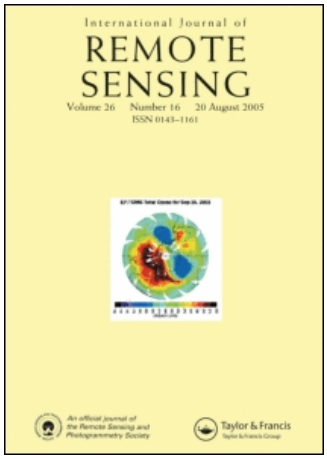

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

The integrated use of optical and InSAR data for urban land-cover mapping

To cite this Article: , 'The integrated use of optical and InSAR data for urban land-cover mapping', International Journal of Remote Sensing, 28:6, 1161 - 1171
To link to this article: DOI: 10.1080/01431160600784267
URL: http://dx.doi.org/10.1080/01431160600784267

International Journal of Remote Sensing
Vol. 28, No. 6, 20 March 2007, 1161–1171

The integrated use of optical and InSAR data for urban land-cover mapping

D. AMARSAIKHAN*†, M. GANZORIG†, P. ACHE‡ and H. BLOTEVOGEL‡
†Institute of Informatics and RS, Mongolian Academy of Sciences, Ulaanbaatar-51, Mongolia
‡Institute of Spatial Planning, University of Dortmund, D-44221 Dortmund, Germany

(*Received 12 December 2005; in final form 1 March 2006*)

The aim of this study is to classify urban land-cover types using the features derived from optical and spaceborne synthetic aperture radar (InSAR) data sets. For the efficient discrimination of the selected classes, a rule-based algorithm that uses the initial image segmentation procedure based on a minimum distance rule and the constraints on spectral parameters and spatial thresholds is constructed. The result of the rule-based method is compared with the results of a standard supervised classification and it demonstrates a higher accuracy. Overall, the research indicates that the integrated features of the optical and InSAR images can significantly improve the classification of land-cover types and the rule-based classification is a powerful tool in the production of a reliable land-cover map.

1. Introduction

As is known, remote sensing (RS) images taken in the optical range of the electro-magnetic spectrum contain information on the reflective and emissive characteristics of the Earth surface features, while the synthetic aperture radar (SAR) images (intensity and coherence) contain information on the surface roughness, texture, dielectric properties and change of the state of natural and man-made objects. In the past, the integrated features of these multisource data sets have been efficiently used for an improved land-cover mapping. It is evident that a combined use of the optical and SAR images will have a number of advantages because a specific object or class which is not seen on the passive sensor image might be seen on the active sensor image and vice versa because of the nature of the used electro-magnetic radiation (Amarsaikhan *et al.* 2000, Amarsaikhan *et al.* 2004).

Optical RS data sets taken from different Earth observation satellites such as Landsat and SPOT have been successfully used for land-cover mapping since the operation of the first Landsat launched in 1972, whereas SAR images taken from space platforms have been widely used for different thematic applications since the launch of the ERS-1/2, JERS-1 and RADARSAT satellites (Amarsaikhan *et al.* 2004). The combined application of data sets from both sources can provide unique information for different thematic studies, because passive sensor images will represent spectral variations of various surface features, whereas microwave data with their penetrating capabilities can provide some additional information. For example, in urban context the optical images provide the information about the

*Corresponding author. Email: amar64@arvis.ac.mn

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160600784267

spectral variations of the urban features, whereas the radar images provide structural information about buildings and street alignment owing to the double bounce scattering (Amarsaikhan and Douglas 2004).

For several decades, digital methods of classification of RS images have been effectively used for land-cover mapping. The early methods mainly involved supervised and unsupervised methods of multispectral images and hence, a great number of techniques have been developed (Amarsaikhan and Ganzorig 1997). Multisource data sets have proved to offer better potential for discriminating between different land-cover types. A number of authors have assessed the potential of multisource images for the classification of different land cover classes (Srinivasan and Richards 1990, Munechika *et al.* 1993, Serpico and Roli 1995, Benediktsson *et al.* 1997, Hegarat-Mascle *et al.* 2000, Gamba and Houshmand 2001, Benediktsson and Sveinsson 2003, Amarsaikhan and Douglas 2004). In RS applications, the most widely used multisource classification techniques are statistical methods, Dempster–Shafer theory of evidence and neural networks (Solberg *et al.* 1996). Any of these techniques can include different rules to involve human expertise in the final decision-making to increase the classification accuracy (Amarsaikhan *et al.* 2000).

