

ADVANCED CLASSIFICATION OF OPTICAL AND SAR IMAGES FOR URBAN LAND COVER MAPPING

Commission I, Symposium

KEY WORDS: Urban land cover, Optical, Synthetic aperture radar (SAR), Feature, Classification method

ABSTRACT:

The aim of this study is to classify urban land cover types using an advanced classification method. As the input bands to the classification, the features derived from the optical and SAR data sets are used. To extract the reliable urban land cover information from the multisource features, a rule-based classification algorithm that uses spatial thresholds defined from the contextual knowledge is constructed. The result of the constructed method is compared with the results of the standard classification technique and it indicates a higher accuracy. Overall, the study demonstrates that the multisource data sets can considerably improve the classification of urban land cover types and the rule-based method is a powerful tool to produce a reliable land cover map.

1. INTRODUCTION

Remote sensing (RS) has provided an important source of information for urban land use and land cover classification, since the appearance of the first digital data sets (Stavroudis *et al.* 2011). As it is one of the advanced techniques used to collect large amount of data with varying spatial resolutions without any physical or direct contact with the object, the extracted information on land cover/use as well as individual objects is essential for rapid urban planning and management at all levels (Erener, 2013). However, urban areas are heterogeneous in nature and are composed of different objects of classes (such as concrete, asphalt, trees, shrubs, grass, water, metal, plastic, glass and soil) which have different spectral characteristics in a RS image. For example, a simple building of building class appears as a complex structure with many architectural details surrounded by gardens, trees, grass, other buildings, roads, social and technical infrastructure and many other temporary objects (Pacifi *et al.* 2009). Therefore, urban area classification and object detection from RS images have been an important research topic for several decades.

Traditionally, multispectral RS data sets have been widely used for urban land-cover mapping (Amarsaikhan *et al.* 2007). However, due to the complex nature and diverse composition of the urban environment, the production of reliable and high quality urban land cover/use maps from the optical images is still a challenging task (Ban *et al.* 2010). In recent years, microwave images have been increasingly used for urban area classification. The studies have shown that SAR images may be the excellent basis for classifying, monitoring and analyzing urban conglomerations and their development over time especially, a large area mapping is under consideration (Dell'Acqua, 2009, Taubenböck *et al.* 2012). As the multispectral and microwave images provide different information, their integration can be efficiently used for an improved urban mapping. It is clear that a combined use of the optical and SAR images will have a number of advantages because a specific urban feature, which is not seen on the passive sensor image may be observable on the microwave image and vice versa because of the complementary information provided by the two sources (Amarsaikhan *et al.* 2007).

During the last decades, a significant progress has been made toward the development of new advanced active and passive RS sensors, with which accurate and detailed mapping of urban land cover and land use could become a reality (Hu and Wang, 2013). However, as the urban areas are complex and diverse in nature and many features have similar spectral characteristics, it is still not easy to separate them by the use of ordinary feature combinations or by applying conventional techniques. Therefore, in urban area mapping, for differentiation of the spectrally similar or mixed classes, reliable features derived from multiple sources and an efficient classification technique should be used (Amarsaikhan *et al.* 2012). The aim of this study is to classify the features derived from optical and SAR data sets and produce a reliable urban land cover map using a rule-based classification method.

2. TEST SITE AND DATA SOURCES

As a test site, Ulaanbaatar, the capital city of Mongolia has been used. Ulaanbaatar city is situated in the central part of Mongolia, on the Tuul River, at an average height of 1350m above sea level and currently has about 1.3 million inhabitants. In the selected image frame of the city, it is possible to define such classes as built-up area, ger (Mongolian traditional dwelling) area, forest, grass, soil and water. The built-up area includes buildings of different sizes, while ger area includes mainly gers surrounded by fences. Figure 1 shows a Landsat image of the test site, and some examples of its land cover.

In the present study, as data sources, a Landsat ETM(+) image of 31 July 2010 and an Envisat image of 25 May 2010 have been used. The Landsat ETM(+) data has seven multispectral bands (B1: 0.45–0.52 μ m, B2: 0.52–0.60 μ m, B3: 0.63–0.69 μ m, B4: 0.76–0.90 μ m, B5: 1.55–1.75 μ m, B6: 10.40–12.50 μ m and B7: 2.08–2.35 μ m). The spatial resolution is 30 m for the reflective bands, while it is 120 m for the thermal band. In the current study, channels 2,3,4,5,7 have been used. The Envisat is a European satellite carrying a cloud-piercing, all weather free polarimetric radar. In the present study, the Envisat C-band (5.6cm) HH polarization image with a spatial resolution of 30m has been used.

