URBAN CLASSIFICATION OF HIGH RESOLUTION IKONOS IMAGERY USING TEXTURE

Shahid Kabir [1], Dong-Chen He [2], Patrice Rivard [3]

ABSTRACT

Traditional spectral-based methods of extracting urban land cover and land use information from remote sensing imagery have proven to be unsuitable for high spatial resolution images. Texture has been widely investigated as a supplement to spectral data for the analysis of complex urban scenes. This research evaluates the Grey Level Co-occurrence Matrix (GLCM) texture analysis technique and the Maximum Likelihood Classification approach for the extraction of texture features to be combined with spectral data, as a method for obtaining more accurate urban land cover and land use information from high spatial resolution images. Classifications were performed on IKONOS imagery using three datasets: a spatial dataset consisting of three texture images (mean, homogeneity and dissimilarity), a spectral dataset consisting of four spectral images (red, green, blue and N-IR), and a combination dataset (spatial and spectral). Results show that the combination dataset produced the highest overall classification accuracy of 86.1%, an improvement of 7.2% over the spectral dataset.

KEYWORDS

texture analysis, classification, high-resolution, satellite imagery, urban environment.

INTRODUCTION

As pressures increase for better land management, due to conflicting land use demands, high-resolution satellite imagery are providing more detailed urban land cover and land use data. As spatial and spectral resolutions of the remote sensor systems increase, however, image-processing algorithms have to be developed in order to determine how to exploit the raising volume of data efficiently in order to extract the desired information.

With the launch of the IKONOS satellites, images with higher spatial resolutions have contributed to improved land assessments. In the field of urban planning, these images are becoming a real alternative to aerial photography. However, traditional techniques used in multispectral classifications are not suitable for higher-resolution imagery. These methods are based on the spectral signatures present in the image, which is adequate for the classification of spectrally homogeneous object classes. Urban scenes, however, are much

[1] Researcher, Groupe de Recherche sur l'Auscultation et l'Instrumentation (GRAI), Univ. of Sherbrooke, Québec, J1K 2R1, Tel 819/821-8000, FAX 819/821-7974, Shahid.Kabir@Usherbrooke.ca

[2] Professor, Centre d'applications et de recherches en télédétection (CARTEL), Univ. of Sherbrooke, Québec, Canada J1K 2R1, Tel 819/821-8000, FAX 819/821-7944, Dong-Chen.He@Usherbrooke.ca

[3] Professor, Civil Engrg. Department, Univ. of Sherbrooke, Québec, Canada J1K 2R1, Tel 819/821-8000, FAX 819/821-7974, Patrice.Rivard@Usherbrooke.ca

more complex; for example, residential areas are typically seen from above as a mixture of tree crowns, rooftops, lawns, paved streets, driveways and parking lots. It is the composite of these features, rather than an inventory of the individual components, that is often of interest. Consequently, for applications involving the mapping of heterogeneous features in complex urban scenes, results obtained from such methods are unsatisfactory. This is mainly because the contribution of spectral information is limited since urban objects are distinguished better through their spatial properties rather than their spectral properties (Kiema 2002).

A significant drawback of these conventional classification approaches is that while the information content of the imagery increases with spatial resolution, the accuracy of the land use classification may decrease (Shaban and Dikshit 2001). This is due to a higher number of detectable sub-class elements resulting in increasing spectral variability within the classes, inherent in more detailed, higher spatial resolution data. Another obstacle is that landscapes are composed of natural and artificial materials that sometimes present close or even identical spectral properties, which can introduce confusion between classes. This confusion can also be caused by the fact that groups of pixels representing the same land cover type will not necessarily have the same spectral information due to noise in the data, atmospheric effects, and natural variation within the land cover type (Smith and Fuller 2001).

Texture has been widely investigated as a possible source of unique information to supplement spectral data. Both aerial photograph interpreters and digital image analysts have long since recognized image texture as a powerful source of information in remote sensing analysis (Moskal and Franklin 2001). One approach to overcome the obstacles of spectral classification of satellite imagery is to integrate textural data into the classification process. However, the efficiency of texture when applied to high-resolution images of complex urban scenes has yet to be determined. This research study proposes the extraction of texture information from high-resolution IKONOS imagery through the grey level co-occurrence matrix texture analysis method, and classification using the texture information combined with the spectral data through the maximum likelihood classification technique, in order to provide a more precise classification of complex urban objects from the high-resolution imagery.