In general, urban areas include complex and diverse environments, in which many features have similar spectral characteristics and it is not easy to separate them by the use of ordinary feature combinations or by applying standard techniques. In this study, we wanted to discriminate urban land-cover types using the features derived from optical and spaceborne InSAR data sets. For the efficient discrimination of the selected classes from the multisource images, a rule-based algorithm has been constructed. The initial multisource data sets consisted of ERS-1/2 tandem pass SAR images, JERS-1 SAR intensity image and SPOT XS data of the urban area in Mongolia. In the feature derivation process, coherence and amplitude images as well as other features have been derived from the InSAR data. The analysis was carried out using personal computer (PC)-based ERDAS Imagine 8.6 and ENVI 3.6.

2. Study area and data sources

As a test site, Ulaanbaatar, the capital city of Mongolia has been selected. The area is about 13.2 km × 9.8 km and represents complex urban environment. The selected part of the capital city is characterized by such classes as building area, ger area, forest, grassland, soil and water. Here, the building class includes all residential, commercial and industrial buildings, because they have very similar spectral characteristics. The ger is Mongolian national dwelling and its history goes back to about 2500–3000 BC. It is still used as dwelling everywhere in Mongolia. The entire outside surface of the ger is covered by materials made of wool. The gers situated in urban areas are surrounded by fences that occupy usually $600m^2$–$800m^2$. The families living in ger districts usually have houses made of brick or wood beside their gers.

In order to perform efficient integration of multisource data sets as well as to avoid radar geometric distortions, in the central part of the selected image frame a flat area has been chosen. The data used consisted of C-band (wavelength is 5.66 cm) ERS-1/2 tandem pass (interferometric) SAR single-look complex (SLC) images acquired on 10 and 11 October 1997 with a spatial resolution of 25 m, L-band (wavelength is 23.5 cm) JERS-1 SAR intensity image of April 1997 with a spatial

Figure 1. SPOT XS image of the selected part of Ulaanbaatar (Red=band2, Green=band3, Blue=band1). The size of the displayed area is about 13.2 km × 9.8 km.

resolution of 18 m and a SPOT XS image of 19 June 1997 with a spatial resolution of 20 m. In addition, for ground truth checking a topographic map of 1984, scale 1 : 50 000 and a general urban planning map were available. Figure 1 shows the test area in a SPOT XS image and some examples of its land cover.

3. Derivation of the InSAR coherence and amplitude images

The InSAR coherence images are generated by using both the amplitude and phase information from a pair of SLC images. The coherence is a measure of the variance of the phase difference of the imaged surface in the time between the two SAR data acquisitions. The coherence values range between 0 and 1. If some land surface changes had occurred in a target area between the two image acquisition periods, then coherence is low and if no changes had occurred, then the coherence is high (Weydahl 2001). In general, the coherence over a dense forest and shrub will be the lowest, while for the bare soil, the coherence will be the highest.

Using the BEST (BEST 2005) software, the coherence and amplitude images have been derived as follows:

(1) Initially, 200 ground control points (GCP) regularly distributed over the images were automatically defined using the satellite orbit parameters and the two SLC images were co-registered with 0.1pixel accuracy. Then, a coarse registration followed by a fine registration was performed.
(2) Coherence has been calculated using a 15 × 3 size average filter and the coherence image was generated.
(3) From the complex images, amplitude images were generated.
(4) The preliminary SLC images were converted from the slant range onto a flat ellipsoid surface.

(5) The true size (5800×5800) SAR images were generated using image undersampling by applying a 3×3 size low pass (average) filter.

4. Derivation of the texture features of the SAR products

To derive texture features of the SAR products, occurrence and co-occurrence measures (using a 3×3 window size) were applied to the coherence and average amplitude images of ERS-1/2 and JERS-1 intensity image. The occurrence measures use the number of occurrences of each grey level within the processing window for the texture calculations, while the co-occurrence measures use a grey-tone spatial dependence matrix to calculate texture values (ENVI 1999).

By applying these measures, initially 40 features (i.e. the results of mean, data range, variance, entropy, skewness, homogeneity, contrast and correlation filters) have been obtained, but after thorough checking of each individual feature only eight features, including the results of the mean and data range filters applied to all three images, and the results of the variance filter applied to the ERS-1/2 average amplitude and JERS-1 intensity images, were selected.

