Land-use classification of SPOT HRV data using a cover-frequency method

P. GONG† and P. J. HOWARTH
Earth-Observations Laboratory, Institute for Space and Terrestrial Science,
Department of Geography, University of Waterloo, Waterloo,
Ontario, N2L 3G1 Canada

(Received 13 August 1990; in final form 26 June 1991)

Abstract. A two-stage classification procedure has been applied to extract land
use in a rural-urban fringe environment from SPOT High Resolution Visible
(HRV) multi-spectral data. In this procedure, the SPOT HRV data were first
classified into twelve land-cover types using a supervised maximum-likelihood
classification (MLC). In the second stage, cover frequencies were extracted by
moving a pixel window over the land-cover map obtained at the first stage. These
cover frequencies were then employed in the classification of 14 land-use classes
using a supervised minimum-city-block classifier. Results obtained with the
cover-frequency method have been compared with those obtained using the
conventional MLC approach. The overall accuracy measured by the Kappa
coefficient was 0.462 for the MLC method; it was significantly improved to 0.663
with the cover-frequency method.

1. Introduction

The problem in computer-assisted land-use mapping is that land use is a cultural
concept which is conceptually different from land cover. Land cover is the physical
evidence on the surface of the Earth (Lillesand and Kiefer 1987); land use is defined
as man’s activities on land which are directly related to land (Clawson and Stewart
1965). What we see on remote sensing imagery is only the physical evidence of land
use in various land-cover types (Driscoll 1985).

It is relatively easy to map land cover with per-pixel classification techniques
because land cover is directly related to the pixel values on an image. Accurate land-
use maps, however, cannot be obtained through a direct transformation from
remotely-sensed data to land-use categories because they require information from
both spectral and spatial contexts to characterize the land use. Per-pixel classifica-
tion techniques do not have such capabilities. This is particularly true when such
techniques are applied to higher spatial resolution satellite data such as Landsat
Thematic Mapper (TM) and SPOT High Resolution Visible (HRV) data (e.g.,

Cover-frequency methods have been developed specifically to derive land-use
information from high spatial resolution data (Wharton 1982, Zhang et al. 1988). In
this procedure, a land-cover map is first derived from remote sensing images using a
clustering method. Cover-frequency vectors are then extracted from each pixel of a

†Present address: Department of Surveying Engineering, The University of Calgary, 2500
University Drive N.W., Calgary, Alberta, T2N 1N4 Canada
land-cover map by passing a pixel window with a specific size over the image. The cover-frequency vectors can be used as new features in the classification of land uses.

Wharton (1983) employed a multi-dimensional histogram clustering procedure (HICAP) to find multi-dimensional peaks of cover-frequency vectors. Subsequently, by associating each cover-frequency vector peak with a land-use class, all the cover-frequency vectors are classified to produce the land-use classification. This method was tested with 7.5 m spatial resolution airborne multi-spectral images obtained over a suburban area in Maryland (Wharton 1982). It was reported that by using HICAP the overall classification accuracy was improved from 55.0 per cent to 75.8 per cent when compared with a per-pixel spectral classifier in the classification of five land-use classes (new single-unit dwelling, old single-unit dwelling, apartment/townhouse, trailer court, and industrial/commercial).

Zhang et al. (1988) classified five land-use classes (forest, paddy, urban, water, and other) from Landsat-TM images. They first derived a land-cover map with 13 land-cover classes using a clustering method. Supervised training was then applied to the land-cover map to define the five land-use types. A minimum-city-block classifier was applied to classify cover-frequency vectors into land-use types. As a result, an overall classification accuracy improvement of 7 per cent, over the figure of 61.5 per cent obtained from classifying these images directly, was achieved using the cover-frequency approach.

In this paper, the application of the cover-frequency method is reported in the identification of 14 land-use classes from SPOT HRV multi-spectral (XS) data acquired for a rural-urban fringe area. The SPOT XS data were first classified into 12 land-cover types using a supervised maximum-likelihood classification (MLC) method. Thus a land-cover map was obtained. Land-use maps were then generated from the land-cover map using a supervised minimum-city-block distance classifier.

