A PRELIMINARY STUDY ON THE METHOD FOR EXTRACTING BAMBOO GROVES IN CHIBA PREFECTURE, JAPAN USING ALOS/AVNIR-2 DATA

Y. TANIGAKI a,*, I. HARADA b, K. HARA a, b

a Graduate School of Informatics, Tokyo University of Information Science, 4-1 Onaridai, Wakaba-ku, Chiba 265-8501 Japan –ho9002ty@edu.tuis.ac.jp
b Dept. of Environmental Information, Tokyo University of Information Science, 4-1 Onaridai, Wakaba-ku, Chiba 265-8501 Japan –jharada, hara@tuis.rsch.ac.jp

Commission VIII, WG VIII/6

KEY WORDS: Bamboo, ALOS/AVNIR-2, Google Earth, Extraction, Land Cover, Landscape management, Environment, Japan

ABSTRACT:

Well-tended bamboo groves, once an integral element in the satoyama, Japan’s traditional countryside mosaic, have over the past half century been abandoned. These abandoned groves are vigorous and are now expanding, causing various problems, including loss of diversity and damage to cropland. To control and manage these groves, an effective yet efficient system for documenting and monitoring their distribution over a wide region is required. Previous researches, using either aerial photograph or satellite images, have been implemented, but these studies suffer from problems with statistical overlearning, or require too much labor and are thus not practical for wide area application. In this study, Google Earth image and field research were used for setting training data and ALOS/AVNIR-2 satellite images were used to classify land cover in a typical countryside area that includes bamboo groves. Both maximum likelihood and decision tree classification methods were employed, with the accuracy test area completely separated from the area used for acquiring training data, thus preventing the classifiers from overlearning and improving their practicability. The accuracy was assessed by comparing the classification results with existing land cover maps. Both classification methods produced almost the same Kappa coefficient (decision tree: 0.41, maximum likelihood: 0.48). Future research can be expected to improve accuracy with higher quality training data and images. The results of this study show that ALOS/AVNIR-2 images can be a useful tool for effective yet cost-efficient mapping and monitoring of bamboo groves at the region or prefecture level.

1. INTRODUCTION

This research focuses on two species of bamboo, Phyllostachys bambusoides Sieb. et Zucc. and P. pubescens Mazel, that from several centuries ago were cultivated for use as food and materials for manufacturing various tools, building, etc. Carefully managed bamboo groves were an important element in the Japanese countryside landscape, or satoyama. Over the past half century, however, bamboo products have been replaced by plastic and foreign imports, resulting in a drop in demand, and many of the traditional bamboo groves have thus been abandoned. Left uncut, bamboo groves tend to expand, and examples of expanding groves have been reported in the regions from the Kanto (Tokyo area) westwards to northern Kyushu. (Okutomi, 1996; Torii, 1998; Miyake et al., 2000; Imai et al., 2004).

Expanding bamboo groves create several problems, including the following:

1. Abandonment allows the groves to be overrun with shoots, preventing sunlight from penetrating the groves and thereby reducing the overall biodiversity. This environmental change prevents other species from getting sunlight and growing, resulting in a loss of biodiversity (Fukui et al., 2004; Isagi et al., 1998; Miyake et al., 2000)
2. Expanding groves infringe on agricultural land, and are very difficult to control
3. Abandoned groves, packed with dead stalks, disrupt the visual aspect of the beautiful satoyama landscape, and also often encourage illegal dumping of household and industrial garbage (Chiba Pref., 2008)

To control bamboo groves, efficient systems for documenting and monitoring their distribution over a wide region are required. Research on effective control techniques is ongoing (for example, Toyota et al., 2005; Suzuki et al., 2008), but a viable system for keeping the groves in check over a wide area is urgently required. The first step in developing this system is to decide where to focus the cutting efforts. For example, if the cutting is implemented in the wrong area the bamboo will soon grow back. Also, the cutting on unsuitable area such as steep slopes or place far from road can result in loss of efficiency.

