM and the building height actually measured on site. According to calculation, the RMS error was 21 m. The figure indicates that there are 28 buildings for which a difference between the assessment value and actually measured value is equal to or less than 10 m, and about half of the buildings have been successfully assessed with an error equal to or less than 10 m. Of these, however, 17 buildings have a difference in error equal to or greater than 20 m. There are

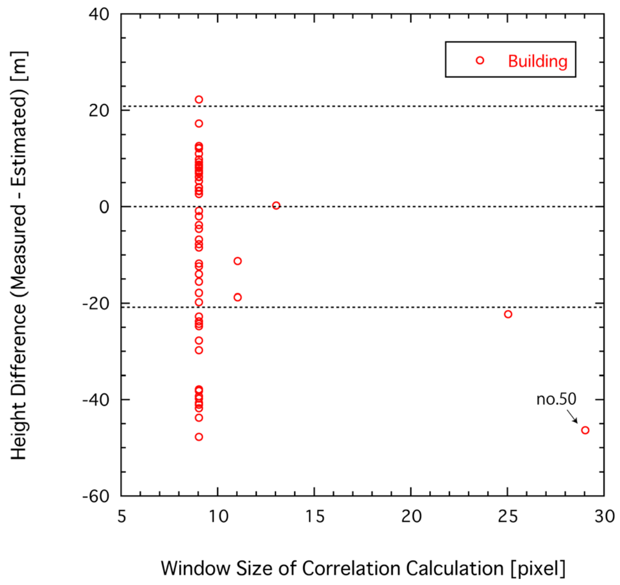

Fig. 27. Relationship between window size of correlation calculation and height difference.

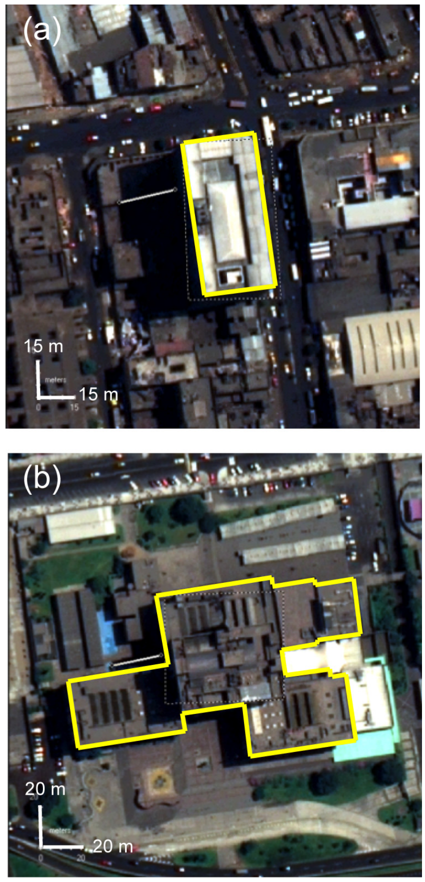

Fig. 28. Buildings carried out height measurement in WV2 image. (a) building No.54, (b) building No.62.

many buildings for which the estimated value in particular is underestimated more than the actually measured value. To discuss DSM value reliability at a point where building height was assessed, **Fig. 25** presents a comparison between the sum of correlation coefficients and errors. It is not that error is great when the value of the sum of correlation coefficients is small but that there is almost no correlation between them.

We now discuss the cause of the large error in height estimation seen in some buildings. **Fig. 26(a)** presents a WV2 image of building No.117, which has the largest error. The size of the building floor is represented by a yellow polygon. The building, located at the corner of an intersection point, is a 22-storey high-rise building 68.6 m high, but the building floor area is relatively small. **Fig. 26(b)** presents building No.50, an 18-storey building 58.2 m high. There are no high buildings surrounding building No.50. The color of the building is close to that of surrounding buildings and empty spaces. Such buildings with a small building floor area or a hue that cannot be distinguished from the surroundings are thought to be difficult to detect. **Fig. 27** presents the size of a correlation window in matching processing. For most points, a corresponding point is determined in a correlation window of 9×9 pixels but the correlation window for building No.50 is 29×29 pixels. Accordingly, building No.50 is missed when stereo-matching points are searched for and the height of surrounding ground is calculated. Building nos. 57 to 60 are high-rises of 50 m high with the same shape that have an error of 40 m. **Fig. 26(c)** presents building No.59 as an example. The accurate height of a building with a complicated shape cannot be calculated due to difficulty in finding stereo-matching points.

