13th World Conference on Earthquake Engineering
Vancouver, B.C., Canada
August 1-6, 2004
Paper No. 678

AUTOMATED BUILDING DETECTION FROM HIGH-RESOLUTION SATELLITE IMAGE FOR UPDATING GIS BUILDING INVENTORY DATA

Hiroyuki MIURA[1], Saburoh MIDORIKAWA[2] and Kazuo FUJIMOTO[3]

SUMMARY

GIS building inventory data is a basic data to practice reliable earthquake damage assessment. In this study, a methodology of automated building detection from a high-resolution satellite image is proposed in order to update GIS building inventory data concisely. Newly constructed buildings not included in the GIS data are extracted from the difference between the image and the GIS data. Considering the location between a building edge and its shadow, the edge-based analysis is conducted to discriminate the building edge from others. We examine the applicability of the method using the IKONOS images in Yokohama city, Japan and Metro Manila, Philippines. The result shows that 90 % of mid-rise and high-rise buildings are detected successfully. The numbers of stories of the buildings are estimated using the shadow length in the image. The GIS building inventory data in Metro Manila is updated using the proposed method together with the land cover classification map. The building damage due to a scenario earthquake is assessed by means of simplified procedure based on the seismic capacity of the buildings and the computed ground motion. We compare the damage estimation between using the existing building inventory data and using the updated building inventory data. The result shows that the updated building inventory data provides more reliable damage estimation.

INTRODUCTION

Influx of population to urban areas is a common problem faced by developing countries. The number of mega-cities, which are vulnerable to disasters, is increasing. The earthquake loss estimation in mega-cities is indispensable to efficient earthquake disaster mitigation planning. The GIS (Geographic Information System) building inventory data is a basic data to practice the reliable loss estimation. Due to the rapid growth in the urban area, the number of buildings has been swelling. Not only the number of low-rise buildings, but also that of high-rise buildings is increasing especially in developed commercial zones. However, the system for updating GIS data continuously is hardly established in developing countries. As a result, the number of buildings that are not included in the GIS data is increasing.

Satellite remote sensing can capture land cover information in urban areas widely and continuously. Recently, high-resolution satellite images (e.g., IKONOS, QuickBird) with a ground resolution 1m or less

[1] Graduate Student, Tokyo Institute of Technology, Japan. Email: hmiura@enveng.titech.ac.jp
[2] Professor, Tokyo Institute of Technology, Japan. Email: smidorik@enveng.titech.ac.jp
[3] Research Associate, Tokyo Institute of Technology, Japan. Email: kazu@enveng.titech.ac.jp

have been available at relatively low cost. Individual building in an urban area can be identified in the images. Therefore, high-resolution satellite images are useful to grasp the individual location of buildings that are difficult to identify in middle-resolution satellite images (e.g., Landsat, SPOT).

In this study, a methodology of automated building detection from high-resolution satellite image is proposed in order to update GIS building inventory data concisely. The applicability of the method is examined using the data in Yokohama city, Japan and Metro Manila, Philippines. Combining the result together with the land cover classification map, the GIS building inventory data in Metro Manila is updated. In order to show the application of the updated building inventory data, the building damage due to a scenario earthquake is assessed by means of simplified procedure based on the seismic capacity of the buildings and the computed ground motion.

STUDY AREA AND DATA SOURCES

The objective of this study is to propose a methodology for updating GIS building inventory data in developing countries. In Metro Manila, the capital of Philippines, the population concentration and urban sprawl have been strongly observed. In order for the reliable loss estimation, the up-to-date building inventory data should be constructed. Therefore, the purposed area of this study is Metro Manila. In order to examine the applicability of the method, the data in Yokohama city, Japan and Metro Manila are used. One of the reasons why we use the data in Yokohama is that the detailed GIS data is constructed. The other is that it is easy to collect the ground truth data for the verification of the analysis.

Yokohama city is located in the central part of Japan. The GIS building inventory data in Yokohama consists of the footprints and their attributes of buildings with a scale of 1/2,500. The GIS data is updated on every five years and latest updating was conducted in 1999. In Metro Manila, the GIS base map was constructed based on the topographic map with a scale of 1/10,000 edited in 1989. In the base map, the footprints of the buildings and the congested housing areas are included. The total number of buildings is about 910,000 as of 1989 (Sarausad [1]). The GIS data has not been updated since 1989. Therefore, the buildings newly constructed after 1989 are not included in the GIS data.

The satellite IKONOS images with the ground resolution of 1 m are used in this study. These images are composed of three bands with visible range (Blue, Green, Red) and one band with near infrared range (NIR). The acquisition conditions of the images are shown in Table 1. In order to register the images to the GIS data, the images are geometrically corrected using ground control points and Affine transformation with the residuals within five pixels. Figure 1 shows the comparison of the IKONOS image in Yokohama acquired in 2001 and the GIS data in 1999. The arrows indicate the buildings newly constructed after 1999. In order to update the GIS building inventory data, the newly constructed buildings should be added in the GIS data.

FLOW OF UPDATING GIS BUILDING INVENTORY DATA

The flow of updating GIS building inventory data employed in this study is shown in Fig. 2. The distribution of mid-rise and high-rise buildings is captured using the method of automated building detection from the high-resolution satellite images. The distribution of low-rise buildings is estimated from the land cover classification map based on the multi-temporal middle-resolution satellite images.