2. Study site and data sources
The area selected for study is the Town of Markham (43°52′N; 79°16′W) which is situated on the rural-urban fringe of north-eastern Toronto. Land cover and land use in this area are typical of the rural-urban fringe of many cities in North America with agricultural and natural land being converted to primarily residential, industrial and commercial uses. The area has been used as a test site for several remote sensing studies of land-cover and land-use classification (Johnson and Howarth 1987, Martin et al. 1988, Gong and Howarth 1989, Gong and Howarth 1990a, b, c).

Data selected for this study consisted of SPOT HRV imagery recorded on 4 June 1987. A 512 by 512 pixel subscene (approximately 10 km by 10 km) of SPOT XS data was chosen for the analysis. Figure 1 is a standard colour composite showing part of the study area (8 km by 10 km) with the built-up area in the middle and lower portions of the figure. In addition, 1:8000 scale black-and-white aerial photographs, acquired less than two months before the SPOT imagery was recorded, were available to assist in land-cover and land-use identification.

3. Cover-frequency method
The cover-frequency method for land-use classification is supported by the physical relationships between the spectral classes (or land-cover classes) and the land-use classes to be identified. Many land-use classes in the rural-urban fringe (e.g., residential and industrial/commercial) consist of a number of spectrally
Figure 1. The standard colour composite of the SPOT multi-spectral images showing part of the study area (approximately 8 km by 10 km).

Figure 2. Land-use classification results obtained using the maximum-likelihood classification.

Figure 3. Land-use classification results obtained using the cover-frequency method with a pixel-window size of 9 by 9.
Table 1. Land-cover types used in the cover-frequency method.

<table>
<thead>
<tr>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residential roof</td>
</tr>
<tr>
<td>Road surface</td>
</tr>
<tr>
<td>Industrial and commercial roof</td>
</tr>
<tr>
<td>Cleared land</td>
</tr>
<tr>
<td>Lawn and tree complex</td>
</tr>
<tr>
<td>Cultivated grass</td>
</tr>
<tr>
<td>Deciduous trees</td>
</tr>
<tr>
<td>Coniferous trees</td>
</tr>
<tr>
<td>Crop cover</td>
</tr>
<tr>
<td>New crops and pasture</td>
</tr>
<tr>
<td>Fallow land</td>
</tr>
<tr>
<td>Water surface</td>
</tr>
</tbody>
</table>

dissimilar land-cover types such as roof, pavement, trees and lawn. The major differences among these land-use classes are the varying proportions of the constituent land-cover types.

When the use of high spatial resolution imagery makes individual land covers separable, what we see on the image is primarily combinations of patterns of these land-cover types. One can therefore count the frequencies of all land-cover types within specific areas and use these frequencies as the spatial signatures for land-use classification. For example, many single-unit residential areas in North America contain large areas of lawns and trees while roofs and pavements occupy little space; in contrast, industrial/commercial sites are often characterized by large buildings and parking lots, and therefore contain more roofs and pavements and fewer trees and lawns. Thus, by comparing the frequencies of the four land covers, it should be possible to distinguish the two types of land use.

3.1. Land-cover map generation

Both Wharton (1982) and Zhang et al. (1988) suggested that the land-cover map be generated using unsupervised classification techniques. In this study, however, a supervised MLC method was used to generate the land-cover map. There are some advantages to this approach. First, it is relatively easy to implement because the supervised approach avoids some difficulties such as assigning labels to ambiguous groupings when using clustering methods. Secondly, a supervised approach allows one to specify the land covers more explicitly. Thus, the various land-cover components can be more physically related to the land-use classes which are to be identified.

Twelve major land-cover types can be defined for the study area (table 1). They have been used in a number of land-cover classification studies of the same area (Gong and Howarth 1989, 1990 b, c). These land covers were defined as pure spectral classes so that the training statistics for each land-cover type would not violate the normal distribution requirement of the MLC method. Most of the land-cover classes are self-explanatory. It should be noted, however, that the lawn and tree complex occurs within urban areas, while the cultivated grass primarily forms the fairways on golf courses. The distinction between crop cover, and new crops and pasture is easily made on the basis of high and low reflectances, respectively, in the infrared band.