To maximize the efficiency of control efforts, the spatial aspects of bamboo expansion must be clearly documented. For documenting the distribution and expansion of bamboo groves over a wide area, GIS analyses of remotely sensed data can be a useful tool. This sort of research, for example, can rely on interpretation of aerial photographs (Okutomi, 1996; Torii, 1998; Miyake et al., 2000), or on satellite imagery such as LANDSAT, MODIS, IKONOS or ASTER (Imai et al., 2004; Cho et al., 2006; Kamagata et al., 2006). When using low spatial resolution satellite imagery such as MODIS, LANDSAT or ASTER, however, the small bamboo groves are included with other elements as mixed pixels, and are thus often difficult to extract (for example, see Koizumi et al. for LANDSAT). On the other hand, when using images of aerial photographs or

* Corresponding author.
high spatial resolution satellite such as IKONOS or QuickBird, the area of a single scene is small. As a result, when applying these images to a wide area such as a prefecture, the cost required for orthorectifying the images and calibrating DN (= pixel value in this study) are high. In addition, in many previous studies the training area (area included in polygon showing land cover for which training data was established) and the area used for checking accuracy are insufficiently defined. In some cases, it is even impossible to verify that these two areas are separate. Another problem is that bamboo groves are no longer considered critical bamboo groves are no longer considered a critical issue for our life and thus funding for control programs is limited, requiring a cost-effective method for extracting the groves. Given this urgent need for effective and efficient mapping of bamboo groves over a wide area, this research develops a system using ALOS AVNIR-2 satellite imagery and Google Earth images (Google, 2009) to acquire training data. The wide area classification is then implemented utilizing the same ALOS/AVNIR-2 imagery. In this study, the high spatial resolution of the imagery prevents many problems with mixels; while the wide area covered by each scene reduces the amount of work required for orthorectification and DN calibration.

2. METHOD

2.1 Study Area

The research was implemented in a 25km × 15km target area located in northern part of Chiba Prefecture, on the outskirts of Tokyo (Figure 1). The geology in the area consists of flat-topped uplands and narrow alluvial valleys. Traditional land-use patterns feature rice paddies on the valley floor and various woodlands, including bamboo groves, on the slopes and uplands. The training data was obtained from land covers including bamboo within the area, and the accuracy check was performed in the Azeta area (Figure 1), where several narrow valleys cut deep into the uplands.

![Figure 1. Location of study area, training areas and accuracy check area](image)

2.2 Data

The following 6 types of data were utilized

1. ALOS/AVNIR-2 radiometric calibrated RGB and near infrared imagery through utilizing stable light source in the satellite (10m spatial resolution, acquired on 15 August, 2007; 15 November, 2007; 15 February, 2008; 1 April, 2008)
2. 10m grid DEM (GSI, 2009)
3. Road/Rail Line vector data on 1/2500 scale (GSI, 2007)
4. Images of Google Earth 5.0 (Google, 2009) which the Google Earth indicated that is taken on 31 December, 2004
5. Field survey data, utilized for setting training area and modifying existing vegetation map for accuracy check
6. Modified vegetation map made of field survey results and existing vegetation maps (Emura et al., 2009) to estimate accuracy of land cover classifiers (Figure 4)

2.3 Orthorectification of Satellite Imagery

Intersections from the Road/Rail Line data were used as reference to set GCPs for orthorectifying the ALOS/AVNIR-2 images. In general, it might be assumed that in cases like this Google Earth images would produce smaller gaps between the reference data and ALOS/AVNIR-2 images. In this research, however, distinct landmarks that could be recognized were very few, making accurate setting of GCPs difficult. In addition, the target bamboo groves were small in size, and usually widely separated from one another. As a result, using Google Earth images as reference would require a great deal of time setting GCPs for each of the many training areas, resulting in a large effort in proportion to the amount and quality of the acquired training data. The images were orthorectified so that the gap between the reference and images was less than 10m. The ortho images include RGB and infrared for four periods (August and November 2007; and February and April 2008), producing 16 layers. In addition, NDVI was calculated for each period, bringing the total to 20 layers for each adjusted AVNIR-2 images.