5. Geometric correction of the multisensor images

Initially, the SPOT image was geometrically corrected to a universal transverse mercator (UTM) map projection using a topographic map of the study area, scale 1 : 50 000. The GCPs have been selected on clearly delineated crossings of roads, streets, rivers and other clear sites. In total 16 points were selected. For the transformation, a second order transformation and nearest neighbour resampling approach have been applied and the related root mean square (RMS) error was 0.86 pixel.

In order to correct the SAR products, 24 more regularly distributed GCPs were selected comparing the locations of the selected points with other information such as SPOT XS image and the topographic map. Then, the images were geometrically corrected to a UTM map projection using the topographic map of the study area. For the actual transformation, a third-order transformation and nearest neighbour resampling approach were applied and the related RMS errors were 0.98 pixel and 0.94 pixel for the ERS SAR products and JERS-1 SAR image, respectively.

6. Evaluation of the features using a supervised classification method

In general, before applying a classification decision rule, the speckle noise of the SAR images should be reduced. The reduction of the speckle increases the spatial homogeneity of the classes, which in turn improves the classification accuracy. Initially, in order to reduce the speckle of the radar images, a 3×3 gammamap filtering (ERDAS 1999) was applied to the JERS-1 SAR and ERS SAR products. Then, from the multisource images, 2–3 areas of interest (AOI) representing the selected six classes such as building area, ger area, forest, vegetation, soil and water have been selected through accurate analysis using a polygon-based approach. The separabilities of the training signatures were first checked on the feature space images and then evaluated using Jeffries-Matusita (JM) distance (Richards 1993). Then the samples which demonstrated the greatest separabilities were chosen to form the final signatures. The final signatures included about 78–592 pixels. The plot of the mean values for the chosen signatures in the selected bands are shown in

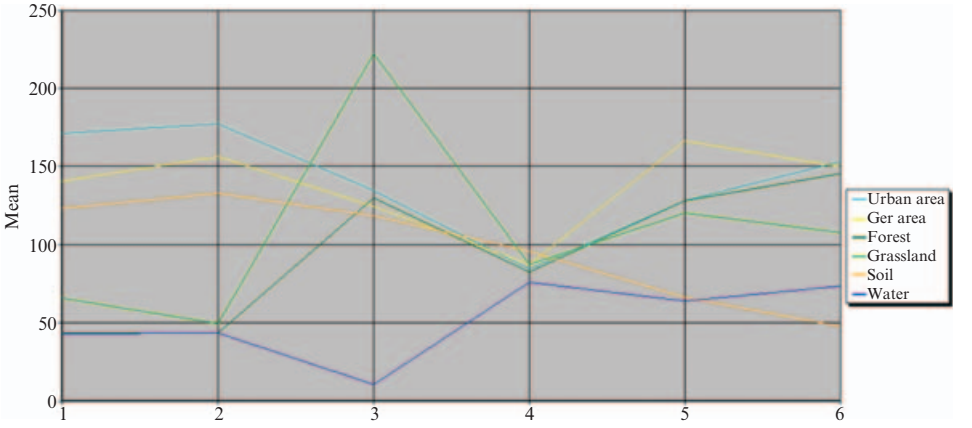

Figure 2. The plot of the mean values for the selected classes. Along the *X*-axes: 1–3 SPOT XS bands; 4 coherence band; 5 average band of ERS1 and 2 SAR; 6 JERS-1 SAR band.

figure 2. For the evaluation of the selected feature combinations, the standard statistical maximum likelihood classifier has been used assuming that the training samples have the Gaussian distribution (Mather 1999). For the classification, the following feature combinations have been used:

(1) Bands 1, 2 and 3 of the SPOT XS.
(2) JERS-1, coherence, average of ERS1 and 2.
(3) The SPOT XS and average of ERS1 and 2.
(4) The SPOT XS and coherence.
(5) The SPOT XS and JERS-1 SAR.
(6) The SPOT XS, coherence and average of ERS1 and 2.
(7) The SPOT XS, JERS-1, coherence, average of ERS1 and 2.
(8) Multiple features, including the original SPOT XS and SAR products as well as eight texture features determined through occurrence and co-occurrence measures.
(9) The first six principal components (PC) of the principal component analysis (PCA) (Richards 1993). The PCA was performed using multiple (14) features.