In contrast to the traditional block-training procedure in which a number of contiguous pixels are used as the training sample, the single-pixel sampling strategy
was employed for training, as described in Gong and Howarth (1990b). After training statistics for each land-cover type were obtained, the SPOT XS data were classified using the MLC method. The resultant land-cover map was then used in the subsequent analysis.

3.2. Land-use classification scheme

Table 2 lists the 14 land-use classes used in this study. These classes were determined based on both the needs of the local planning agencies and the discriminating capability of the SPOT XS data. It has to be noted that some of the land-use categories in table 2 are similar to the land-cover types, according to the definitions of land cover and land use stated in §1. These categories include idle land, new crop and pasture, mature crop, parks, cleared land, deciduous trees, and water. They could have been excluded from the land-use classification process by setting a threshold for the remainder of the land-use classes. However, it was decided to retain them because they can be used to test the potentials of the cover-frequency classification algorithm for classifying land cover. In spite of the land-cover classes specified in table 2, the classification scheme is referred to as a land-use classification to differentiate it from a pure land-cover classification, such as the one listed in table 1.

Descriptions of the 14 land-use classes in table 2 are presented in table 3. These land-use classes can be divided into two groups:

1. Those which are dominated by one spectral class on the SPOT XS image.
2. Those which are composed of more than one spectral class.

Five of the 14 classes can be placed in the first group. These are mature crop, parks, cleared land, deciduous trees, and water. All the other classes belong to the second group.

3.3. Cover-frequency extraction and classification

For each pixel at the centre of a pixel window, a cover-frequency vector was extracted. This was done for every pixel by moving the pixel window over the land-

<table>
<thead>
<tr>
<th>Code</th>
<th>Land-use classes</th>
<th>Colour</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Old urban residential</td>
<td>Red</td>
</tr>
<tr>
<td>2</td>
<td>New urban residential</td>
<td>Green</td>
</tr>
<tr>
<td>3</td>
<td>Rural residential</td>
<td>Blue</td>
</tr>
<tr>
<td>4</td>
<td>Industrial/commercial/institutional</td>
<td>Yellow</td>
</tr>
<tr>
<td>5</td>
<td>Idle land</td>
<td>Pink</td>
</tr>
<tr>
<td>6</td>
<td>New crop and pasture</td>
<td>Turquoise</td>
</tr>
<tr>
<td>7</td>
<td>Mature crop</td>
<td>Orange</td>
</tr>
<tr>
<td>8</td>
<td>Golf course</td>
<td>Light grey</td>
</tr>
<tr>
<td>9</td>
<td>Parks</td>
<td>Dark green</td>
</tr>
<tr>
<td>10</td>
<td>Cleared land</td>
<td>Dark blue</td>
</tr>
<tr>
<td>11</td>
<td>Land under construction</td>
<td>Purple</td>
</tr>
<tr>
<td>12</td>
<td>Deciduous trees</td>
<td>Light blue</td>
</tr>
<tr>
<td>13</td>
<td>Hazard land</td>
<td>Dark red</td>
</tr>
<tr>
<td>14</td>
<td>Water</td>
<td>Pale green</td>
</tr>
</tbody>
</table>
Table 3. Spatial-spectral characteristics of land-use classes in Scheme 1 when observed on SPOT XS image.