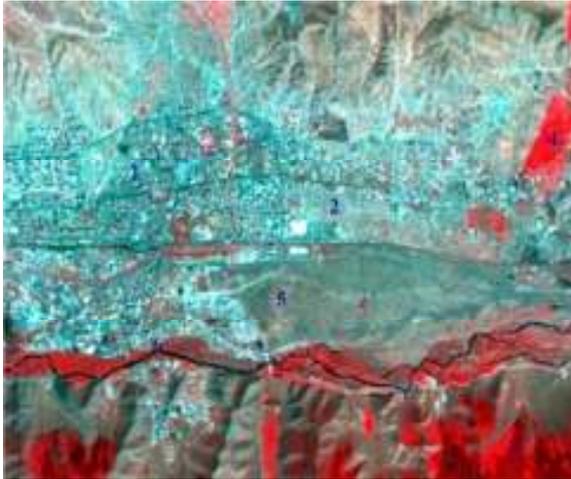


Figure 1. 2010 Landsat image of the selected part of Ulaanbaatar city. 1-built-up area; 2-ger area; 3-forest; 4-grass; 5-soil; 6- water.

3. CO-REGISTRATION OF THE LANDSAT AND ENVISAT IMAGES

At the beginning, a Landsat image was georeferenced to a UTM map projection using 12 ground control points (GCPs) defined from a topographic map of the study area, scale 1:100,000. The GCPs have been selected on clearly delineated crossings of roads, streets and city building corners. For the transformation, a second-order transformation and nearest-neighbour resampling approach were applied and the related root mean square (RMS) error was 0.84 pixel. Then, the Envisat image was geometrically corrected and its coordinates were transformed to the coordinates of the georeferenced Landsat image. In order to correct the Envisat image, 18 more regularly distributed GCPs were selected from different parts of the image. For the actual transformation, a second-order transformation was used. As a resampling technique, the nearest-neighbour resampling approach was applied and the related RMS error was 1.16 pixel.

4. SPECKLE SUPPRESSION OF THE ENVISAT IMAGE

As microwave images have a granular appearance due to the speckle formed as a result of the coherent radiation used for radar systems; the reduction of the speckle is a very important step in further analysis. The analysis of the radar images must be based on the techniques that remove the speckle effects while considering the intrinsic texture of the image frame (Serkan *et al.* 2008). In the current study, four different speckle suppression techniques such as local region, kuan, frost and gammamap filters (ERDAS, 1999) of 3x3 and 5x5 sizes were compared in terms of delineation of urban features and texture information. After visual inspection of each image, it was found that the 3x3 gammamap filter created the best image in terms of delineation of different features as well as preserving content of texture information. In the output image, speckle noise was reduced with very low degradation of the textural information.

5. FEATURE DETERMINATION AND A SUPERVISED CLASSIFICATION METHOD

In any classification process, it is desirable to include different orthogonal features to increase its decision-making. In the present study, for this aim, the following feature combinations were determined:

1. All original spectral bands of the Landsat TM image.
2. Red, near infrared and second middle infrared (i.e., 3,4 and 7) bands of the Landsat TM image.
3. The combined features of Envisat and Landsat TM data sets.
4. The principal component (PC) 1, PC2, PC3 and PC4 of the principal component analysis (PCA).

As seen, for the feature determination, in addition to the original data sets, the first four PCs obtained from the PCA have been chosen. The PCA is a data compression technique used to reduce the dimensionality of the multidimensional datasets. The bands of the PCA data are non-correlated and are often more interpretable than the source data. In n dimensions, there are n principal components. Each successive principal component is the widest transect of the ellipse that is orthogonal to the previous components in the n -dimensional space, and accounts for a decreasing amount of the variation in the data which is not already accounted for by previous principal components. Although there are n output bands in a PCA, the first few bands account for a high proportion of the variance in the data (Richards and Jia, 1999).

In the present study, the PCA has been performed using all available bands and the results are shown in table 1. As can be seen from table 2, PC1 is totally dominated by the variance of the Envisat image and other bands have almost no influence on it. Although, it contained 84.69% of the overall variance, a visual inspection revealed that it contained less information related to the selected classes. The first middle infrared band of Landsat has a high negative loading in PC2. Here, the second middle infrared band of Landsat also has the second highest negative loading. In PC3, near infrared band has a high negative loading and red band has moderately high loading. Although PC3 contained only 2.07% of the overall variance, visual inspection showed that it contained some useful information related to the urban texture. PC4 is dominated by the variances of the red and near infrared bands and it contained only 0.95% of the overall variance. The inspection of the PC5 and PC6 indicated that they mainly contained noise from the total data set. Sometimes, useful information can be gathered from the bands with the least variances and these bands can show subtle details in the image that were obscured by higher contrast in the original image. Therefore, for the final analysis, the first four PCs that contained 98.87% of the overall variance have been selected. Figure 2 shows the comparison of the images obtained by the selected feature combinations.