2.4 Setting Training Data

In general, land classification accuracy is high when the accuracy check is implemented in the same area where the training data is obtained. This accuracy, however, is often due to overlearning rather than true accuracy in classification. In these cases, the accuracy scores will drop if the check is performed outside the training area. To avoid this problem, the accuracy check was performed in an area separate from the training areas in this research. Based on field data and Google Earth images, training areas for the following land covers were set: bamboo grove, bare land, water, urban, grass and paddy field, deciduous forest and evergreen forest. All gaps between the intersections in the Road-Rail Line Data and Google Earth images was equal to or less than 10 meters. As a result, the gap between the adjusted AVNIR-2 images and the Google Earth images were equal to or less than 20m (Figure 2). To reduce the effect of these gaps, all training areas were shrunk by 20 meter (Figure 3). Finally, training data was built using adjusted AVNIR-2 images from the training areas (Figure 1, Table 1).

<table>
<thead>
<tr>
<th>Bamboo</th>
<th>Evergreen</th>
<th>Deciduous</th>
<th>Water</th>
<th>Grass &amp; paddy field</th>
<th>Dry field</th>
<th>Urban</th>
<th>Bare land</th>
</tr>
</thead>
<tbody>
<tr>
<td>total area (ha)</td>
<td>1.14</td>
<td>1.4</td>
<td>1.1</td>
<td>546.4</td>
<td>20.8</td>
<td>9.6</td>
<td>36.9</td>
</tr>
<tr>
<td>number</td>
<td>26</td>
<td>3</td>
<td>10</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

![Table 1. Number of training areas and total area for each land cover category](image)
2.5 Supervised Land Cover Classifications

Two classifications were performed and compared; one using the maximum likelihood classifier (MLC) which assumes normal distribution of DN and equal probability of each land cover (for example; Wadachi et al., 1976); and the other a decision tree classifier (DTC), employing the well-tested Gini coefficient and CART algorithms (Breiman et al., 1984). The condition for leaf pruning was

\[
\frac{A_t}{A_l} < 0.01
\]  

Where \( A_t \) = area assigned to the leaf
\( A_l \) = total area in training areas occupied by the same land cover as that of the leaf

2.6 Accuracy Check

Accuracies of MLC and DTC were assessed from the error matrices, including the Kappa Coefficient (Cohen, 1960), which was calculated by comparing the classification results to the vegetation map. In addition, to confirm the importance of separating the training and accuracy check areas, DTC and MLC accuracy checks were carried out utilizing the training data instead of the vegetation map.

3. RESULTS

3.1 Training Data

The distributions of DN and NDVI for each land cover class are shown in Figure 5. With the exception of a few instances, such as DN for February 2008, a clear separation appeared between bamboo groves and other land cover categories.
indicating that data and method used in this study have the capability of extracting bamboo groves.

3.2 Land Cover Classifications

The results of classifications by MLC and DTC are shown in Figure 6, and the error matrixes in Table 2 and Table 3. As a whole, DTC showed a lower sensibility towards the bamboo groves, but a higher specificity. The Kappa coefficient indicating overall accuracy for both methods was essentially the same, with only minor differences which could vary from place to place. In addition, the Kappa coefficient for bamboo (DTC: 0.81, MLC: 0.97) and overall Kappa coefficient (DTC: 0.97, MLC: 0.98) were very high when the accuracy check area was set in the training area. Table 2 and Table 3, however, indicate that deciduous forest, evergreen forest, and grass and paddy field are difficult to distinguish from bamboo. Also, when the MLC was implemented without shrinking the training area, almost all land covers were recognized as bamboo.

4. DISCUSSION

4.1 Low NDVI of Vegetation in April

In the training data, NDVI in April 2008 was less than that in February 2007 (Figure 5). NDVI for all land cover categories in April approached -0.4 despite radiometric calibration utilizing stable light source in AVNIR-2. This may be due to yellow sand because April is the month in which the most yellow sand was blown from China to Japan (Japan Meteorological Agency, 2008a). Therefore, sand in the upper atmosphere may have caused NDVI for all the categories to approach -0.4, which may be the value for the yellow sand, though yellow sand wasn’t observed at 1 April 2008 by the point which is on the ground and nearest to training areas (Japan Meteorological Agency, 2008b). This NDVI’s approach occurred in the training data may be responsible for low accuracy in classifiers that rely on seasonal differences in DN and NDVI values. However, classifiers in this study use only DN and NDVI’s hierarchical relationships among seasons and thus are not affected by the approach.
4.2 Importance of Shrinking in Setting Training Area

The MLC test implemented without shrinking recognized almost all land covers as bamboo, indicating the importance of shrinking the training area an adequate length when utilizing images which have a large positional gap.