The locations of the newly constructed mid-rise and high-rise buildings are detected from the difference between the IKONOS image and the GIS data. In this method, the edges are extracted from the image in order to detect the boundaries of the buildings. Considering the location between a building edge and its shadow, the edge-based analysis is conducted to discriminate the building edge from others. The regions

Table 1 Conditions of IKONOS images used in this study

Area	Acquisition Date	Sun Azimuth(deg.) Sun Elevation(deg.)	Sensor Azimuth(deg.) Sensor Elevation(deg.)
Yokohama	2001/Jan./6 AM 10:21	N157E 29	N112E 82
Metro Manila	2001/Sep./28 AM 10:29	N152E 71	N130E 65

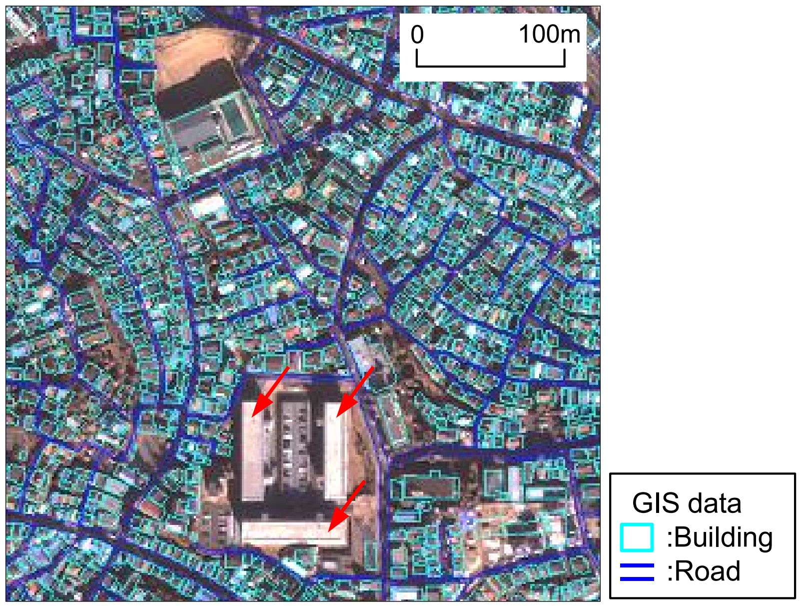


Fig. 1 Comparison of IKONOS image and GIS data

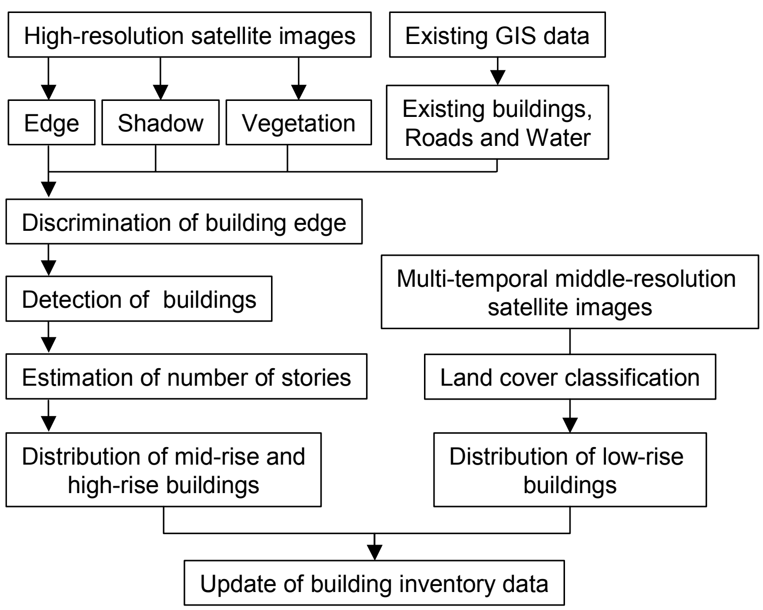


Fig. 2 Flow of Updating GIS building inventory data

including the building edges are extracted in the analysis. The numbers of stories of the buildings are estimated using the shadow length in the image.

It is difficult to detect individual locations of small low-rise buildings from the high-resolution satellite image. In this study, the distribution of low-rise buildings is estimated from the land cover classification map. The map was constructed from the multi-temporal Landsat images by Yamazaki *et al.* [2]. They classified the built-up areas into three categories according to the developing period. Considering that the newest built-up areas almost correspond with the distribution of the buildings constructed after the edit of the GIS data, we estimate the distribution and amount of the low-rise buildings. Combining the results of the analysis, the GIS building inventory data is updated.