We now discuss buildings for which the height is estimated accurately. **Fig. 28(a)** presents building No.54, which has an estimation error of 4 m. The building has a bright-colored roof, which is different from the color of the surroundings, and has a slightly larger building floor area. The value of the sum of correlation coefficients is 1.9, which indicates that the stereo-matching point has been calculated accurately. **Fig. 28(b)** presents building No.62, which also has a large floor area. Texture given by irregularity on the roof is thought to work as a cue for a corresponding point search.

As seen above, we have successfully estimated the building height of about half of the surveyed buildings with errors of 10 m or less. We have not, however, accurately extracted small buildings, buildings with complicated shapes, or buildings with hue and texture similar to those of surroundings in stereo matching. As a result, the height of buildings assessed from the DSM tends to be smaller than that of actual buildings. This suggests that PRISM images have insufficient resolution for assessing the height of each building in a highly accurate manner. Areas with a concentration of large-size, mid-to-high-rise buildings can, however, be assessed to some extent. In future, we will attempt to evaluate the height of buildings in city blocks and create building inventory data that contribute to building damage estimation.

5. Conclusions

This study has intended to create building inventory data that contribute to earthquake damage estimation. Targeting Lima City, Peru, this study has compared aerial photographs and field survey results in an attempt to extract the spatial distribution of building age from time-series Landsat satellite images and assessing building height from ALOS/PRISM stereovision images. The main results are as follows:

Using Landsat images shot in 1987, 1998, and 2006, we calculated indices that represent vegetation area and bare ground and, through threshold value processing of differences in indices among age, assessed areas that have changed from vegetation to built-up and from bare ground to built-up. By comparing an aerial photograph observed in 1984 and a high-resolution satellite image observed in 2010, we have verified the assessment accuracy of changed points, and have successfully extracted them with a correct classification rate of about 80%.

By calculating built-up area distribution by age for all of Lima City, we have found that about 70% of built-up area had been developed before 1987 and that newly developed built-up areas sprawled out to the suburb of northern, southern, and eastern Lima.

Distribution of feature height acquired by difference between the DEM (Digital Elevation Model) generated from a land-use map and the DSM (Digital Surface Model) of PRISM imagery has the problem that the DEM value becomes extremely large in low land near a hill. This is caused by the low density of elevation measurement data in low land on the land-use map.

Comparing the distribution of feature height acquired through filter processing of PRISM DSM images with actual building height acquired from field surveys, both exhibit relatively good correspondence for large-size, mid-to-high-rise buildings. It may be difficult, however, to detect buildings with small floor area, with hue similar to that of surroundings, and with complicated shape.

Acknowledgements

Landsat images are owned by the United States and provided by the United States Geological Survey. WorldView-2 images are owned by DigitalGlobe, Inc. ALOS/PRISM images are owned by the Japan Aerospace Exploration Agency. The images and DSM data are provided by the Remote Sensing Technology Center of Japan. We have been taught the DSM generation method by Mr. Jun'ichi Takaku and Mr. Akira Mukaida of the Remote Sensing Technology Center of Japan. We have received cooperation in the data analysis from Mr. Takeshi Takase and Mr. Yusuke Hirano, graduates of the Tokyo Institute of Technology. This study was supported in part by the JST-JICA Science and Technology Research Partnership for Sustainable Development (SATREPS), Enhancement of Earthquake and Tsunami Disaster Mitigation Technology in Peru (Principal: Prof. Fumio Yamazaki, Chiba University). We express our appreciation to all of those who have been involved in this work.

References:

[1] H. Miura and S. Midorikawa, "Updating GIS Building Inventory Data Using High-resolution Satellite Images for Earthquake Damage Assessment: Application to Metro Manila, Philippines," Earthquake Spectra, Vol.22, No.1, pp. 151-168, 2006.