	PC1	PC2	PC3	PC4	PC5	PC6
Band 2	0.06	-0.21	0.27	0.42	-0.06	-0.82
Band 3	0.09	-0.37	0.44	0.53	-0.26	0.54
Band 4	-0.01	-0.14	-0.75	0.58	0.24	0.06
Band 5	0.00	-0.70	-0.32	-0.38	-0.48	-0.08
Band 7	0.04	-0.53	0.20	-0.20	0.79	0.05
SAR	0.99	0.07	-0.07	-0.05	0.00	0.00
Eigenvalue	4317.7	617.6	105.9	48.7	6.0	2.3
Variance (%)	84.69	12.11	2.07	0.95	0.13	0.05

Table 1. PC coefficients from Landsat TM and Envisat images

To classify the defined features, initially, the training samples had to be selected. To define the sites for the training signature selection from the combined optical and SAR images, several areas of interest (AOI) representing the available six classes (builtup area, ger area, forest, grass, soil and water) have been selected through accurate analysis. The separability of the training signatures was firstly checked in feature space and then evaluated using Jeffries–Matusita distance (table 2). The values

of Jeffries–Matusita distance range from 0 to 2.0 and indicate how well the selected pairs are statistically separate. The values greater than 1.9 indicate that the pairs have good separability (ERDAS, 1999). After the investigation, the samples that demonstrated the greatest separability were chosen to form the final signatures. The final signatures included about 83–835 pixels.

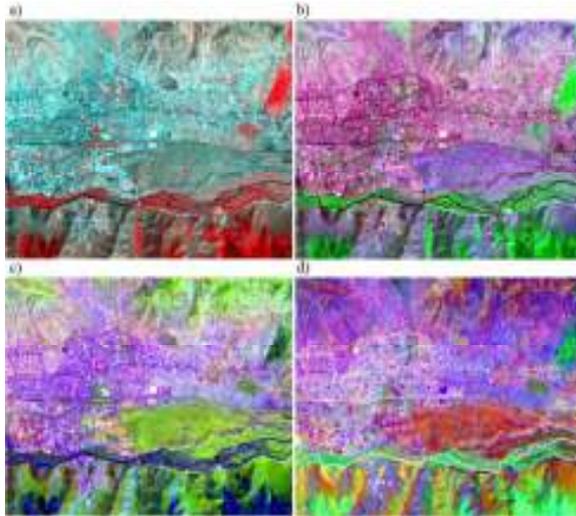


Figure 2. Comparison of the images obtained by the selected feature combinations: (a) the image obtained by Landsat band combinations (R=B4, G= B3, B= B2); (b) the image obtained by selected Landsat band combinations (R= B3, G= B4, B= B7); (c) the image obtained by optical and SAR features (R= B2, G= B5, B=SAR); (d) PC image (R=PC1, G=PC2, B=PC3).

	Builtup area	Ger area	Forest	Grass	Soil	Water
Builtup area	0.000	1.318	1.999	1.999	1.998	1.969
Ger area	1.318	0.000	2.000	2.000	1.999	1.999
Forest	1.999	2.000	0.000	1.915	2.000	1.999
Grass	1.999	2.000	1.915	0.000	2.000	1.999
Soil	1.998	1.999	2.000	2.000	0.000	1.998
Water	1.969	1.999	1.999	1.999	1.998	0.000

Table 2. The separabilities measured by Jeffries–Matusita distance

For the actual classification, a Mahalanobis distance classifier has been used. The Mahalanobis distance classifier is a parametric method, in which the criterion to determine the class membership of a pixel is the minimum Mahalanobis distance between the pixel and the class centre. The sample mean vectors and variance-covariance matrices for each class are estimated from the selected training signatures. Then, every pixel in the dataset is evaluated using the minimum Mahalanobis distance and the class label of the closest centroid is assigned to the pixel (Mather and Koh, 2011).

The classified images for the selected feature combinations are shown in figure 3(a–d). As seen from figure 3(a–d), the classification result of all bands of Landsat TM gives the worst result, because there are high overlaps among two urban classes: built-up area and ger area. However, these overlaps decrease on the classified image of red and infrared bands. It can be explained by a fact that a fewer bands with statistically separable features can produce a better result than many bands

with high overlaps. The combined use of optical and SAR data sets as well as the PC bands produced better results than the results of the Landsat TM bands, but they still contain many mixed pixels for different classes. As seen, although multisource images give some improvement, it is still very difficult to obtain a reliable urban land cover map by the use of the standard technique, specifically on decision boundaries of the statistically overlapping classes.