The images classified by each of the above mentioned feature combinations (except the results of the SPOT XS and original SAR products) are shown in figure 3(*a*)–(*g*). Visual inspection of the classification results of the separate SPOT XS and original SAR products showed that the results were not reliable, because the classified XS image contained high mixtures among building area, ger area and soil classes, whereas the classified image of the original SAR products contained different mixed classes. 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, for each class several regions containing the purest pixels have been selected and overall 4214 pixels (i.e. building area-1028, ger area-463, forest-1297, grassland-216, soil-1142 and water-68 pixels,

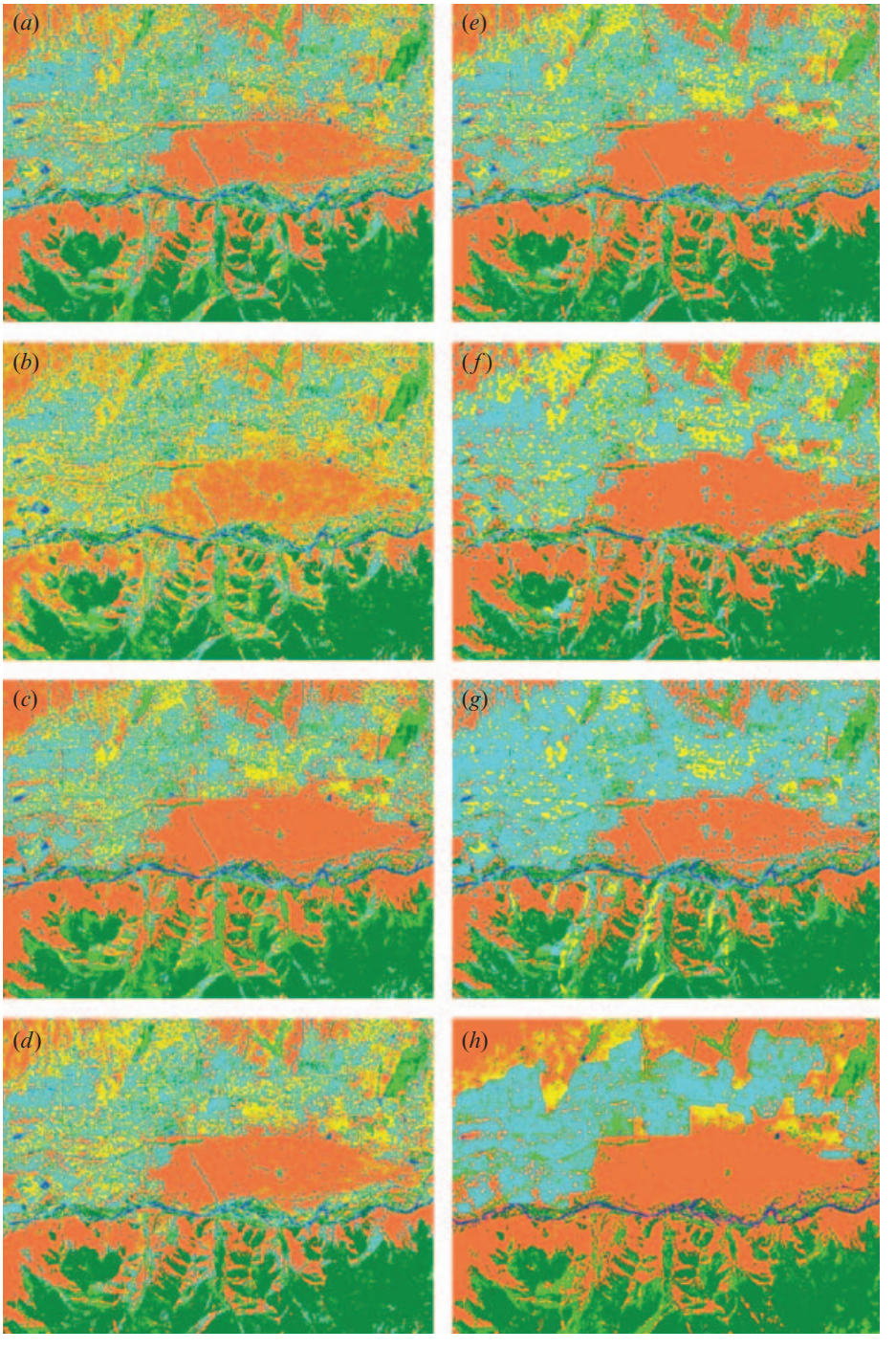

Figure 3. Comparison of the classification results for the selected classes (urban area-cyan, ger area-yellow, forest-dark green, grassland-green, soil-brown, water-blue). (*a*) Classified image using SPOT XS and average of ERS1 and 2. (*b*) Classified image using SPOT XS and coherence. (*c*) Classified image using SPOT XS and JERS-1 SAR. (*d*) Classified image using SPOT XS, coherence and average of ERS1 and 2. (*e*) SPOT XS, JERS-1, coherence, average of ERS1 and 2. (*f*) Multiple features. (*g*) Classified image using the first six PCs of the PCA. (*h*) Result of the rule-based method.

Table 1. The overall classification accuracy of the classified images.

The feature combinations used for the MLC	Overall accuracy (%)
SPOT XS and average of ERS1 and 2	84.12
SPOT XS and coherence	70.09
SPOT XS and JERS-1	84.86
SPOT XS, coherence and average of ERS1 and 2	81.16
SPOT XS, JERS-1, coherence, average of ERS1 and 2	88.87
Multiple features, including the original SPOT XS and SAR products as well as 8 texture features	79.91
The first 6 PCs of the PCA	74.56

respectively) have been chosen. The overall classification accuracies for the selected classes are shown in table 1.

As seen from figure 3 (*a*)–(*g*) and table 1, compared with other results the classification accuracy of combination of the SPOT XS and coherence image is not high. In urban context, coherence cannot add much information, but it might have some contribution to separate the fuzzy boundaries between grassland and forest, because as the state of the grassland is more stable than that of forest, it would result in higher coherence values than forest. However, it is seen that a combination of the SPOT XS with either JERS-1 or original ERS products gave improved classification results. Specifically, the combination of the XS and JERS-1 images gave better result in terms of separation among the statistically overlapping classes such as building area, ger area and soil. This is most probably owing to the nature of the longer wavelength of JERS-1, which can totally separate backscatter of soil that has specular reflection from the backscatter of urban or forest classes