[2] F. Yamazaki, M. Matsuoka, H. Mitomi, and Y. Yusuf, "Macro- and Micro-scale Urban Classification of Metro Manila for Seismic Risk Assessment Using Satellite Images," 6th Multi-Lateral Workshop on Development of Earthquake and Tsunami Disaster Mitigation Technologies and Their Integration for the Asia-Pacific Region, EDM Technical Report, No.18, CD-ROM, Earthquake Disaster Mitigation Research Center, p. 8, 2004.

[3] I. F. Shaker, A. Abd-Elrahman, A. K. Abdel-Gawad, and M. A. Sherief, "Building Extraction from High Resolution Space Images in High Density Residential Areas in the Great Cairo Region," Remote Sensing, Vol.3, No.4, pp. 781-791, 2011.

[4] W. Takeuchi and Y. Yasuoka, "Development of Normalized Vegetation, Soil and Water Indices Derived from Satellite Remote Sensing Data," Journal of the Japan Society Photogrammetry and Remote Sensing, Vol.43, No.6, pp. 7-19, 2005 (in Japanese).

[5] J. Chen, L. Manchun, Y. Liu, and C. Shen, "Extract Residential Areas Automatically by New Built-Up Index," Proceedings of 18th International Conference on Geoinformatics, pp. 1-5, 2010.

[6] I. Saito, M. Piao, and O. Ishihara, "Extraction of Land Covering Changes by Landsat TM Data," Journal of Architecture, Planning and Environmental Engineering, Vol.561, pp. 79-85, 2002 (in Japanese).

[7] J. Takaku and T. Tadano, "PRISM On-Orbit Geometric Calibration and DSM Performance," IEEE Transaction on Geoscience and Remote Sensing, Vol.47, No.12, pp. 4060-4073, 2009.

[8] USGS (US Geological Survey): GTOPO30 – Global 30 Arc Second Elevation Data Set, http://www1.gsi.go.jp/geowww/globalmap-gsi/gtopo30/gtopo30.html [access available on Nov. 1, 2012]

[9] GSI (Geospatial Information Authority of Japan): Basic Geospatial Information (Digital Elevation Model 10 m and 5 m), http://www.gsi.go.jp/kiban/index.html (in Japanese) [access available on Nov. 1, 2012]

Name:
Masashi Matsuoka

Affiliation:
Associate Professor, Department of Built Environment, Tokyo Institute of Technology

Address:
4259-G3-2, Nagatsuta, Midori-ku, Yokohama 226-8502, Japan

Brief Career:

1992 Research Associate, Tokyo Institute of Technology
1996 Engineer, Remote Sensing Technology Center of Japan
1998 Deputy Team Leader, RIKEN
2004 Team Leader, National Research Institute for Earth Science and Disaster Prevention
2007 Senior Research Scientist, National Institute of Advanced Industrial Science Technology
2010 Division Chief, National Institute of Advanced Industrial Science Technology
2012- Associate Professor, Tokyo Institute of Technology

Selected Publications:

- Matsuoka and Yamazaki, "Use of Satellite SAR Intensity Imagery for Detecting Building Areas Damaged due to Earthquakes," Earthquake Spectra, EERI, Vol.20, No.3, pp. 975-994, 2004.
- Matsuoka and Nojima, "Building Damage Estimation by Integration of Seismic Intensity Information and Satellite L-band SAR Imagery," Remote Sensing, MDPI, Vol.2, No.9, pp. 2111-2126, 2010.
- Matsuoka and Yamazaki, "Comparative Analysis for Detecting Areas with Building Damage from Several Destructive Earthquakes Using Satellite Synthetic Aperture Radar Images," Journal of Applied Remote Sensing, SPIE, Vol.4, 041867, 2010.