AUTOMATED BUILDING DETECTION FROM HIGH-RESOLUTION SATELLTE IMAGE

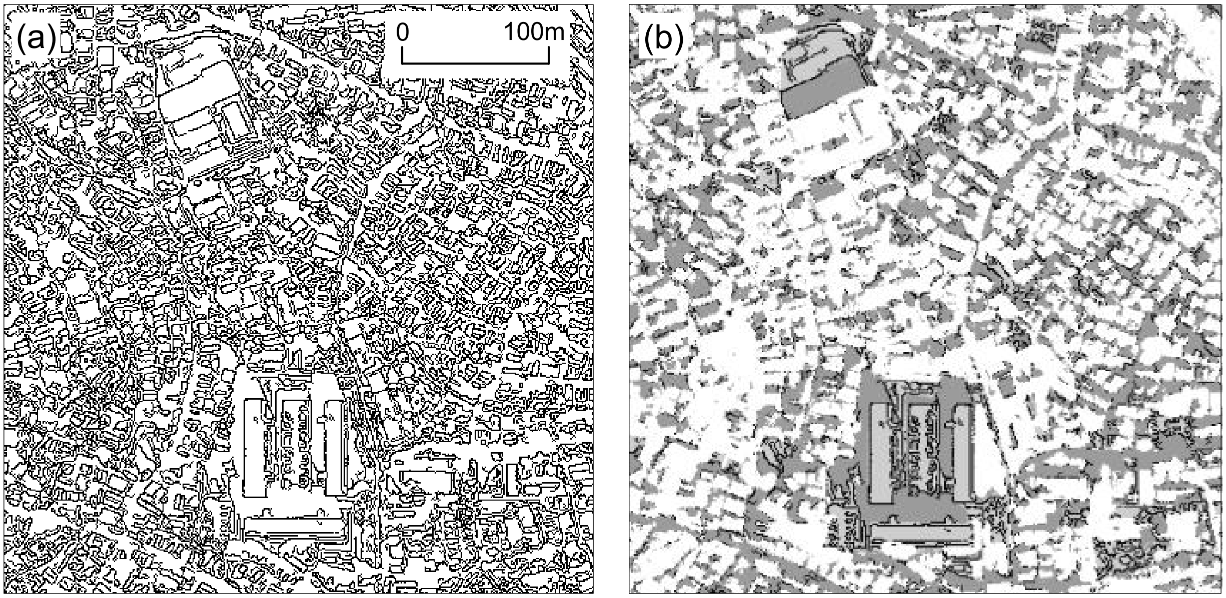
The edges are extracted from the IKONOS image using Canny's [3] method as shown in Fig. 3 (a). The edges of the target buildings are well extracted. However, the edges of other objects are also included. The shadow areas are derived from the threshold of the NIR band image. The lower peak in the brightness histogram is determined at the threshold value to separate the shadow areas. The vegetation areas are extracted using NDVI (Normalized Difference Vegetation Index) of the image. NDVI is computed from the NIR and Red band images (differences divided by sums). We determine the pixels whose NDVI show higher than 0.05 at the vegetation areas. In order to extract the newly constructed buildings, the existing buildings, the roads, and the water areas included in the GIS data should be avoided in the analysis. We define the shadow areas, the vegetation areas, the existing buildings, the roads, and the water areas as "the known areas". The edges in the known areas are eliminated from the edge image as shown in Fig. 3 (b). Black lines indicate the remaining edges. The edges of other objects (e.g., vacant ground) are still remaining. The building edges should be discriminated from the others.

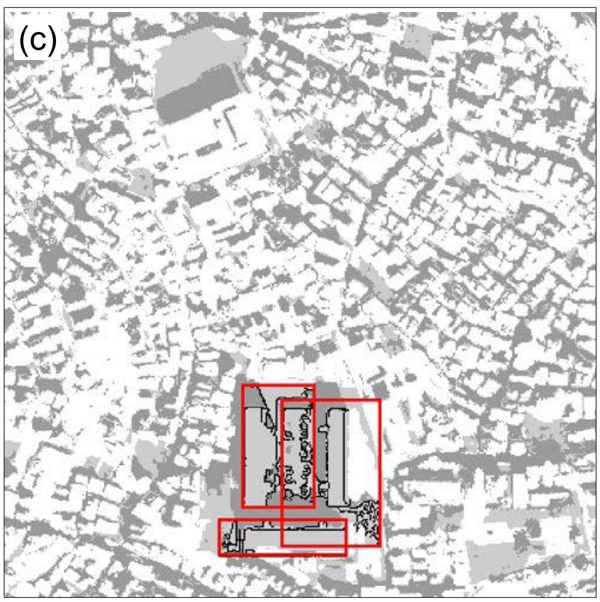
The characteristic of a building edge in the image is examined. Three-dimensional structures such as buildings generate shadows close together the structures in the image. The shadows are useful information to discriminate building edges from others. Considering the sun azimuth, the criterion to identify the building edge shown in Fig. 4 is included in the analysis. When all the gray pixels in the opposite direction of the sun correspond with the shadow pixels, the edge is identified as one part of the building edges.

Another criterion to identify the building edge is the edge length. The buildings have a certain length of edges in the image. If the threshold value for the edge length is small, numbers of mis-detections are extracted in the analysis. In order to reduce the mis-detections, the threshold value should be determined appropriately. According to the preliminary analysis using the images, the result in case of threshold value at 25 pixels shows less mis-detections and less un-detected buildings. Therefore, we determine the threshold value for edge length to be 25 pixels in the analysis.

The edges satisfying the criteria are extracted as the building edges. Figure 3 (c) shows the result of the analysis. The red squares indicate the regions circumscribing the building edges. Compared with Fig. 1, the approximate locations of the newly constructed buildings are detected successfully.

We apply the method to the built-up area in Yokohama. Figure 5 (a) shows the IKONOS image used in this study. Using the proposed image and the existing GIS data, newly constructed buildings are detected. The result of the analysis is shown in Fig. 5 (b). The rectangles indicate the locations of correctly detected buildings. The triangles and the circles represent the locations of mis-detected objects and un-detected buildings, respectively. In this area, 28 newly constructed buildings are included. As a result, 22 buildings, which correspond with almost 80 % of the target buildings, are detected in the analysis. The number of mis-detections is nine. The mis-detected objects represent the three-dimensional structures