RESEARCH DATA

High spatial resolution images of the city of Sherbrooke, Quebec, Canada, located between 45º18′ and 45º27′ latitude north, and 71º48′ and 72º02′ longitude west, were acquired on May 20th, 2001 at 10:50 am local time by the IKONOS-2 satellite of Space Imaging. This study site was selected for its various types of land use and land cover, which provide a good region for the purposes of urban classification analysis. The scenes have an image dimension of approximately 12x13 km and consist of four multispectral 4x4 meter bands in red, green, blue, and N-IR, and one panchromatic 1x1 meter band. Supplementary data used in this study were obtained from the National Topographic Data Bank of the Sherbrooke region and a topographic map of the Sherbrooke area, both at a scale of 1:50 000, produced in 2000 by Natural Resources Canada, black and white aerial photographs of the area taken in September 1998 and August 2000, at a scale of 1:15 000 and 1:40 000 respectively, obtained from the Québec Ministry of Natural Resources, as well as data collected during field visits.

METHODS AND RESULTS

In this study, high-resolution IKONOS imagery was processed using the grey level co-occurrence matrix textural analysis method and then classified through the maximum likelihood classification technique. Three datasets were created in order to evaluate the effectiveness of the texture information on the classification of the complex urban objects in the high-resolution imagery: a spectral dataset consisting of the original red, green, blue and NIR images, a spatial dataset consisting of texture images and a combined dataset consisting of the multispectral and texture images.

Statistical texture methods analyze the spatial distribution of grey values by computing local features at each point in the image, and deriving a set of statistics from the distributions of the local features. Second-order statistics operate on a probability function that measures the probability of co-occurrence of two pixel grey values in the image, separated by a certain distance and direction; this function is also known as the grey level co-occurrence matrix (GLCM) (Haralick et al. 1973).

In order to derive the texture images, the variation coefficient for each object class was calculated as a function of the window size using the randomly chosen homogeneity texture feature. This revealed the most appropriate window size to be 11x11 pixels. The direction of 0º between pixels was used since it is the most common choice found in literature. For the pixel distance, small distances have been found to produce the best results (Karathanassi et al. 2000); as such, a distance of 1 pixel was employed. Using the 11x11 window, the direction of 0º between pixels and a distance of 1 pixel, eight texture features were extracted from the original panchromatic image: contrast, correlation, dissimilarity, entropy, homogeneity, mean, second moment, and variance (Figure 1).

