

# High-Resolution Urban Area Mapping of World's Cities Using Satellite Images

Hiroyuki Miyazaki, Koki Iwao and Ryosuke Shibasaki

**Abstract:** We present development of global urban area map using ASTER satellite images, which has much higher resolution than that of existing global urban area maps. Moreover, we validated accuracy of the map using ground truth database developed from global gazetteer. Our map will contributed to measuring geographical feature of urban (e.g. compactness, contiguity, and openness) and modelling land-cover and land-use change with socio-economical variables.

**Keywords:** Urban area mapping, ASTER satellite images, gazetteer

## 1. Introduction

Urbanization has been a main concern for regional and global environmental change (Foley et al., 2005) and socio-economic problems (Angel et al., 2005). Various kinds of studies have used satellite-derived global urban area maps to evaluate critical aspects of urbanization for global environmental change, such as size, scale and form of cities and conversion of land cover. The studies using global urban area map had provided valuable information of urbanization especially for less documented regions. As the studies on urbanization progressed, however, 1-km spatial resolution of global urban area map have gotten obsolete for measuring spatial structure of urban area in fine scale (Angel et al., 2005) and for modeling land use conversion with socio-economical variables (Nelson and Robertson, 2007).

Developing high-resolution urban area map would be key issue for promoting new insights on urban dynamics. Several studies had developed urban area map only for their cases using using high-resolution satellite images (e.g. Landsat, Terra/ASTER, IKONOS and Quickbird) and shown valuable outcome; however mapping urban area using high-resolution satellite images involves considerable time and labor cost. It prevents comprehensive and comparative studies on urban dynamics on world's cities.

Another issue is scarcity in ground truth data for global urban area maps. Several ground truth database had been developed; however they don't have enough amount of data for urban area. For consistent evaluation of global urban area maps, it is needed to develop ground truth database specifically for urban area.

In this paper, we present developing global urban area

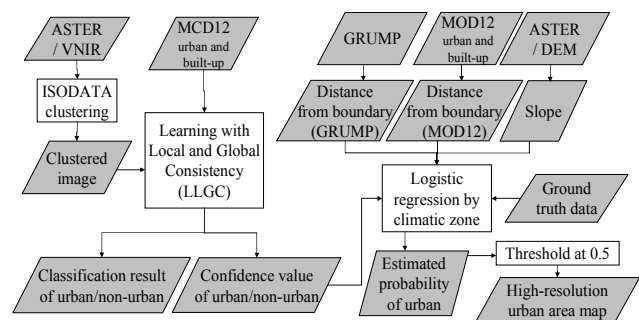


Figure 1. Overview of mapping urban area in high resolution

map in 15-m resolution from ASTER satellite images and accuracy assessment of the new map using global gazetteer. Regarding definition of urban, we introduced definition discussed by (Potere and Schneider, 2007) ; they defined urban with presence of built-up area.

## 2. Method

### 2.1. Global Urban Area Mapping using High-Resolution Satellite Images

To classify pixels of satellite image into urban or non-urban, we have to acquire training data in advance of classification, because pixel itself does not have definitive information of urban or non-urban. Traditionally, training data is acquired by human visual interpretation; however it is extraordinary time- and labor-consuming for global coverage. To reduce such cost, we employed existing coarse-resolution global urban area maps as training data. However gap in spatial resolution causes inconsistency among pixel values and classes. For instance, if a pixel likely to be urban and another pixel likely to be non-urban are covered within an urban pixel of coarse resolution, each pixel is assigned to training data of urban even though they should be separated into urban and non-urban.

To deal with the gap in spatial resolution, we introduced Learning with Local and Global Consistency (LLGC). LLGC constructs a function to correct roughly labeled classification into smoothly labeled result (Zhou