Academic Societies & Scientific Organizations:

- Earthquake Engineering Research Institute (EERI)
- Architectural Institute of Japan (AIJ)
- Remote Sensing Society of Japan (RSSJ)

Name:
Hiroyuki Miura

Affiliation:
Associate Professor, Department of Architecture, Hiroshima University

Address:
A2-823, 1-4-1 Kagamiyama, Higashi-Hiroshima 739-8527, Japan

Brief Career:
2004 Postdoctoral Research Fellow, Tokyo Institute of Technology
2007 Assistant Professor, Tokyo Institute of Technology
2012- Associate Professor, Hiroshima University

Selected Publications:
- Miura, Midorikawa and Kerle, "Detection of Building Damage Areas of the 2006 Central Java, Indonesia Earthquake through Digital Analysis of Optical Satellite Images," Earthquake Spectra, Vol.28, 2012.
- Miura, Midorikawa, Fujimoto, Pacheco and Yamanaka, "Earthquake Damage Estimation in Metro Manila, Philippines Based on Seismic Performance of Buildings Evaluated by Local Experts' Judgments," Soil Dynamics and Earthquake Engineering, Vol.28, 2008.
- Miura and Midorikawa, "Updating GIS Building Inventory Data Using High-Resolution Satellite Images for Earthquake Damage Assessment: Application to Metro Manila, Philippines," Earthquake Spectra, Vol.22, 2006.

Academic Societies & Scientific Organizations:
- Earthquake Engineering Research Institute (EERI)
- Architectural Institute of Japan (AIJ)
- Japan Association for Earthquake Engineering (JAEE)

Name:
Saburoh Midorikawa

Affiliation:
Professor, Department of Built Environment, Tokyo Institute of Technology

Address:
4259-G3-3, Nagatsuta, Midori-ku, Yokohama 226-8502, Japan

Brief Career:
1988 Associate Professor, Tokyo Institute of Technology
1989 Visiting Professor, Catholic University of Chile
1995- Professor, Tokyo Institute of Technology

Selected Publications:
- Joshi and Midorikawa, "Attenuation Characteristics of Ground Motion Intensity from Earthquakes with Intermediate Depth," Journal of Seismology, Vol.9, 2005.
- Matsuoka, Wakamatsu, Midorikawa, and Fujimoto, "Average Shear-wave Velocity Mapping Using Japan Engineering Geomorphologic Classification Map," Journal of Structural Eng. and Earthq. Eng., JSCE, Vol.23, 2006.
- Midorikawa, "Recent Seismic Microzoning Maps in Japan," Journal of Disaster Research, Vol.1, No.2, 2006.

Academic Societies & Scientific Organizations:
- Earthquake Engineering Research Institute (EERI)
- Architectural Institute of Japan (AIJ)
- Institute of Social Safety Science (ISSS)

Name:
Miguel Estrada

Affiliation:
General Director, CISMID
Associate Professor, Faculty of Civil Engineering, National University of Engineering

Address:
Av. Tupac Amaru 1150, Rimac, Lima, Peru

Brief Career:
1998-2000 Master of Engineering in the field of Civil Engineering, The University of Tokyo
2000-2004 Ph.D. of Civil Engineering, The University of Tokyo
2004-present Associate Professor, Faculty of Civil Engineering, National University of Engineering
2013-present General Director, Japan-Peru Center for Earthquake Engineering Research and Disaster Mitigation (CISMID), Faculty of Civil Engineering, National University of Engineering

Selected Publications:
- M. Estrada, H. Miura, F. Yamazaki, and S. Midorikawa, "Evaluation of Social Seismic Vulnerability through High Resolution Satellite Imagery," 15th World Conference on Earthquake Engineering, Portugal, 2012.
- M. Estrada, C. Zavala, and Z. Aguilar, "Use of Geomatics for Disaster Management – Case Study 2007 Peru, Pisco Earthquake," 7th International Workshop on Remote Sensing and Disaster Response, USA, 2009.
- M. Estrada, M. Matsuoka, and F. Yamazaki, "Use of Optical Satellite Images for the Recognition of Areas Damaged by Earthquakes" 6th International Conference on Seismic Zonation, USA, 2000.

Academic Societies & Scientific Organizations:
- Peruvian Board of Engineers
- Earthquake Engineering Research Institute