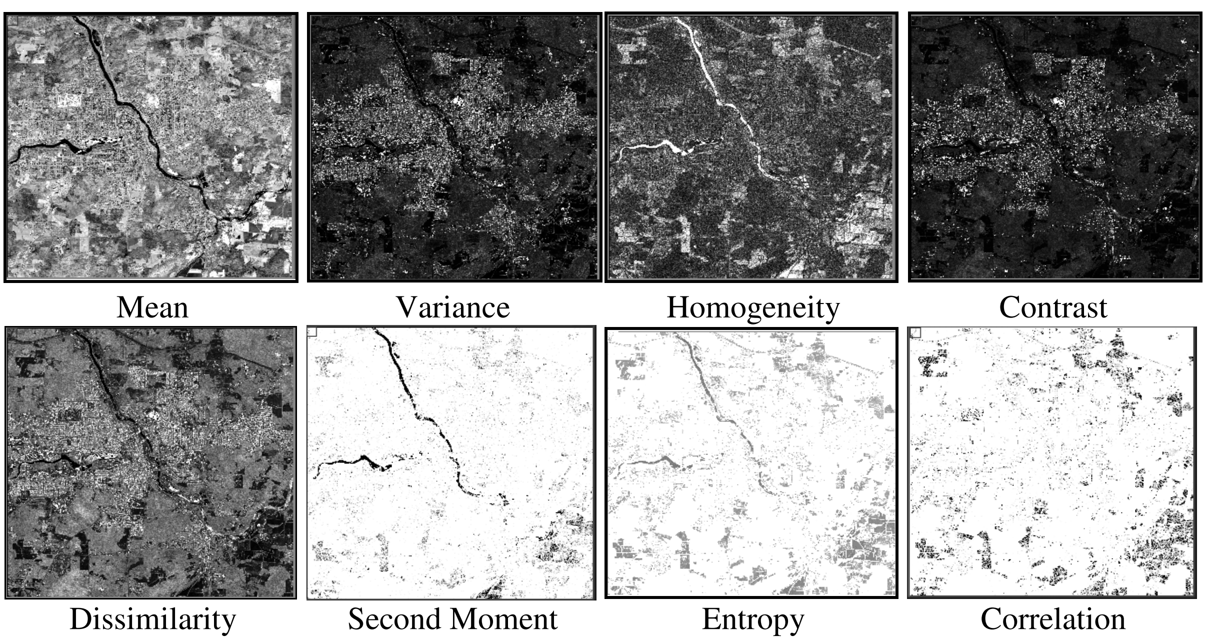

Figure 1: Texture Images of the Eight GLCM Features

Since many of these texture features contain the same information, an elimination process was employed in order to choose the most useful features for good urban class discrimination. For the first step in the process of elimination, the visual quality of these texture images was analysed and three features, correlation, entropy, and second moment, were initially considered for discarding; these features were found to be unsuitable for more heterogeneous classes. After displaying the histograms of the texture images, it was confirmed that the three features, correlation, entropy, and second moment, were to be eliminated due to the minimal amount of information they presented. The possible elimination of another two features, contrast and variance, was also considered from the histogram analysis because of the same reason. Finally, through calculation of the correlation matrix, it was confirmed that the two features, contrast and variance, as well as the first three features, correlation, entropy, and second moment, were to be discarded due to their relatively high correlation with the other features. As a result, only three texture features, dissimilarity, homogeneity, and mean, were selected for use.

Supervised classification techniques use suitable algorithms to label the pixels in an image as representing particular ground cover types, or classes (Richards and Jia 1999). They require a priori knowledge of the object classes in the image in order to create training sites, which are then used to "train" the system in order to generate the spectral signatures for these classes. The system thereafter labels all pixels belonging to each particular class according to a decision rule. The maximum likelihood classification (MLC) technique calculates the greatest probability that a pixel belongs to a given class, thus minimizing pixel misclassifications. In this study, training and verification sites for the following twelve urban land use and land cover classes were created: Agriculture, Asphalt and Parking Lot, Bare Soil, Commercial Area, Coniferous Forest, Deep Water, Deciduous Forest, Grass, Residential Area, Road Networks, Shallow Water, Shrubs. Each of the three datasets was then classified into these twelve object classes.

The final step of the classification is the evaluation of the accuracy of the results obtained, which indicates how well the classification performed. Once the spectral space is segmented into different regions associated with classes of objects, each pixel of the verification sites is assigned the label of the class that represents it in the segmented spectral space. The overall result of this process is presented in the form of a confusion matrix. From this matrix many classification precision indexes can be calculated. From a comparative study done on the different methods of evaluating the classification accuracy, it was found that the most appropriate index to provide classification precision is the Kappa coefficient, because it takes account of all the elements of the confusion matrix (Fung and Ledrew 1988).

CLASSIFICATION RESULTS

The results obtained from the classification stage of this research show that the classification done with the purely spatial dataset of mean, homogeneity and dissimilarity texture bands, produced limited accuracies ranging from 59.8% to 84.9% for all classes, with an overall accuracy of 73.5%. The best accuracies obtained for this dataset are for the Deep Water, Bare Soil, and Grass classes, which have 84.9%, 73.9% and 72.5% accuracies respectively. Classes that produced low accuracies are the Commercial class, Coniferous Forest, Asphalt and Parking Lot, Residential, and Shrubs classes, with 59.8%, 61.1%, 61.2%, 61.4% and 61.8% accuracies respectively.

The classification of the purely spectral dataset of red, green, blue and NIR bands produced somewhat higher accuracies for all of the classes compared to the spatial dataset. Here, the accuracies range from 62.4% to 87.5% for all classes, with an overall accuracy of 78.9%. This indicates an increase in accuracy ranging from 0.3% to 6.1% for each class and an overall increase of 5.4%. The highest classification accuracies obtained with this dataset was for the Deep Water 87.5%, and Grass 77.3% classes. The Asphalt and Parking Lot 64.3%, Coniferous Forest 62.4%, and Shrubs 62.4% classes once again produced the lowest accuracies.