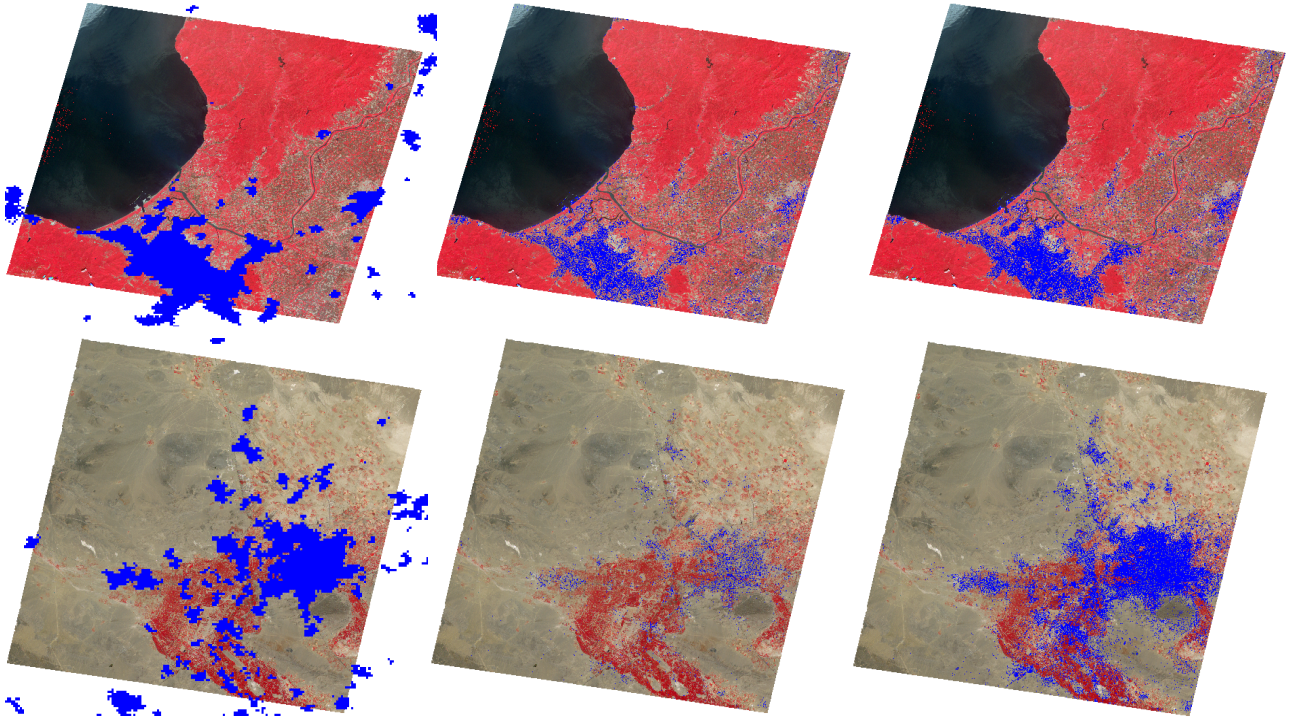


Figure 2. Urban area maps derived from MCD12Q1 (left), LLGC (middle) and integrated map (right) for Sapporo, Japan (upper; 43.28°, 141.46°) and Esfahan, Iran (lower; 32.73°, 51.5°). Blue pixel represents urban area. Background images are false-color composite image of ASTER/VNIR.

et al., 2003). The method was thus suitable for our case, in which pixels of high-resolution satellite images were initially labeled with coarse-resolution urban area map.

Although the clusters were successfully classified by LLGC, the result would include misclassifications due to similarity in surface reflectance among different land covers. For example, urban area surrounded with desert could be classified as non-urban because surface reflectance of urban is similar to desert. Also, cloud cover would considerably lead to misclassification.

To reduce the misclassifications, we integrated the result of LLGC with existing urban area maps by introducing logistic regression. Logistic regression has an advantage in representing presence of urban in form of probability, which could represent spatial gradient of built-up density between urban and rural area. We constructed a model to estimate probability of presence of urban as equation (1).

$$P_i(\text{urban}) = \frac{\exp(U_i)}{1 + \exp(U_i)} \quad (1)$$

where  $P_i(\text{urban})$  is the probability of presence of urban at  $i$ th pixel and  $U_i$  is defined in form of polynomial expression as equation (2).

$$U_i = \beta_0 + \beta_1 \times LLGC_i + \beta_2 \times SLOPE_i + \beta_3 \times DIST_{MOD12,i} + \beta_4 \times DIST_{GRUMP,i} \quad (2)$$

where  $\beta$  is coefficient for each variable,  $LLGC_i$  is confidence value of LLGC at  $i$ th pixel,  $SLOPE_i$  is slope at  $i$ th pixel,  $DBU_{MOD12,i}$  and  $DBU_{GRUMP,i}$  is DBU at  $i$ th pixel in MOD12Q1 (available at <http://www-modis.bu.edu/landcover/userguide/c/>), GRUMP Urban Extent Grid (GRUMP UE; available at <http://sedac.ciesin.columbia.edu/gpw/>), respectively (positive value for outside urban area; negative value for inside urban area).

## 2.2. Accuracy Assessment using Global Gazetteer

Accuracy assessment is essential for publishing the map to end users. It requires ground truth data, which represents urban or non-urban in more trustful way than the mapping method mentioned above. Several ground truth database had been developed for validation of global land cover maps; however they includes only less than 100 ground truth data, the number of which is not sufficient to validate urban area maps of the world.