4.3 Importance of Separating Setting Training Area from Accuracy Check Area

Although the Kappa coefficient for bamboo (DTC: 0.81, MLC: 0.97) and overall (DTC: 0.97, MLC: 0.98) was very high when the accuracy check area was set in training area, the Kappa coefficient for bamboo in the classifications were low (Table 2,

<table>
<thead>
<tr>
<th>Classified land cover</th>
<th>Name and area of landcover in vegetation map (unit : ha)</th>
<th>Total area (ha)</th>
<th>User's Accuracy (%)</th>
<th>Kappa coefficient (Cohen)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bamboo</td>
<td>Bambo 16.51 Evergreen 4.13 Deciduous 7.84 Grass &amp; paddy field 1.24 Dry field 0.53 Bare land 0.12 Water 0.00 Urban 0.84</td>
<td>30.80</td>
<td>0.53</td>
<td>0.67</td>
</tr>
<tr>
<td>Evergreen</td>
<td>Bamboo 5.14 Evergreen 12.74 Deciduous 6.36 Grass &amp; paddy field 0.17 Dry field 0.17 Bare land 0.02 Water 0.00 Urban 0.19</td>
<td>25.01</td>
<td>0.51</td>
<td>0.47</td>
</tr>
<tr>
<td>Deciduous</td>
<td>Bamboo 2.72 Evergreen 3.16 Deciduous 31.53 Grass &amp; paddy field 4.84 Dry field 0.29 Bare land 0.21 Water 0.00 Urban 0.34</td>
<td>35.32</td>
<td>0.57</td>
<td>0.41</td>
</tr>
<tr>
<td>Grass &amp; paddy field</td>
<td>Bamboo 2.33 Evergreen 1.15 Deciduous 4.69 Grass &amp; paddy field 49.74 Dry field 5.13 Bare land 1.60 Water 0.03 Urban 3.49</td>
<td>75.57</td>
<td>0.64</td>
<td>0.35</td>
</tr>
<tr>
<td>Dry field</td>
<td>Bamboo 0.86 Evergreen 1.18 Deciduous 5.56 Grass &amp; paddy field 10.28 Dry field 16.49 Bare land 1.00 Water 0.11 Urban 5.17</td>
<td>66.64</td>
<td>0.47</td>
<td>0.35</td>
</tr>
<tr>
<td>Bare land</td>
<td>Bamboo 0.00 Evergreen 0.00 Deciduous 0.00 Grass &amp; paddy field 0.00 Dry field 0.00 Bare land 0.00 Water 0.00 Urban 0.00</td>
<td>30.87</td>
<td>0.67</td>
<td>0.35</td>
</tr>
<tr>
<td>Water</td>
<td>Bamboo 0.00 Evergreen 0.00 Deciduous 0.00 Grass &amp; paddy field 0.00 Dry field 0.00 Bare land 0.00 Water 0.00 Urban 0.00</td>
<td>24.92</td>
<td>0.57</td>
<td>0.35</td>
</tr>
<tr>
<td>Total area (ha)</td>
<td>Bamboo 30.87 Evergreen 24.92 Deciduous 61.55 Grass &amp; paddy field 74.59 Dry field 35.32 Bare land 5.24 Water 0.94 Urban 66.67</td>
<td>300.05</td>
<td>0.53</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Overall Accuracy: 0.611 Overall Kappa Coefficient (Cohen): 0.522