The highest accuracies obtained in this study were with the classification of the combined spectral and spatial datasets, which produced accuracies ranging from 70.6% to 90.9% for all classes and an overall accuracy of 86.1%. The improvement in classification accuracies with this dataset over the spectral dataset ranges from 3.4% to 22.2% for each class with an overall increase of 7.2%. For this dataset also, the Deep Water and Grass classes once more have the highest classification accuracies at 90.9% and 89.0% respectively. The classes that saw the greatest improvement with the combined dataset are the Asphalt and Parking Lot class, and the Commercial class; due to the irregular patterns of these classes, as well as their high spectral variance, combination of the data improved their discrimination.

The classification accuracies obtained for each of the three datasets, as well as the overall accuracies and Kappa coefficients are presented in Table 1.

Table 1: Classification Accuracies

	Spectral Dataset	**Spatial Dataset**	**Combined Dataset**
Kappa Coefficient	0.74	0.68	0.83
Overall Accuracy (%)	78.9	73.5	86.1
Classes	**Classification Accuracies (%)**		
Agriculture Land	73.9	70.6	88.9
Asphalt and Parking Lot	64.3	61.2	86.5
Bare Soil	74.2	73.9	84.3
Commercial, Industrial & Institutional	68.5	59.8	83.1
Coniferous Forest	62.4	61.1	70.6
Deciduous Forest	73.9	67.8	82.7
Deep Water	87.5	84.9	90.9
Grass	77.3	72.5	89.0
Residential Area	65.8	61.4	82.8
Road Network	68.2	62.6	82.1
Shallow Water	74.7	70.7	80.9
Shrubs	62.4	61.8	71.9

Interpretation of Results

The combination dataset classification produced the highest increases in accuracy, the results of which are presented in the form of a thematic map in Figure 2. The Asphalt and Parking Lot class, as well as the Commercial class, showed the most increase in classification accuracy with the combination dataset. Other classes that also produced comparably high increases in accuracy are the Residential and Road Network classes. This is the expected performance of the input of textural data in the multispectral classification, since these classes obtained relatively poor accuracies with the spectral and textural datasets alone. The lowest increases in classification accuracy with the combined data were obtained by the classes that produced relatively high accuracies with the purely spectral and textural datasets. The Deep Water class saw an increase in accuracy of only 3.4% and the Shallow Water class only 6.2%. Since these classes are spatially and spectrally distinguishable anyhow, the addition of texture did not make much of a contribution. This indicates that the combination of textural and spectral information is needed for those classes that produce low accuracies with purely spectral or textural data.