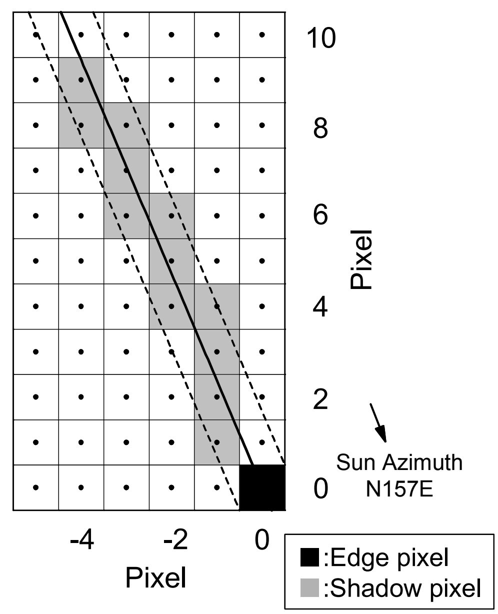
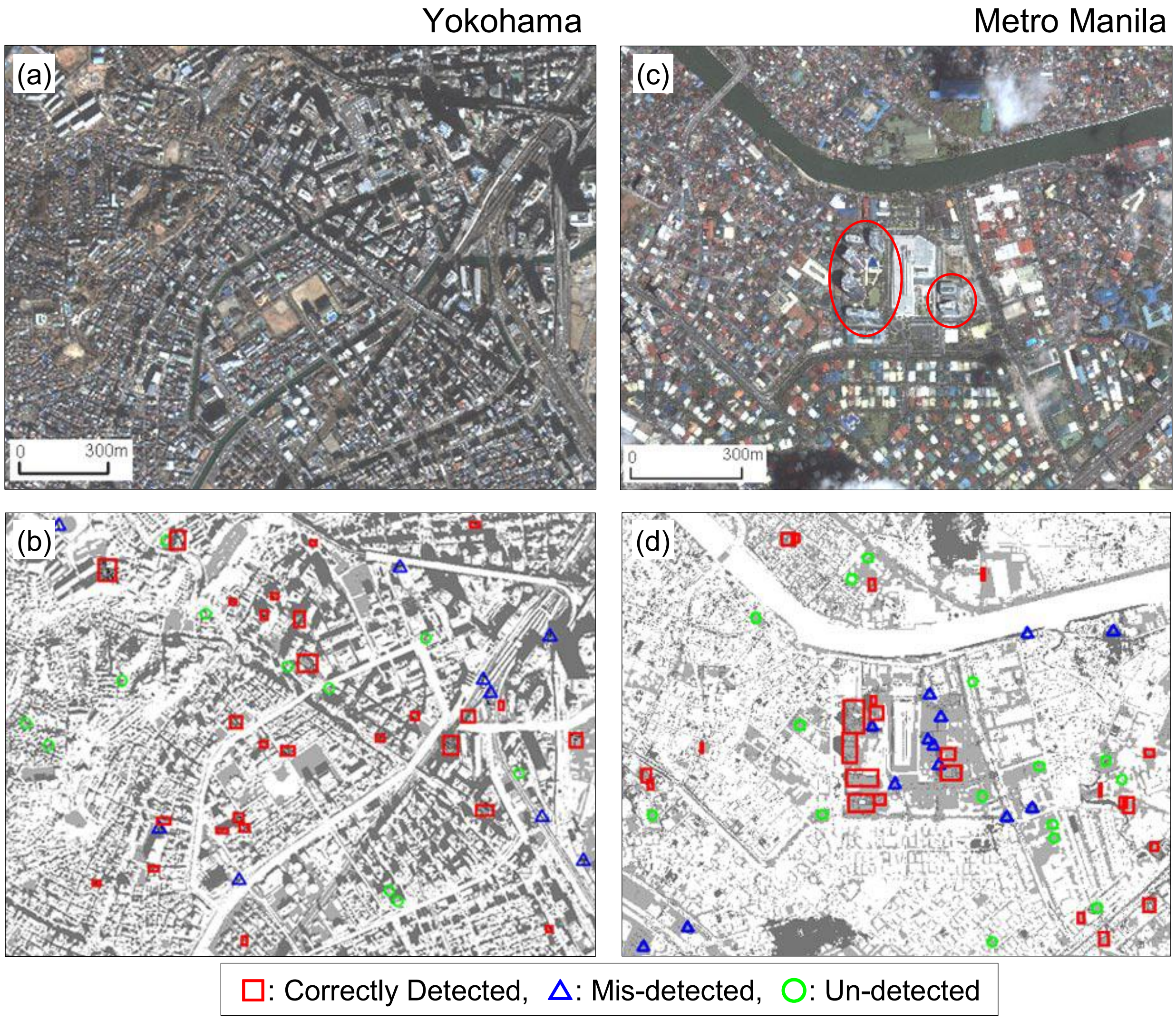


Fig. 3 (a) Extracted edges from image
(b) Elimination of edges in known areas
(Black: edge, Dark gray: shadow,
White: known area, Light gray: unknown)
(c) Result of analysis
(Red: extracted region, Black: edge)

Fig. 4 Digital criterion for discriminating building edge
(In case of image in Yokohama)

except buildings (e.g., walls in parking lot). Figure 5 (c) shows the IKONOS image including the recently developed commercial area in Metro Manila. The high-rise buildings shown in the circles are newly constructed after the edit of the GIS data. The result of the analysis shown in Fig. 5 (d) demonstrates that all of the high-rise buildings shown in Fig. 5 (c) are detected successfully.