<table>
<thead>
<tr>
<th>Code</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Well landscaped residential areas where trees, lawns, driveways, and roof tops dominate</td>
</tr>
<tr>
<td>2</td>
<td>Fewer trees when compared to old residential areas, a very regular pattern</td>
</tr>
<tr>
<td>3</td>
<td>A low density of roof tops which are surrounded with vegetation</td>
</tr>
<tr>
<td>4</td>
<td>Large building roof tops and little vegetation</td>
</tr>
<tr>
<td>5</td>
<td>No vigorous growth of vegetation</td>
</tr>
<tr>
<td>6</td>
<td>Fields where vegetation does not fully cover the soil, or where the surface shows moderate vegetative growth</td>
</tr>
<tr>
<td>7</td>
<td>Fields in which high density vegetation is growing providing high spectral reflectance in the infrared band</td>
</tr>
<tr>
<td>8</td>
<td>Dominated by three types of land covers: well maintained grass, normal grass, and trees, and the spectral reflectance of the well maintained grass is very high in the infrared band</td>
</tr>
<tr>
<td>9</td>
<td>Large areas of grassland in the urban area</td>
</tr>
<tr>
<td>10</td>
<td>Denuded of vegetation and top soil showing evenly high reflectance in every band (Martin et al. 1988)</td>
</tr>
<tr>
<td>11</td>
<td>Lands where on which construction is underway, and varied reflectance associated with building foundations and superstructures, construction materials, and partially installed roads (Martin et al. 1988)</td>
</tr>
<tr>
<td>12</td>
<td>A few patches of forest land dominated by deciduous trees, other trees such as those scattered through old urban residential area, parks, and valleys do not belong to this class</td>
</tr>
<tr>
<td>13</td>
<td>Valley land which is composed of rivers or streams, wet grass, and trees</td>
</tr>
<tr>
<td>14</td>
<td>Several relatively large water surfaces such as reservoirs, and ponds</td>
</tr>
</tbody>
</table>

cover map. To obtain the cover-frequency vector, the number (or frequency) of pixels having a specific land-cover code can be determined. After the frequency of pixels within the pixel window was counted for each of the 12 land-cover codes, 12 frequencies were recorded to form the cover-frequency vector for identifying the land use to be assigned to the central pixel. Detailed descriptions of the concept of cover-frequency extraction are to be found in Wharton (1982) and Zhang et al. (1988).

As supervised training had been adopted for the land-use classification, it was decided to use a traditional block-training strategy. The advantage of this type of training is the ease with which one can specify training areas. By so doing, the image analyst can also implicitly specify the spatial structure for a class. For each land-use class, an average cover-frequency vector was calculated from the land-cover map by using the land-use training pixels to characterize the class.

When average cover-frequency vectors for all land-use classes had been obtained, the entire land-cover map was classified by comparing the city-block distances (Gonzalez and Wintz 1987) between the cover-frequency vector for each pixel and the average cover-frequency vectors for all 14 land-use classes. A pixel was classified into the land-use class for which the average cover-frequency vector had the shortest city-block distance to the cover-frequency vector of the pixel. The city-block distance $d(X,M_i)$ between a cover frequency vector $X=(x_1,x_2,\ldots,x_{12})^T$ and the average cover-frequency vector for class $i$, $M_i=(m_{i1},m_{i2},\ldots,m_{i12})^T$ is calculated from:
\[ d(X, M_i) = \sum_{j=1}^{12} |x_j - m_{ij}| \]

In the cover-frequency method, the size of pixel window used is an important factor. If the pixel window is too small, insufficient spatial information will be included to characterize a specific land-use class. If the pixel window size is too large, it will result in too much spatial information from other land-use classes being included. To examine the effect of pixel window size, 10 pixel sizes ranging from 3 by 3 to 21 by 21 with lateral increments of two pixels were used in this study. For each pixel window size, a land-use classification map was produced to give a total of 10 maps.

To provide a basis for comparison, the SPOT XS data were classified directly using the supervised MLC method. Accuracies of land-use classification derived from the MLC method were then compared with the accuracies obtained from the cover-frequency method.

3.4. Accuracy assessment

Accuracy assessment was done through a comparison of test pixels. Simple pixels were selected using the stratified systematic unaligned sampling strategy (SUSS) (Jensen, 1983) as a guideline such that one random sample pixel was picked from every 16 by 16 block on the image. Consequently, a total of 1024 sample pixels was selected. Since SUSS allocates sample pixels according to the area of each class, classes with small areas will receive too few sample pixels while classes with large areas will have too many samples. In order to maintain a similar confidence level for the accuracy estimate for each class, it was decided to select approximately 30 sample pixels for each class. Therefore, instead of using all the sample pixels allocated by SUSS, only part of them were used. For classes too small for the SUSS to allocate the desired number of test pixels, additional pixels were selected arbitrarily. The land class for each test pixel was identified so that it could be used as the reference data in the accuracy assessment. The identification was aided by using the 1:8000 scale aerial photographs.