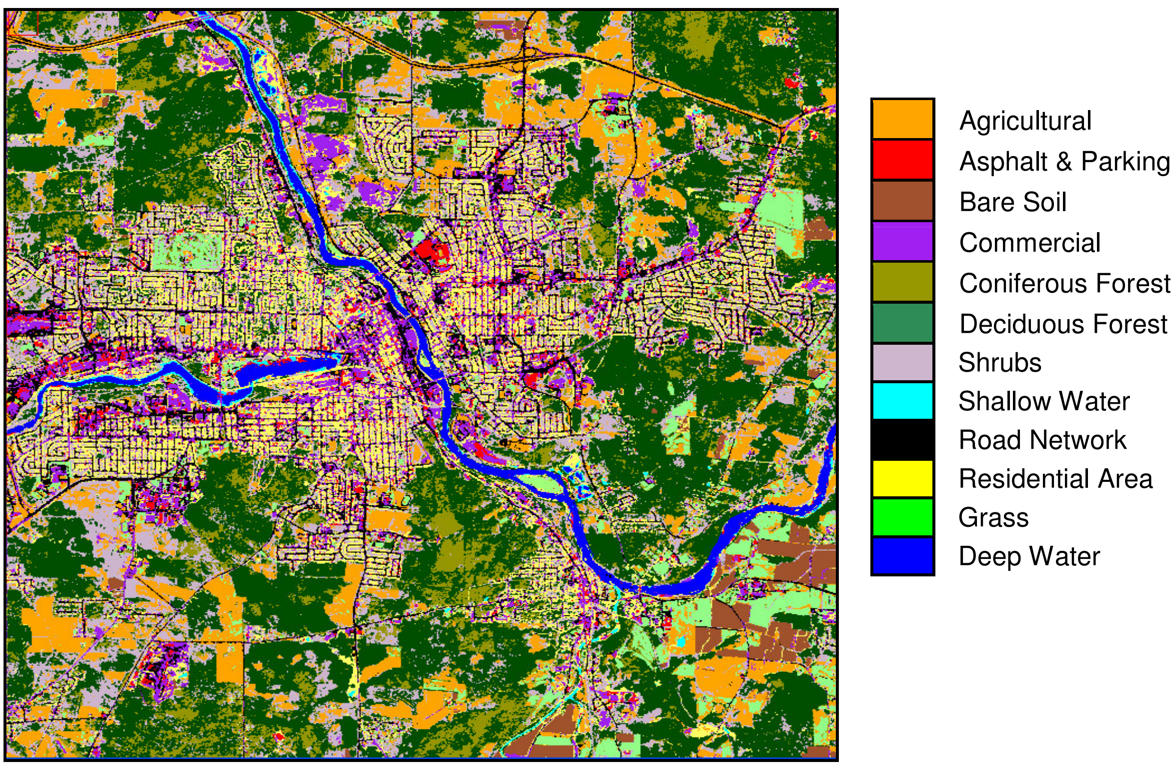

Figure 2: Classification of Combined Dataset

The lowest classification accuracies produced for the combination dataset was for the Coniferous Forest and Shrubs classes. These relatively low accuracies are reflected in the low percentages covered by these two classes in the image, where the Coniferous Forest class comprises only 7.5% and the Shrubs class 15.0%. The Deciduous Forest class, on the other hand, is shown to occupy more than 30% of the whole image. The Coniferous Forest class should actually make up almost half of the total of the two Forest classes. This indicates that many pixels belonging to the Coniferous Forest class were misclassified as Deciduous Forest, which may be due to the small window size employed. The combined dataset classifications of some specific classes are shown in Figure 3.

Residential area | Road network | Forest | Water network

Figure 3: Combined Dataset Classification of Certain Classes

The Deep Water, Bare Soil, and Grass classes obtained the highest accuracies in the spatial classification. The lowest classification accuracies obtained with this dataset are by the Commercial, Asphalt and Parking Lot, Residential, Coniferous Forest, and Shrubs classes. The textural heterogeneity of the Commercial class can be explained by the irregular structures of the buildings, as well as the presence of more than one building intermingled with parking areas, such as the case of colleges and universities. The Asphalt and Parking Lot class presents heterogeneous textures because of the presence of vehicles, which, especially in the case of parking lots, do not always have an even distribution. For the Residential class, the random mixture of roofs and treetops is likely the cause of the varying textures. As for the heterogeneity of the textures described by the Coniferous Forest and Shrubs classes, this may well be due to the fact that these two classes, as well as the Deciduous Forest class, do not occupy distinct areas of the image; most of the forests in the images are a composite of these three classes. The low classification accuracies of all these classes indicate that they need the input of spectral information for greater discrimination.

In the spectral classification, the classes that produced the highest accuracies are again the Deep Water and Grass classes, which mean that with either spatial or spectral information, these classes are highly discriminable. The classes that produced the lowest classification accuracies with the spectral data are the Asphalt and Parking Lot, Coniferous Forest, and Shrubs classes. This means that these classes are not easily distinguishable spectrally. The inability to produce representative spectral signatures for these classes may be due to various reasons. In the case of the Asphalt and Parking Lot class, this is most likely due to the presence of vehicles, which produce spurious diffuse and specular reflections that degrade the spectral signature of the pixels in this class. The fact that the forests in the image are generally mixed is probably the reason that the Coniferous Forest and Shrubs classes failed to produce representative spectral signatures. Since these classes also produced low accuracies with the spatial dataset, this means that they are not distinguishable with only spectral or textural data alone.

In Figure 4, some examples taken from the classified image of the combined dataset are presented. Figure 4(a) shows an irrigation pond on an agricultural plot taken from the original panchromatic image and Figure 4(b) is the corresponding section from the classified image. Also taken from the panchromatic image, Figure 4(c) shows reserved water in a gravel production company at the corner of Bel-Horizon and Dunant streets, and Figure 4(d) is the classified section that corresponds to it. Both of these examples demonstrate the Shallow Water classification. Figure 4(e) presents a segment of the Magog River and Figure 4(f) shows part of the Saint François River near Sherbrooke North; these are examples of the Shallow Water and Deep Water classes. Figure 4(g) shows the intersection of highways 10 and 216 and Figure 4(f) shows the Jacques-Cartier Bridge; these examples demonstrate the Road Network classification.