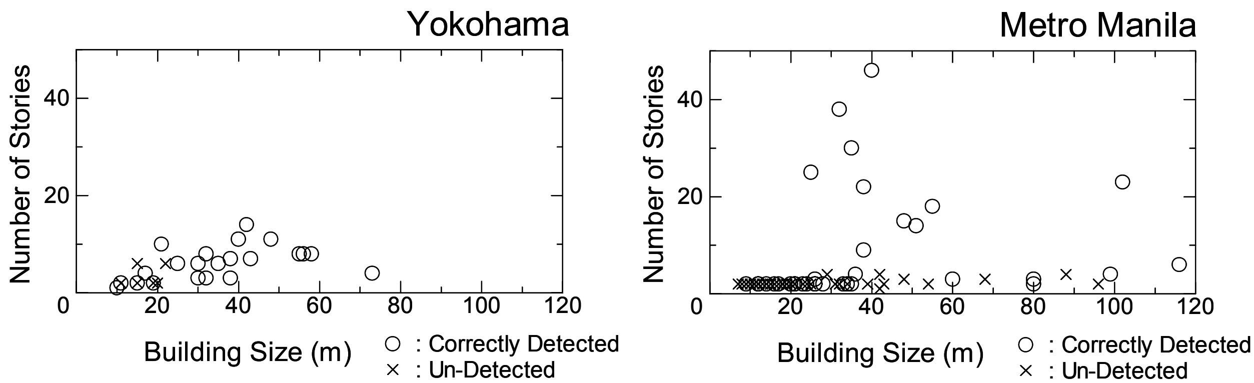
Figures 6 show the results of the analysis arranged according to the building size and the number of stories. A lot of low-rise buildings whose building sizes are less than 25 m are included especially in



Figs. 5 IKONOS images and results of analysis

Metro Manila. It is difficult to detect such small low-rise buildings individually by using the proposed method because the edges of the buildings and their shadows are too short to detect from the image. However, almost 90 % of the mid-rise and high-rise buildings are detected successfully. The schematic of the accuracy of the method is shown in Fig. 7. The numbers in the figure indicate the detection percentage of the buildings presented in each building size and number of stories. The percentages of the buildings lower than 3-story show 20 to 50 %. The lager the number of stories is, the higher the detection percentages show. The detection percentages of buildings higher than 4-story show almost 90 % or more. The result suggests that the proposed method is useful to grasp the individual locations of mid-rise and high-rise buildings.

We estimated the number of stories of detected buildings using the shadow length observed in the image. The number of stories is calculated from the shadow length, the sun elevation, and the average floor height set as 3.2 m. The result shows that the numbers of stories are estimated in almost all the buildings successfully within the error range of 2 stories.



Figs. 6 Results of analysis arranged according to building size and number of stories

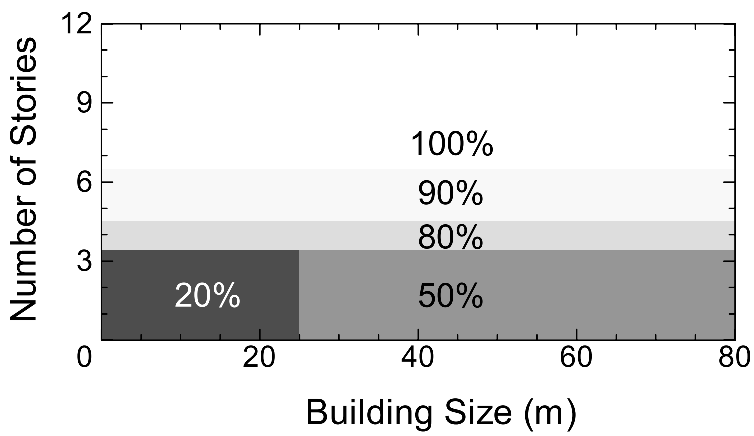


Fig. 7 Schematic of detection percentage in analysis

UPDATING OF GIS BUILDING INVENTORY DATA

In order for the reliable damage estimation of buildings in Metro Manila, the GIS building inventory data is updated using satellite remote sensing data. As shown in Fig. 2, the distribution of mid-rise and high-rise buildings is captured by the high-resolution satellite images using the proposed method. The distribution of low-rise buildings is estimated from the land cover classification map based on the multi-temporal middle-resolution satellite images.

The IKONOS images used in this study are shown in Fig. 8. These images covers about 75 % of Metro Manila. The proposed method is applied to the whole area of the images. Figure 9 shows the result of the analysis. The regions indicate the locations of detected buildings. About 2,600 newly constructed buildings are detected from the images. The numbers of stories of the buildings are estimated using their shadow lengths. The number of the detected mid-rise buildings (4-9 story), that of the high-rise buildings (10-30 story), and that of the super high-rise buildings (over 31 story) are 666, 202, and 24, respectively. In order to provide quantitative assessment of the analysis, the un-detected buildings are extracted by the manual interpretation with the images. The detection percentage of mid-rise and high-rise buildings is computed as shown in Table 2. The result shows that all the detection percentages are more than 90%. All the footprints of the newly constructed buildings are included in the GIS data.