which are dominated by double bounce and volume scattering, respectively. Moreover, it is seen that the classification result of combination of the SPOT XS and original SAR products is the highest among the selected feature combinations. This indicates that multisource information collected from optical as well as different radar frequencies can significantly improve the classification result, however, using a standard technique it is very difficult to obtain the most reliable result.

Furthermore, as seen from figure 3 (*a*)–(*g*) and table 1, the accuracy of the classification result of the created 14 features is not high. This is probably owing to the influence of texture features whose variances make confusions in decision boundaries in multidimensional space. Likewise, the accuracy of the PC bands is not high. When the PCA was performed to 14 features, it is interesting to notice that the first three PCs were dominated by the variance of the SAR texture features and contained 99.78% of the overall data variance. When they were classified, the result was not reliable, because they contained different mixed classes. Further inspection of the PCs indicated that PC4 was dominated by the variance of the visible bands of the SPOT XS, while PC5 was dominated by the variance of the original SAR bands. The inspection of other PCs indicated that PC6 was dominated by the variance of the infrared band of the SPOT XS and the rest contained noise from the total data set. Therefore, for the classification the first 6 PCs which contained 99.98% of the total data variance were used. As seen from the classification result of the PC images, there are different mixed classes. This is probably because of the influences of the texture features.

As seen, although the multisource data sets produced better results than the single-source images, they could not separate the two statistically overlapping urban

classes such as building area and ger area. An attempt to separate these classes will be made in the next section, which describes a rule-based classification method.

7. Rule-based classification

Over the years, knowledge-based techniques have been widely used for the digital classification of RS images. The knowledge in image classification can be represented in different forms depending on the type of knowledge and necessity of its usage (Amarsaikhan and Douglas 2004). The most commonly used techniques for knowledge representation are a rule-based approach and neural network classification (Amarsaikhan *et al.* 2000).

In the present study, for separation of the statistically overlapping classes, a rule-based approach has been used. A rule-based approach 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 the high-level object classes. The primary objects contain the user-defined variables and include geographical objects represented in different structures, external programmes, scalars and spatial models (ERDAS 1999).