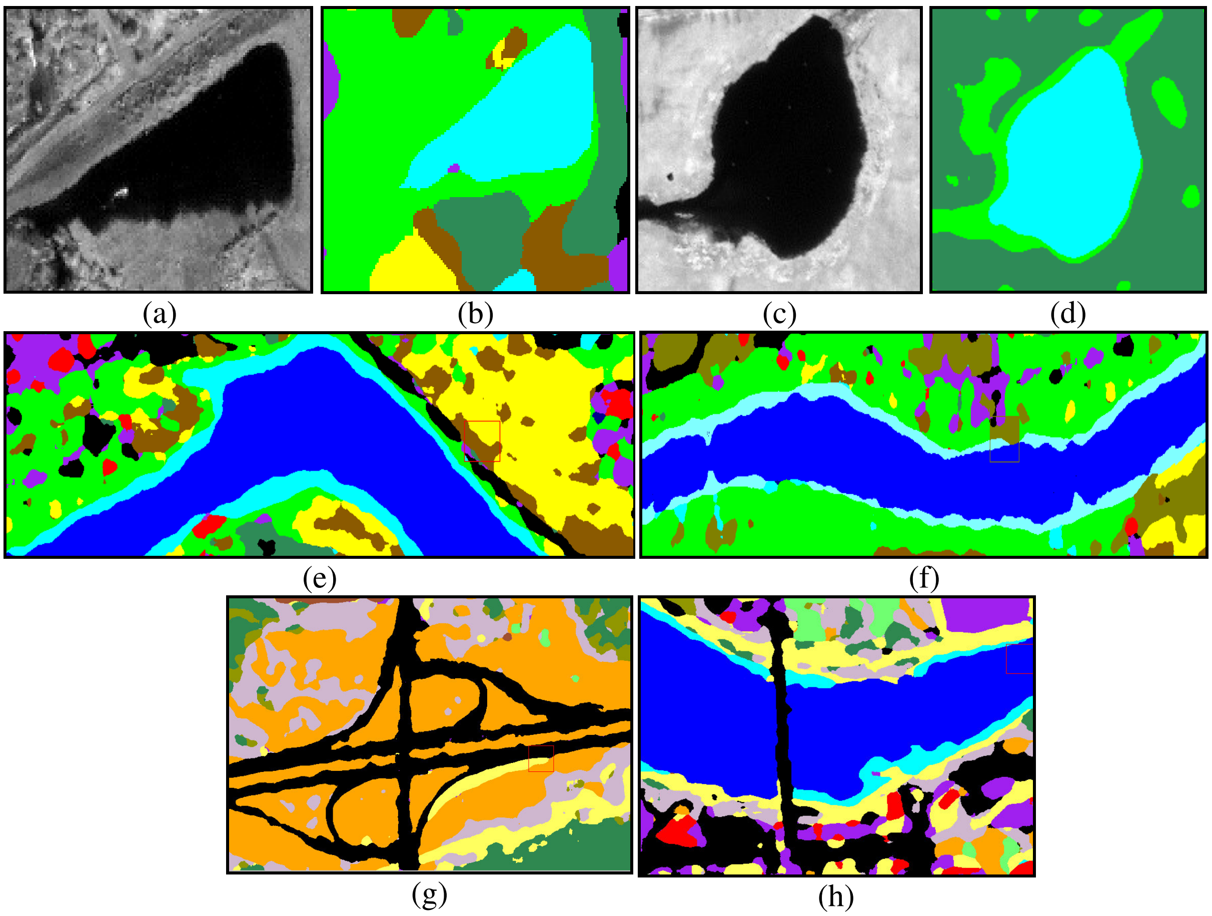

Figure 4: Classification Examples Taken from Classified Image

A statistical representation of the classification done with the combined data can be seen in Table 2. It shows how many pixels belong to each class and what percent of the image each class occupies.