For each land-use map, a confusion matrix was produced by comparing the classification results for the test samples with the reference data. The Kappa coefficient, \( \kappa \) (Cohen 1960), and its estimated variance, \( \hat{\kappa} \) (Fleiss et al. 1969), were calculated for each confusion matrix to evaluate the agreement between the classification results and the reference data. The Kappa coefficient was used as an index of accuracy for each classification. It has been recommended as a suitable accuracy measure in thematic classification for representing the whole confusion matrix (Congalton and Mead 1983, Rosenfield and Fitzpatrick-Lins 1986, Fung and LeDrew 1988). This is because it takes all the elements in the confusion matrix into consideration, rather than just the diagonal elements which occurs with calculation of overall classification accuracy. The variance was used when the significance tests were undertaken.

To determine the significance level of the difference between a classification result obtained with the cover-frequency method and one produced with the MLC method, the difference between the two Kappa values from the two classifications was first derived. The square-root of the sum of the variances between the two classifications was then calculated. The ratio, determined by dividing the difference by the square-root, was used as an index for the significance tests (Cohen 1960). A
ratio over 2.58 indicates that a difference is significant at the 0.99 probability confidence level.

In order to determine classification accuracies on a class-by-class basis, the conditional Kappa coefficient (Bishop et al. 1975) was used as a class accuracy index. It can be derived from a confusion matrix.

4. Results

Figure 2 shows the classification results obtained using the MLC method. It can be seen that all the urban land-use classes are intermixed. In agricultural areas, field boundaries have been classified as urban land use by the MLC. As a result, the entire map looks fragmented. The greatest confusion is between golf course (light grey) and parks (dark green). While golf course has been partly allocated to parks and to mature crop (orange), because of their similar spectral signatures, parks has been omitted almost entirely by the MLC. Rural residential (blue) and new crop and pasture (turquoise) have also been partly allocated to the parks class by the MLC. Hazard land (dark red) has been identified as golf course. Other confusions that occur are as follows: mature crops and deciduous trees (light blue), industrial land (yellow) and idle land (pink), rural residential and new crop and pasture, and cleared land (dark blue) and land under construction (purple).

Table 4 shows the confusion matrix, Kappa value, and the estimated variance of the Kappa value for the classification results obtained using the MLC method. The row entries of the confusion matrix represent the reference data and the column entries represent the classified results.

Table 5 displays the Kappa values of the classification results obtained using the cover-frequency method. A comparison of tables 4 and 5 shows that Kappa values obtained from the cover-frequency method with all 10 pixel window sizes are higher than the classification results obtained using the MLC method. The symbol * indicates that the improvement in classification accuracy with the cover-frequency method at a particular pixel window size was significant at the 0.99 probability confidence level when compared with the Kappa value derived from the MLC method.

The best classification accuracy was obtained using a pixel window size of 9 by 9. The Kappa value obtained with this pixel window was 0.663 which improved the Kappa value derived with the MLC method by 0.201. Figure 3 shows the classification results obtained using the cover-frequency method with a pixel window size of 9 by 9. Land-use classes look very homogeneous and the map appears more like a product of manual interpretation. The differences between this map and the one obtained by the MLC (figure 2) are readily apparent. The ‘pepper and salt’ effect observed in figure 2 has been reduced dramatically. Confusion between rural and urban land-use classes has also been reduced. Golf course was classified with a few pixels being assigned to other vegetation classes. Table 6 displays the confusion matrix. The row and column entries are the same as in table 3.

Figure 4 shows the conditional Kappa values obtained from the cover-frequency method and from the MLC method. Conditional Kappa values for each class are plotted against pixel window size. This allows a conditional Kappa value curve to be produced for each class. The resultant 14 curves have been grouped and displayed as three graphs: figure 4(a) shows the three residential classes and the industrial/commercial/institutional class; figure 4(b) shows curves for the three agricultural classes, deciduous trees, hazard land, and water; and figure 4(c) shows curves for the
two recreational land-use classes and two transitional land uses. Each curve starts with the conditional Kappa value of the particular class obtained from the MLC method. This serves as a reference for comparison and is plotted against the pixel window size of 1 by 1.