<table>
<thead>
<tr>
<th>Classified land cover</th>
<th>Name and area of landcover in vegetation map (unit : ha)</th>
<th>Total area (ha)</th>
<th>User's Accuracy (%)</th>
<th>Kappa coefficient (Cohen)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bamboo</td>
<td>Bambo 11.90 Evergreen 2.48 Deciduous 3.12 Grass &amp; paddy field 2.43 Dry field 0.12 Bare land 0.14 Water 0.05 Urban 0.49</td>
<td>25.01</td>
<td>0.53</td>
<td>0.67</td>
</tr>
<tr>
<td>Evergreen</td>
<td>Bamboo 8.43 Evergreen 14.22 Deciduous 11.59 Grass &amp; paddy field 1.19 Dry field 1.01 Bare land 0.12 Water 0.02 Urban 0.63</td>
<td>30.87</td>
<td>0.57</td>
<td>0.47</td>
</tr>
<tr>
<td>Deciduous</td>
<td>Bamboo 5.70 Evergreen 5.42 Deciduous 39.21 Grass &amp; paddy field 17.16 Dry field 3.18 Bare land 1.03 Water 3.66 Urban 1.90</td>
<td>24.92</td>
<td>0.57</td>
<td>0.47</td>
</tr>
<tr>
<td>Grass &amp; paddy field</td>
<td>Bamboo 0.87 Evergreen 0.16 Deciduous 1.55 Grass &amp; paddy field 36.45 Dry field 2.39 Bare land 0.35 Water 0.01 Urban 1.65</td>
<td>35.32</td>
<td>0.57</td>
<td>0.47</td>
</tr>
<tr>
<td>Dry field</td>
<td>Bamboo 3.03 Evergreen 1.78 Deciduous 4.62 Grass &amp; paddy field 12.81 Dry field 23.89 Bare land 2.25 Water 0.39 Urban 13.66</td>
<td>43.43</td>
<td>0.57</td>
<td>0.47</td>
</tr>
<tr>
<td>Bare land</td>
<td>Bamboo 0.35 Evergreen 0.44 Deciduous 0.44 Grass &amp; paddy field 0.75 Dry field 2.55 Bare land 0.71 Water 0.00 Urban 9.95</td>
<td>19.05</td>
<td>0.57</td>
<td>0.47</td>
</tr>
<tr>
<td>Water</td>
<td>Bamboo 0.00 Evergreen 0.00 Deciduous 0.01 Grass &amp; paddy field 0.12 Dry field 0.40 Bare land 0.00 Water 0.25 Urban 0.16</td>
<td>15.20</td>
<td>0.57</td>
<td>0.47</td>
</tr>
<tr>
<td>Total area (ha)</td>
<td>Bamboo 30.80 Evergreen 25.01 Deciduous 61.55 Grass &amp; paddy field 74.59 Dry field 35.32 Bare land 5.22 Water 0.94 Urban 66.64</td>
<td>300.05</td>
<td>0.53</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Overall Accuracy: 0.543 Overall Kappa Coefficient (Cohen): 0.454
Table 3). This indicates the importance of the separating training area from the accuracy check area.

4.4 Extraction Bamboo

The results of this research showed that bamboo groves could be readily distinguished from most other land covers, but that some difficulty is experienced in distinguishing bamboo from other types of vegetation. The RGB and NDVI characteristics of bamboo are close to those of both deciduous and evergreen forest. In particular, willows and other deciduous trees on south facing slopes appear a similar color of light green to that of bamboo. In addition, both bamboo and evergreen forests are green in winter. Looking at the areas that were misclassified as bamboo, some evergreen forest in vegetation map were areas where bamboo was invading under a canopy formed of different trees. In the future, to improve the accuracy of this system, methods for obtaining higher quality training and image data, and for further orthorectifying the gaps between ground and sensed image positions, need be developed.

5. CONCLUSION

In this study, bamboo was extracted with 0.41-0.48 Kappa coefficient accuracy, with the classification for accuracy check implemented in an area completely separate from that used for setting training data. Although various technical issues need to be addressed to improve classification accuracy in future research, the results indicate that ALOS/AVNIR-2 and Google Earth images can be a useful tool for effective yet cost-efficient monitoring of Japan's expanding bamboo groves at the region or prefecture level.

ACKNOWLEDGMENTS

The authors would like to thank Kazuyuki Takahashi and Kohei Emura of Tokyo University of Information Sciences (TUIS) for kindly allowing us to use their vegetation data in the accuracy test. We also thank Kevin Short of TUIS for his help on the English version of this paper.

REFERENCES


The Geospatial Information Authority of Japan (GSI), 2007. Digital Map 2500 (Spatial Data Framework)