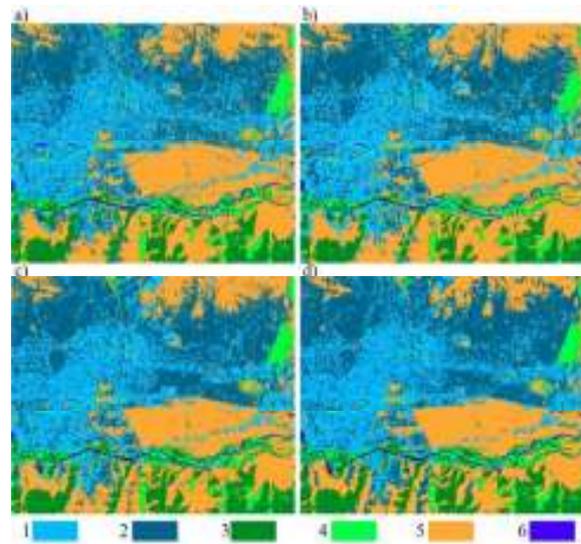


Figure 3. Comparison of the standard classification results for the selected classes (1-built-up area; 2-ger area; 3-forest; 4-grass; 5-soil; 6-water). Classified images (a) using Landsat TM bands, (b) using bands 3, 4 and 5 of the Landsat TM image, (c) using PCs, (d) using multisource bands.

For the accuracy assessment of the classification results, the overall performance has been used. This approach creates a confusion matrix in which reference pixels are compared with the classified pixels and as a result an accuracy report is generated indicating the percentages of the overall accuracy (ERDAS, 1999). As ground truth information, different AOIs containing 1241 purest pixels have been selected. AOIs were selected on a principle that more pixels to be selected for the evaluation of the larger classes such as ger area and soil than the smaller classes such as water. The overall classification accuracies for the selected classes are shown in table 3.

No	The used band combinations	Overall accuracy (%)
1	Landsat TM original bands	84.16
2	Bands 3, 4 and 7 of the Landsat TM	86.43
3	Landsat TM and SAR bands	88.57
4	PC 1, 2, 3 and 4	89.95

Table 3. The overall classification accuracy of the classified images

6. RULE-BASED CLASSIFICATION METHOD

In general, it is very important to design an appropriate image processing procedure in order to successfully classify any digital data into a number of class labels. The effective use of different features derived from different sources and the selection of a reliable classification technique can be a key significance for the improvement of classification accuracy (Lu

and Weng, 2007). In the present study, for the classification of urban land cover types, a rule-based algorithm has been constructed. As the reliable features, the first four PCs defined from the multisource images have been selected.

A rule-based approach is a part of knowledge-based techniques and it uses a hierarchy of rules or a decision tree describing the conditions under which a set of low-level primary objects becomes abstracted into a set of high-level object classes. The primary objects contain the user defined variables and include geographical objects represented in different structures, external programs, scalars and spatial models (Amarsaikhan *et al.* 2012).

The constructed rule-based approach consists of a set of rules, which contains the initial image segmentation procedure based on a Mahalanobis distance rule and the constraints on spatial thresholds. It is clear that a spectral classifier will be ineffective if applied to the statistically overlapping classes such as built-up area and ger area, because they have very similar spectral characteristics in both multispectral and SAR images. For such spectrally mixed classes, classification accuracies can be improved if the spatial properties of the classes of objects could be incorporated into the classification process. These thresholds can be determined on the basis of different knowledge. In the current study, the spatial thresholds were defined on the basis of contextual knowledge about the test area. The contextual knowledge is based on the spectral and textural variations of the selected classes in different parts of the combined optical and SAR images.

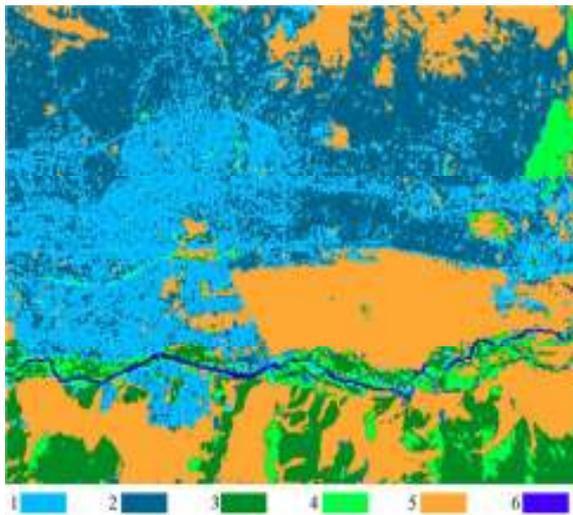


Figure 4. Classification result obtained by the rule-based classification.
(1-built-up area; 2-ger area; 3-forest; 4-grass; 5-soil; 6-water).