### 5. Discussion

By comparing tables 4 and 6, it can be seen that the agreements between classified results and reference data have been improved for most of the land-use classes by use of the cover-frequency method. This is particularly true for classes such as rural residential (3), idle land (5), new crop and pasture (6), and parks (9). However, the agreements for mature crop (7) and golf course (8) have been slightly decreased with the cover-frequency method. This indicates that the pixel window size of 9 by 9 is not appropriate for these two classes.

More interestingly, the commission errors for classes such as rural residential, new crop and pasture, and hazard land (13) have been increased by use of the cover-frequency method (table 6), even though the agreements between the classified

Table 5. Kappa values and their variances for the classification results obtained using the cover-frequency method.

<table>
<thead>
<tr>
<th>Pixel window size</th>
<th>3 x 3</th>
<th>5 x 5</th>
<th>7 x 7</th>
<th>9 x 9</th>
<th>11 x 11</th>
<th>13 x 13</th>
<th>15 x 15</th>
<th>17 x 17</th>
<th>19 x 19</th>
<th>21 x 21</th>
</tr>
</thead>
<tbody>
<tr>
<td>K (x 100)</td>
<td>60.3*</td>
<td>63.4*</td>
<td>64.9*</td>
<td>66.3*</td>
<td>61.0*</td>
<td>58.1*</td>
<td>57.5*</td>
<td>55.9</td>
<td>55.5</td>
<td>53.9</td>
</tr>
<tr>
<td>V (x 1000)</td>
<td>0.711</td>
<td>0.683</td>
<td>0.663</td>
<td>0.647</td>
<td>0.688</td>
<td>0.705</td>
<td>0.704</td>
<td>0.711</td>
<td>0.713</td>
<td>0.717</td>
</tr>
</tbody>
</table>

*Significant at the 0.99 probability confidence level.
both groups of land-use classes, the cover-frequency method can further improve classification accuracies obtained using the MLC method. However, as can be seen from figure 4, the improvement for spatially heterogeneous classes was generally greater than for spatially homogeneous classes. In addition, for spatially homogeneous classes, a small pixel window size was required while for spatially heterogeneous classes a relatively large pixel window size was necessary.

6. Conclusions
From the above analyses, the cover-frequency method used in this study proved to be a superior procedure for land-use classification when compared with the
conventional maximum-likelihood classification (MLC). Land-use classification accuracies obtained with the MLC method can be improved significantly by using the cover-frequency method. The major drawback of the pixel-window-based cover-frequency method is the pixel-window effect. Alternative methods need to be developed if further improvements in accuracy are to be expected.

The pixel window size used in the cover-frequency approach plays an important role in improving land-use classification accuracies. In the rural-urban fringe, such as the area selected for this study, land-use classes can be divided into two groups, spatially homogeneous classes and spatially heterogeneous classes. For the first group, the accuracy improvements obtained with the cover-frequency method, when compared with the classification accuracies obtained with the MLC method, tend to be relatively low. Optimal pixel window sizes for these classes are likely to be small. For the second group, the accuracy improvement can be relatively high, but larger pixel window sizes are needed. Given the variability in spatial extent of land uses in different environments, it would be beneficial to have a procedure for selecting the optimal pixel window size.

Acknowledgments
The authors gratefully acknowledge the assistance of SPOT Image Corporation of France and the Canada Centre for Remote Sensing in supplying the SPOT data used in this study as part of the Programme d’Évaluation Préluminaire SPOT (PEPS), Project No. 229. This research is funded by a Centre of Excellence grant from the Province of Ontario to the Institute for Space and Terrestrial Science and NSERC Operating Grant A0766 awarded to P. J. Howarth. P. Gong’s studies were supported by the International Development Research Centre (IDRC) of Ottawa.

References
Driscoll, E. C., Jr., 1985, Spatial classification of urban land uses from landcover probabilities derived through remote sensing. PhD Dissertation, Stanford University, Stanford, California, U.S.A.