Therefore we developed ground truth database for global urban area map. We employed GRUMP Settlement Points (GRUMP SP; available at

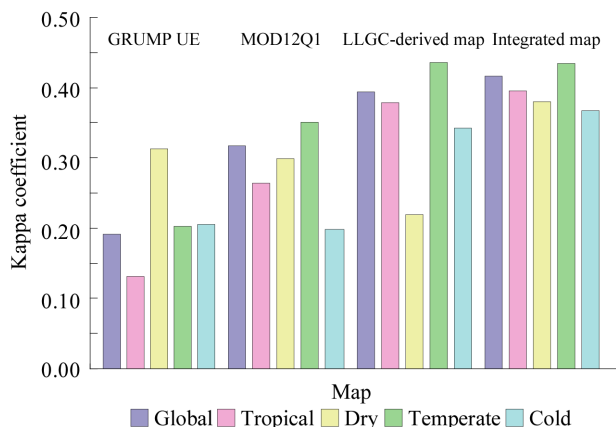


Figure 3. Kappa coefficient calculated by maps and climatic zone.

<http://sedac.ciesin.columbia.edu/gpw/>), a global gazetteer of populated places, as a source of ground truth database. It includes 55417 point coordinates of populated places, which could be placed on urban area. We visually interpreted land cover (urban or non-urban) at 3734 point coordinates of populated places of more than 0.1 million people based on color and texture of 15-m-resolution satellite images taken by ASTER.

To complement non-urban ground truth data, we also employed Degree Confluence Project data, which contains ground information collected at a  $1^\circ \times 1^\circ$  grid of latitude and longitude, called confluence (Iwao et al., 2006). We interpreted land cover of urban or non-urban at 4211 confluences.

### 3. Result and Discussion

#### 3.1. Urban Area Mapping using ASTER Satellite Images

We applied the method on 775 scenes of ASTER satellite images. Figure 2 shows some results of them. Urban clusters refined into 15-m resolution by applying LLGC and logistic regression, representing spatial structure of urban area in detail whereas existing coarse-resolution map represents the clusters with only a few pixels.

We found that some urban areas surrounded by bare land or desert were omitted by LLGC. It is because bare land and desert have similar characteristic in surface reflectance to urban area; thus LLGC misclassified actual urban area as bare land or desert.

However, by integrating LLGC result with existing urban area map and DEM, the omitted urban area had successfully detected. Therefore we might conclude the integration improved accuracy of urban area map.

#### 3.2. Accuracy Assessment on the New Map

We visually interpreted land cover of urban or

non-urban at 3734 point coordinates of populated places derived from GRUMP SP and 4211 confluences of DCP. As the result, we acquired ground truth data of urban at 7744 sites of the world (2185 data points of urban and 5559 of non-urban). We calculated kappa coefficient (Foody, 2006), which is commonly used for accuracy assessment of land cover classification using remote sensing, for global coverage and each climatic zone (Figure 3).

The result shows that accuracies of the integrated map were better than existing urban area maps and LLGC-derived map for global coverage and for all of climatic zones. It indicates that the integration worked well for improving accuracy of urban area map.

Another interest feature of the result was significant improvement for dry zone. LLGC-derived maps of dry zone were much poorer than existing urban area maps due to misclassifications caused by similarity in surface reflectance. However, as mentioned above, integration with existing map corrected the misclassifications, and it was reflected to improvement in accuracy.

#### 3.3. Mapping Urban Area over Broader Region

The final goal of our work is archiving global coverage of high-resolution urban area map; thus we applied the method to broader region rather than each scene of ASTER satellite image.

Figure 4 shows the result for Tokyo Metropolitan Region in Japan and Taipei city in Taiwan. The method classified 15-m-resolution pixels into urban and non-urban correctly as a whole, indicating that the method is compatible not only to extent of ASTER scene, but broad urban area whose size is larger than extent of ASTER scene.

The considerable problem is that the maps include much amount of minor misclassification, though such misclassifications of the other regions could be identified and solved by evaluating agreement with existing urban area maps.

Another possible problem is availability of cloud-free images. Some cities, especially in Southeast Asia, have not been taken any satellite image without either cloud or haze. To complete cloud-free images for such cities, we will assess image quality by pixel and it enables to patch poor-quality pixels with good-quality pixels.

### 4. Conclusion

We developed method for mapping urban area of world's cities in high resolution using LLGC and logistic regression. The method worked well on 775 scenes of ASTER satellite images; it improved not only spatial resolution into 15 m, but also classification accuracy.