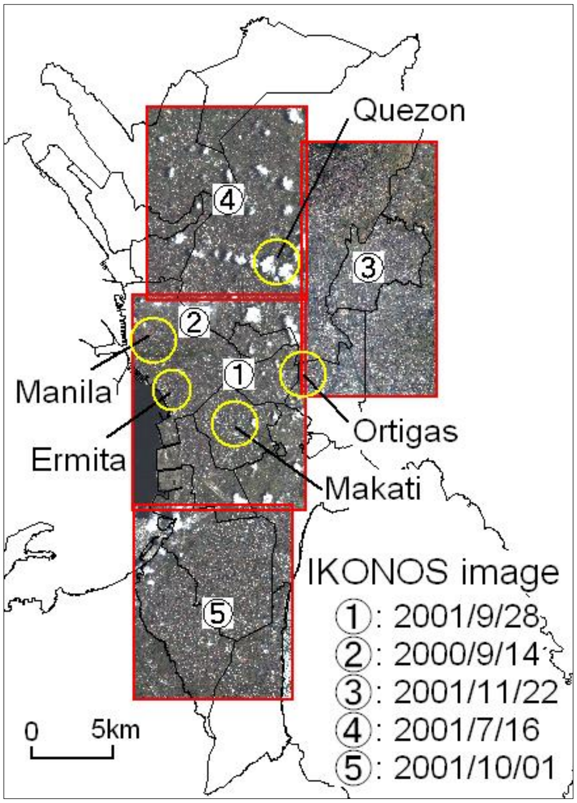


Fig. 8 Coverage area of IKONOS images

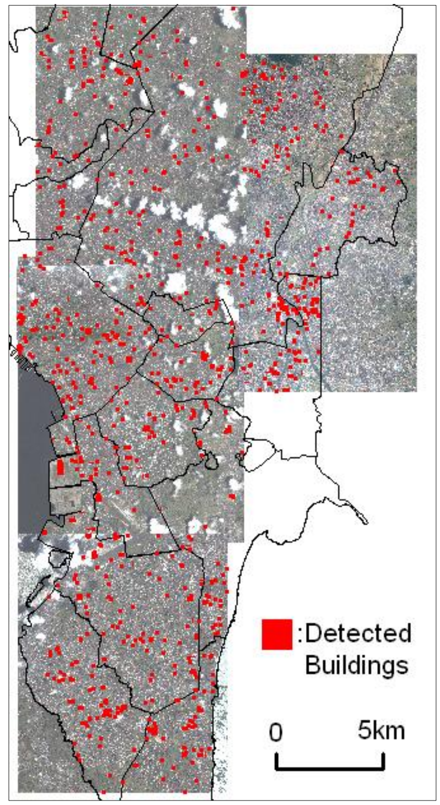


Fig. 9 Detected buildings from IKONOS images

Table 2 Detection percentage of analysis

The number of stories	Detected	Un-detected	Detection percentage
4 - 9	666	74	90%
10 - 30	202	7	97%
31 -	24	0	100%

It is difficult to detect the individual location of low-rise buildings accurately by using the proposed method. Yamazaki *et al.* [2] revealed the distribution of built-up area in Metro Manila using multi-temporal Landsat images (30m-resolution). They classified the built-up areas into the three categories according to the developing period (-1972, 1972-1992, 1992-). The older built-up areas before 1992 almost coincide with the building distribution of the existing GIS data. The built-up areas after 1992 correspond with the building distribution developed after the edit of the GIS data. We compare the number of the buildings of the existing data and the built-up areas. The number of the buildings in the existing data is about 910,000 (Sarausad [1]) while the built-up areas before 1992 is about 295 km^2 (330,000 pixels). Therefore, the number of buildings that are located in the one pixel is estimated at three. Considering that the built-up area after 1992 is about 124 km^2 (130,000 pixels), the number of the buildings in the built-up area after 1992 is estimated at about 380,000. Assuming that the most of the estimated buildings consist of low-rise buildings that are uniformly distributed in the built-up areas, they are added in the GIS data. Combining the distribution of mid-rise and high-rise buildings and that of low-rise buildings, the building inventory data is updated. The total number of the buildings is estimated at about 1,290,000, which almost corresponds with the result of the recent survey by JICA *et al.* [4].

APPLICATION TO BUILDING DAMAGE ESTIMATION

Using the updated GIS building inventory data, building damage assessment due to a scenario earthquake is conducted considering the seismic capacity of the buildings. The flow of building damage estimation is shown in Fig. 10. The ground motion due to a scenario earthquake was computed by the hybrid simulation method with the three-dimensional underground structure model and the soil response analysis using the surface soil profiles (Yamada *et al.* [5]). The building response is evaluated by the capacity spectrum method (FEMA [6]). The buildings in Metro Manila are classified into several categories. The nonlinear response of the buildings is estimated from the capacity curve and the ground motion spectrum. The damage rate for each building category is determined by the building response and the fragility curve. Multiplying the damage rate of each building category and the updated building inventory data, the distribution of the damaged buildings is computed.

The building types proposed by Vibrametrics, Inc. [7] are used in this study. They classified the buildings into three major categories; CHB (concrete hollow block), C1 (concrete moment frame building), and C2 (concrete shear wall building), and twenty detailed categories considering the possible height range and the design vintage for each structural type. Taking account of the height range and the typical design vintage, we classify the buildings into seven categories shown in Table 3. The capacity curves and the fragility curves proposed by Vibrametrics, Inc. [7] are used in the capacity spectrum method.

The West Valley Fault is selected as the source of a scenario earthquake because the fault is closer to the central part of Metro Manila. Figure 11 (a) shows the computed peak ground velocity on the surface due to the scenario earthquake (M 6.7) with the 500 m mesh system. The large ground shaking is computed in the areas that are located on the thick soft soil such as the Coastal lowland and the Marikina valley.

The number of the damaged buildings is computed by multiplying the damage rate and the number of the buildings. Figure 11 (b) shows the distribution of the low-rise buildings (1-3 story) with the complete or extensive damage level. The damage of low-rise buildings is concentrated in the Coastal lowland and the Marikina valley. The number of the damaged buildings is about 181,600, which corresponds with almost 14% of the low-rise buildings in Metro Manila.