Table 2: Statistics of the Combined Dataset Classification

Classes	Number of Pixels	Percentage (%) of Whole Image
Agricultural Land	12 292 332	10.16
Bare Soil	2 193 414	1.81
Commercial	8 756 036	7.24
Coniferous Forest	9 055 436	7.48
Deciduous Forest	36 916 561	30.51
Deep Water	675 924	0.56
Grass	4 315 099	3.57
Parking Lot	2 633 992	2.18
Residential Area	15 340 348	12.68
Road Network	9 581 520	7.92
Shallow Water	1 136 101	0.94
Shrubs	18 103 237	14.96
Total	121 000 000	100.00

CONCLUSION AND DISCUSSION

This study has produced results that show classifications based only on textural information provide lower accuracies than classifications performed with purely spectral data. The combination of both types of data for classification, however, produces the highest classification accuracies. These findings are supportive of the concept proposed in this study, that both texture channels and high spatial resolution imagery can provide improved spectral classification accuracies.

Overall, the results of this research work support previous studies (Moskal and Franklin 2001, Shaban and Dikshit 2001, and Kiema 2002) in respect to the improvement of spectral classifications through the addition of textural data, though they differ somewhat in areas that are directly related to the texture analysis stage, and mainly from previous research conducted with imagery of a lower spatial resolution. As such, the use of GLCM texture analysis on high spatial resolution IKONOS imagery, combined with the MLC approach, for the improvement of spectral classifications of urban land cover and land use classes provides some interesting results, such as better discrimination for classes that have high spatial and spectral variation, and not much improvement for classes that are already spectrally distinct, as well as the need for a different window size for large classes.

Future studies within the GLCM texture analysis approach can, therefore, focus on the use of different pixel distances and directions, and various window sizes in order to examine their relationship to different types of urban land cover and land use classes for the determination of their contribution to urban texture discrimination of high spatial resolution imagery. Another study also within the scope of texture analysis that may prove to be very interesting is the separate assessment of the most useful texture features to determine their role in the classification, which might provide some insight into which ground classes they complement the most, and their impact on urban classifications.

The application of GLCM texture analysis and multispectral MLC techniques for the classification of combined spatial and spectral data for the urban land use and land cover classification of high spatial resolution IKONOS imagery produced very promising results. Some problem areas were encountered, however, related to the limitations of this study. The texture analysis applied in this study was not comprehensive as it relied on the use of only one window size, which did not permit good textural discrimination of certain ground cover classes, and the use of only one direction and distance between pixels, the effects of which have not been determined. These aspects need to be further studied, based on smaller samples to avoid large computational costs, in order to optimize their application to high spatial resolution imagery.

ACKNOWLEDGEMENTS

This research was supported by a grant to Dr. Dong-Chen He from the Natural Science and Engineering Research Council of Canada.

REFERENCES

Fung, T., and Ledrew, E. (1988). “The Determination of Optimal Threshold Levels for Change Detection using Various Accuracy Indices.” *Photogrammetric Engineering and Remote Sensing*, 54 (10) 1449-1454.

Haralick, R.M., Shanmugan, K., and Dinstein, I. (1973). “Textural Features for Image Classification.” *IEEE, Transactions on Systems, Man, and Cybernetics*, (8) 610-621.

Karathanassi, V., Iossifidis, C.H., and Rokos, D. (2000). “A Texture-Based Classification Method for Classifying Built Areas According to their Density.” *International Journal of Remote Sensing*, 21 (9) 1807-1823.

Kiema, J.B.K. (2002). “Texture Analysis and Data Fusion in the Extraction of Topographic Objects from Satellite Imagery.” *International Journal of Remote Sensing*, 23 (4) 767-776.

Moskal, L.M., and Franklin, S.E. (2001). “Classifying Multilayer Forest Structure and Composition using High Resolution Compact Airborne Spectrographic Imager CASI Image Texture.” *Research Report*, Department of Geography, University of Kansas, Lawrence, KS, USA.

Richards, J.A., and Jia, X. (1999). *Remote Sensing Digital Image Analysis: An Introduction*. 3rd edition, Springer-Verlag, New York, 363 pp.

Shaban, M.A., and Dikshit, O. (2001). “Improvement of Classification in Urban Areas by the use of Textural Features: The Case Study of Lucknow City, Uttar Pradesh.” *International Journal of Remote Sensing*, 22 (4) 565-593.

Smith, G.M., and Fuller, R.M. (2001). “An Integrated Approach to Land Cover Classification: An Example in the Island of Jersey.” *International Journal of Remote Sensing,* 22 (16) 3123-3142.