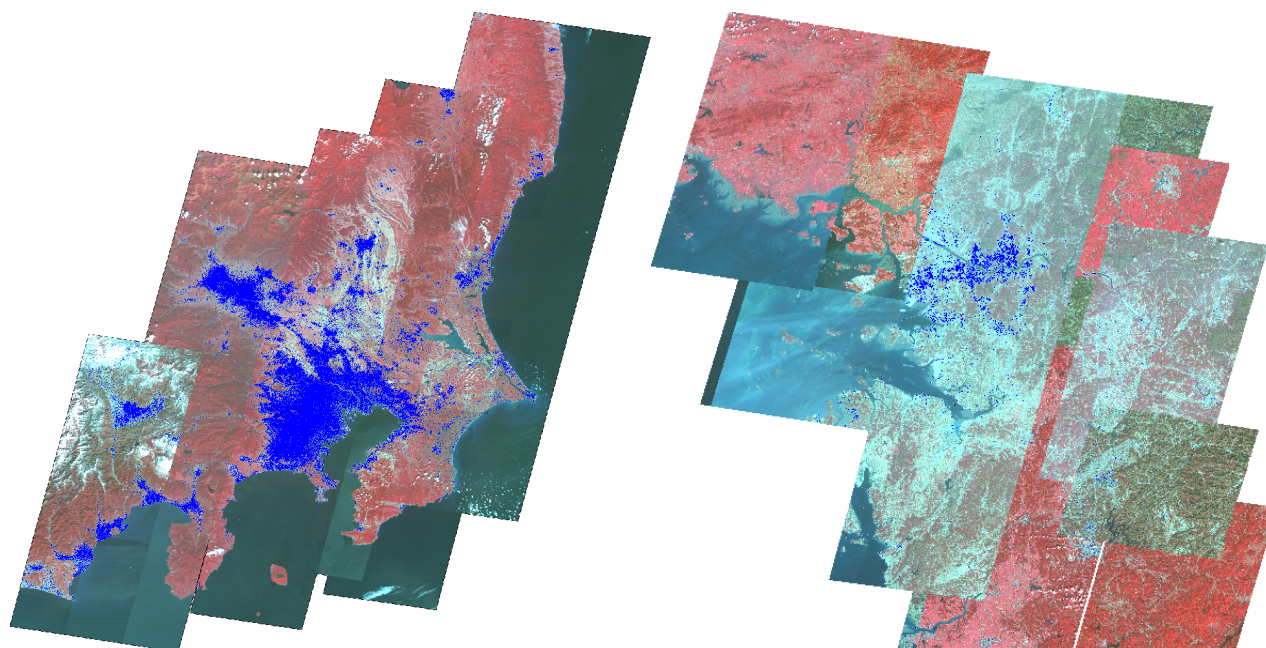


Figure 4. Experimental product of urban area map of high resolution for Tokyo Metropolitan Region, Japan (left) and Seoul, South Korea (right). Blue pixels represent urban area; back ground image is false-color composite of ASTER/VNIR.

We also developed ground truth database for urban area of the world. It includes 2185 data points of urban and 5559 of non-urban, the number of which was much more than that of existing ground truth database. The database would contribute to evaluating quality and characteristics of global urban area maps, including our new urban area maps.

Finally, we experimentally developed urban area map of broader region rather than each scene of ASTER satellite images. The method refined urban cluster of coarse-resolution map into 15-m resolution very well, indicating that the method would be compatible to global coverage of world cities.

For further works, we will complete the new global urban area maps of high resolution by processing on GEO Grid at AIST of Japan. After a little tune up and revise of the method, we will publish the urban area map for end users of urban studies.

#### Acknowledgement

This research used ASTER Data beta processed by the AIST GEO Grid from ASTER Data owned by the Ministry of Economy, Trade and Industry of Japan. The study was supported by a grant from Grant-in-Aid for JSPS Fellows (22-2598).

#### References

Angel, S., Sheppard, S. C. and Civco, D. L. (2005). The

- Dynamics of Global Urban Expansion, Transport and Urban Development Department, The World Bank: 200.
- Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N. and Snyder, P. K. (2005). Global Consequences of Land Use. *Science* **309**(5734), 570-574.
- Foody, G. M. (2006). What is the difference between two maps? A remote sensor's view. *J. Geograph. Syst.* **8**(2), 119-130.
- Iwao, K., Nishida, K., Kinoshita, T. and Yamagata, Y. (2006). Validating land cover maps with Degree Confluence Project information. *Geophys. Res. Lett.* **33**, L23404.
- Nelson, G. C. and Robertson, R. D. (2007). Comparing the GLC2000 and GeoCover LC land cover datasets for use in economic modelling of land use. *Int. J. Remote Sens.* **28**(19), 4243-4262.
- Potere, D. and Schneider, A. (2007). A critical look at representations of urban areas in global maps. *GeoJournal* **69**(1), 55-80.
- Zhou, D., Bousquet, O., Lal, T. N., Weston, J. and Schölkopf, B. (2003). Learning with Local and Global Consistency. *Adv. Neural Inf. Process. Syst.* **16**, 321-328.