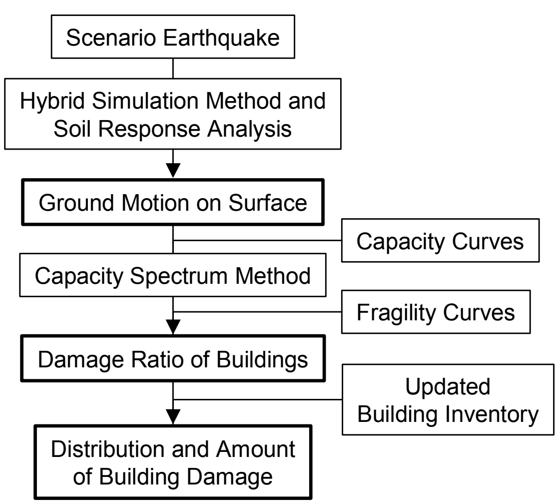


Fig. 10 Flow of building damage estimation

Table 3 Building types used in this study

Building Type	Structural Type	Number of Stories	Design Vintage
CHB	Concrete Hollow Block	1 - 3	Sub-Type 3
C1L	Concrete Moment Frame	1 - 3	Sub-Type 3
C1M		4 - 7	Sub-Type 3
C1H		8 - 15	Sub-Type 2
C2V	Concrete Shear Wall	16 - 25	Sub-Type 1
C2E		26 - 35	Sub-Type 1
C2S		36 -	Sub-Type 1

Sub-Type 1 : Constructed after 1992
Sub-Type 2 : Constructed from 1972 to 1991
Sub-Type 3 : Constructed before 1971

The distribution of mid-rise buildings (4-7 story) with the complete or extensive damage level is shown in Fig. 11 (c). The damage is concentrated in Manila because a lot of mid-rise buildings are located in Manila. About 22% of mid-rise buildings are damaged. The distributions of the high-rise buildings with the moderate damage level are shown in Fig. 11 (d)-(f). The building damages for the higher types are relatively slight. Most of the high-rise buildings may suffer less than moderate damage level. One of the reasons is the spectral characteristic of ground motion. The magnitude of the scenario earthquake (6.7) is not large enough to generate the strong ground motion with longer period, which contributed to the response of higher buildings.

Table 4 shows the comparison of the damage estimation between using the existing building inventory data and using the updated building inventory data. The numbers in Table 4 indicate the number of the damaged buildings, the damage ratio, and the increase of the damaged buildings, respectively. The numbers of the damaged 1-3 story buildings based on the updated building inventory increase by about 40 % compared with the estimation based on the existing building inventory. The numbers of 4-7 story buildings and 8-15 story buildings with the complete or extensive damage also increase about 30 % and 50 %, respectively. The number of over 16 story building with the moderate damage increase by more than 70 %. These results indicate that the updated building inventory data reveals the wider area damage and provides more reliable result of the damage estimation.

Table 4 Comparison of damage estimation between using existing building inventory data and using updated building inventory data

	Using Existing Building Inventory		Using Updated Building Inventory		Increase	
	A: No. of Damaged Bldgs	**B**: Ratio	**C**: No. of Damaged Bldgs	**D**: Ratio	**C-A**: No.	**(C-A)/A**: Ratio
1-3 story Bldgs						
Complete + Extensive	132,800	14%	181,600	14%	48,800	37%
Moderate	49,700	5%	123,300	10%	35,100	40%
Total No. of Bldgs	908,500	-	1,281,400	-	372,900	41%
4-7 story Bldgs						
Complete + Extensive	501	37%	647	22%	136	27%
Moderate	310	22%	413	14%	103	33%
Total No. of Bldgs	1,382	-	2,869	-	1,487	108%
8-15 story Bldgs						
Complete + Extensive	63	15%	96	12%	33	52%
Moderate	235	58%	452	56%	217	92%
Total No. of Bldgs	408	-	812	-	404	99%
Over 16 story Bldgs						
Complete + Extensive	0	0%	1	0.2%	1	100%
Moderate	46	19%	80	19%	34	74%
Total No. of Bldgs	239	-	411	-	172	72%