In the initial image classification, for separation of the statistically overlapping classes, only pixels falling inside of the spatial thresholds and the first four PCs of the PCA were used. The pixels falling outside of the spatial thresholds were temporarily identified as unknown classes and further classified using the rules in which other spatial thresholds were used. The final urban land cover map was created by a combined use of Spatial Modeler and Knowledge Engineer modules of Erdas Imagine. The image classified by this method is shown in figure 4.

As seen from the classified image, the rule-based approach could very well separate the built-up area from the ger area

compared to the results obtained by the traditional supervised method. The overall classification accuracy has been evaluated using the same set of regions containing the purest pixels as in the previous classification and it demonstrated an improvement to 92.06%.

7. CONCLUSIONS

The main purpose of the research was to classify the multisource data sets using a superior classification method and produce an improved urban land cover map. For the classification decision rule, four different feature combinations such as all original spectral bands of the Landsat TM, bands 3,4 and 7 of the Landsat TM, the combined bands of Envisat and Landsat TM data sets, and first four PCs, were determined. To extract the reliable urban land cover information from the selected multispectral and SAR features, a rule-based classification algorithm that uses spatial thresholds defined from the contextual knowledge was constructed. The result of the rule-based method was compared with the results of the standard supervised classification and it indicated a higher accuracy. Overall, the study demonstrated that the combined use of optical and microwave data sets can considerably improve the classification of urban land cover types and the rule-based method is a powerful tool to produce a reliable land cover map.

REFERENCES

- Amarsaikhan, D., Ganzorig, M., Ache, P. and Blotevogel, H., 2007. The Integrated Use of Optical and InSAR Data for Urban Land Cover Mapping, *International Journal of Remote Sensing*, 28, pp.1161-1171.
- Amarsaikhan, D., Saandar, M., Ganzorig, M., Blotevogel, H.H. and Enkhjargal, D., 2012. Comparison of multisource image fusion methods and land cover classification. *International Journal of Remote Sensing*, 33(8), pp. 2532-2550.
- Ban, Y., Hu, H. and Rangel, I.M, 2010. Fusion of Quickbird MS and RADARSAT SAR data for urban land cover mapping: object-based and knowledge-based approach. *International Journal of Remote Sensing*, 31(6), pp. 1391-1410.
- Dell'Acqua, F., 2009. The role of SAR sensors. In P. Gamba, & M. Herold (Eds.), *Global mapping of human settlements: Experiences, data sets, and prospects*, pp. 209-319.
- ERDAS 1999, *Field Guide*, 5th edn (Atlanta, Georgia: ERDAS, Inc.).
- Erener, E., 2013. Classification method, spectral diversity, band combination and accuracy assessment evaluation for urban feature detection. *International Journal of Applied Earth Observation and Geoinformation*, 21, pp. 397-408.
- Hu, S. and Wang, L., 2013. Automated urban land-use classification with remote sensing. *International Journal of Remote Sensing*, 34(3), pp. 90-803.
- Lu, D. and Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, Volume 28,5, pp.823 – 870.
- Mather, P.M. and Koh, M., 2011. *Computer Processing of Remotely-Sensed Images: an Introduction*. Fourth edition (Wiley-Blackwell).

Pacifici, F., Chini, M. and Emery, J., 2009. A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sensing of Environment*, 113, pp. 1276–1292.

Richards, J.A. and Jia, X., 1999. *Remote Sensing Digital Image Analysis-An Introduction*, Third Edition, (Berlin: Springer-Verlag).

Serkan, M.,Musaoglu, N.,Kirkici, H. and Ormeci, C., 2008. Edge and fine detail preservation in SAR images through speckle reduction with an adaptive mean filter. *International Journal of Remote Sensing*, 29(23), pp.6727 – 6738.

Stavrakoudis, D.G., Theocharis, J.B. and Zalidis, G.C., 2011. A Boosted Genetic Fuzzy Classifier for land cover classification of remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66, pp. 529–544.

Taubenböck, H., Esch, T., Felbier, A., Wiesner, M., Roth, A. and Dech, S., 2012. Monitoring urbanization in mega cities from space. *Remote Sensing of Environment*, 117, pp. 162-176.