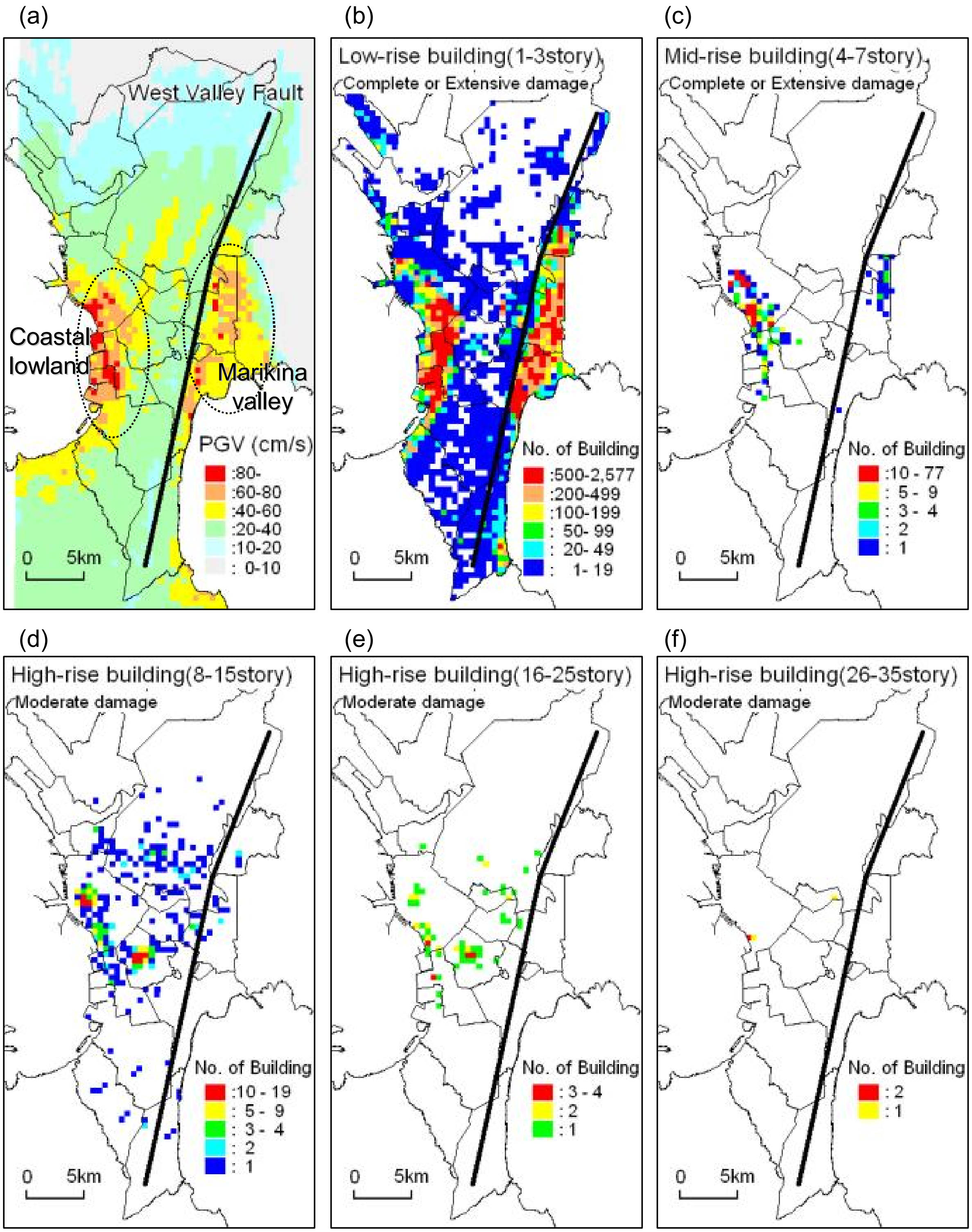


Fig. 11 (a) Computed peak ground velocity on surface
(b) Distribution of Low-rise building with complete or extensive damage
(c) Distribution of Mid-rise building with complete or extensive damage
(d) Distribution of High-rise building (8-15 story) with moderate damage
(e) Distribution of High-rise building (16-25 story) with moderate damage
(f) Distribution of High-rise building (26-35 story) with moderate damage

CONCLUSIONS

A methodology of automated building detection from high-resolution satellite image is proposed in order to update GIS building inventory data concisely. We examined the applicability of the method using the IKONOS images in Yokohama city, Japan and Metro Manila, Philippines. The result shows that the detection percentage for mid-rise and high-rise buildings is almost 90 %. The method is useful to grasp the locations of mid-rise and high-rise buildings. The GIS building inventory data in Metro Manila is updated using the proposed method together with the land cover classification map. The distribution of mid-rise and high-rise buildings is captured using the proposed method. The distribution of low-rise buildings is estimated from the land cover classification map based on the multi-temporal Landsat images. The building damage due to a scenario earthquake is assessed by means of simplified procedure based on the seismic capacity of the buildings and the computed ground motion. The result shows that the updated building inventory data provides the reliable result of the damage estimation.

ACKNOWLEDGMENTS

The authors acknowledge Ass. Prof. Hiroaki Yamanaka (Tokyo Institute of Technology) and Dr. Masashi Matsuoka (Earthquake Disaster Mitigation Research Center) for providing the ground motion simulation data and the land cover classification data, respectively. This study was done as a part of the Development of Earthquake and Tsunami Disaster Mitigation Technologies and Their Integration For the Asia-Pacific Region (EqTAP) project.

REFERENCES

1. Sarausad, F. “Infrastructure Condition in Metro Manila”, *Disaster Prevention and Mitigation in Metropolitan Manila Area*, 1993; 93-107
2. Yamazaki, F. *et al.* “Urban Classification of Metro Manila for Seismic Risk Assessment Using Satellite Images”, *The 5th Multi-Lateral Workshop on Development of Earthquake and Tsunami Disaster Mitigation Technologies and Their Integration for the Asia-Pacific Region*, EDM Technical Report No.16, Earthquake Disaster Mitigation Research Center, NIED, 2003, (CD-ROM)
3. Canny, J. “A Computational Approach to Edge Detection”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1986; 8(6), 679-698
4. Japan International Cooperation Agency (JICA) *et al.* “*Earthquake Impact Reduction Study for Metropolitan Manila, Republic of the Philippines*”, Progress Report 2, 2003
5. Yamada, N. *et al.* “Strong Ground Motion Simulation in Metro Manila”, *Journal of Structural Engineering*, 2003; 49B, 1-6 (in Japanese)
6. Federal Emergency Management Agency (FEMA) “*HAZUS99 SR2 Technical Manual*”, Washington D.C., 1999
7. Vibrametrics, Inc. “*Development of Fragility Curves for Selected Building Types*”, EqTAP Metro Manila Case Study, Final Report, 2003