The constructed rule-based approach consists of a set of rules, which contains the initial image segmentation procedure based on a minimum distance rule (Richards 1993) and the constraints on spectral parameters and spatial thresholds. In the minimum distance estimation, for the initial separation of the classes, only pixels falling within 1.5 standard deviation (SD) (these pixels more clearly represented the related classes) and the features extracted through a feature extraction process, were used. As the reliable features in which the selected classes could be more separable, the PCs extracted through the PCA have been chosen. The PCA has been performed using six bands, including the SPOT XS and original SAR products. The result is shown in table 2.

As seen from table 2, in the PC1 that contains 39.45% of the overall variance, JERS-1 has a very high negative loading, whereas in the PC2 that contains 31.24% of the overall variance, variance of the visible green band of SPOT XS has a total dominance. In the PC3 that contains 17.68% of the overall variance, coherence has a very high negative loading and further visual inspection revealed that this feature contained less information related to the selected classes. Moreover, as seen from the table 2, in the PC4 that contains 8.83% of the overall variance, infrared band of the SPOT XS has a very high negative loading. The inspection of the last two PCs indicated that they contained noise from the total data set. As the PC3 had less information in terms of describing the selected classes, for the final features, PC1,

Table 2. Principal component coefficients from the SPOT XS and original SAR features.

	PC1	PC2	PC3	PC4	PC5	PC6
XS1	0.025	0.988	−0.065	0.093	0.093	−0.019
XS2	0.049	0.104	0.170	0.004	−0.978	−0.031
XS3	0.036	0.090	−0.035	−0.991	0.003	−0.080
Coherence	−0.541	−0.052	−0.820	0.003	−0.175	0.001
Av.ERS-1/2	0.019	0.029	−0.007	−0.077	−0.029	0.995
JERS-1	−0.837	0.014	0.540	−0.050	0.052	0.017
Eigenvalue	6434.11	5094.30	2883.24	1440.38	355.58	103.12
Variance %	39.45	31.24	17.68	8.83	2.18	0.62

PC2 and PC4 have been chosen. For the initial image segmentation, these selected features were evaluated using a minimum distance rule. The decision rule can be written as

$$d_j(x) = (x - m_j)^t (x - m_j) \tag{1}$$

and if SD(x)≤1.5, then a pixel (x) is assigned to the class (*Cj*) where *dj* is minimum.

The pixels falling outside of 1.5 SD were temporarily identified as unknown classes and further classified using the rules in which different spectral and spatial thresholds were used. In general, it is very difficult to separate the classes if they have the same or very similar spectral characteristics. In such a case, the usage of thoroughly defined spectral and spatial thresholds can play an important role for separating the overlapping classes. The spectral thresholds can be defined from the statistics of the training signatures of the available classes, while the spatial thresholds can be determined on the basis of historical data sets or from local knowledge about the site. In this study, the spectral thresholds were determined based on the knowledge about spectral and scattering characteristics of the selected six classes that is using mainly the lower and upper limits of the pixel values falling in between 1.5 SD and 2.0 SD, whereas the spatial thresholds were determined using polygon boundaries defined on the basis of local knowledge. The final land-cover map was created by a combined use of Spatial Modeler and Knowledge Engineer modules of Erdas Imagine. The flowchart for the constructed rule-based approach is shown in figure 4 and the image classified by this method is shown in figure 3(*h*). As seen from the image, the rule-based approach could very well separate the building area from the ger area compared to the results obtained by the maximum likelihood classifier. 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 95.16%.

8. Conclusions

The aim of the study was to discriminate the urban land-cover classes using the features extracted from optical and spaceborne InSAR data sets. The initial multisource data sets included ERS-1/2 tandem pass SAR images, JERS-1 SAR intensity image and SPOT XS data of the urban area in Mongolia. In the feature derivation process, coherence and amplitude images as well as other texture features were derived from the SAR data.

For the efficient classification of the selected classes, a rule-based algorithm that uses the initial image segmentation procedure based on a minimum distance rule and the constraints on spectral parameters and spatial thresholds was constructed. The result of the constructed rule-based algorithm was compared with the results of a standard maximum likelihood classification and it demonstrated a higher accuracy.

Overall, the research demonstrated that the integrated features of the optical and InSAR images can significantly improve the classification of land cover types and the rule-based classification is a powerful tool in the production of an accurate land-cover map that can be directly used for a decision-making process.

Acknowledgement

A part of this research was conducted under the exchange programme between the DAAD and Mongolian Academy of Sciences.

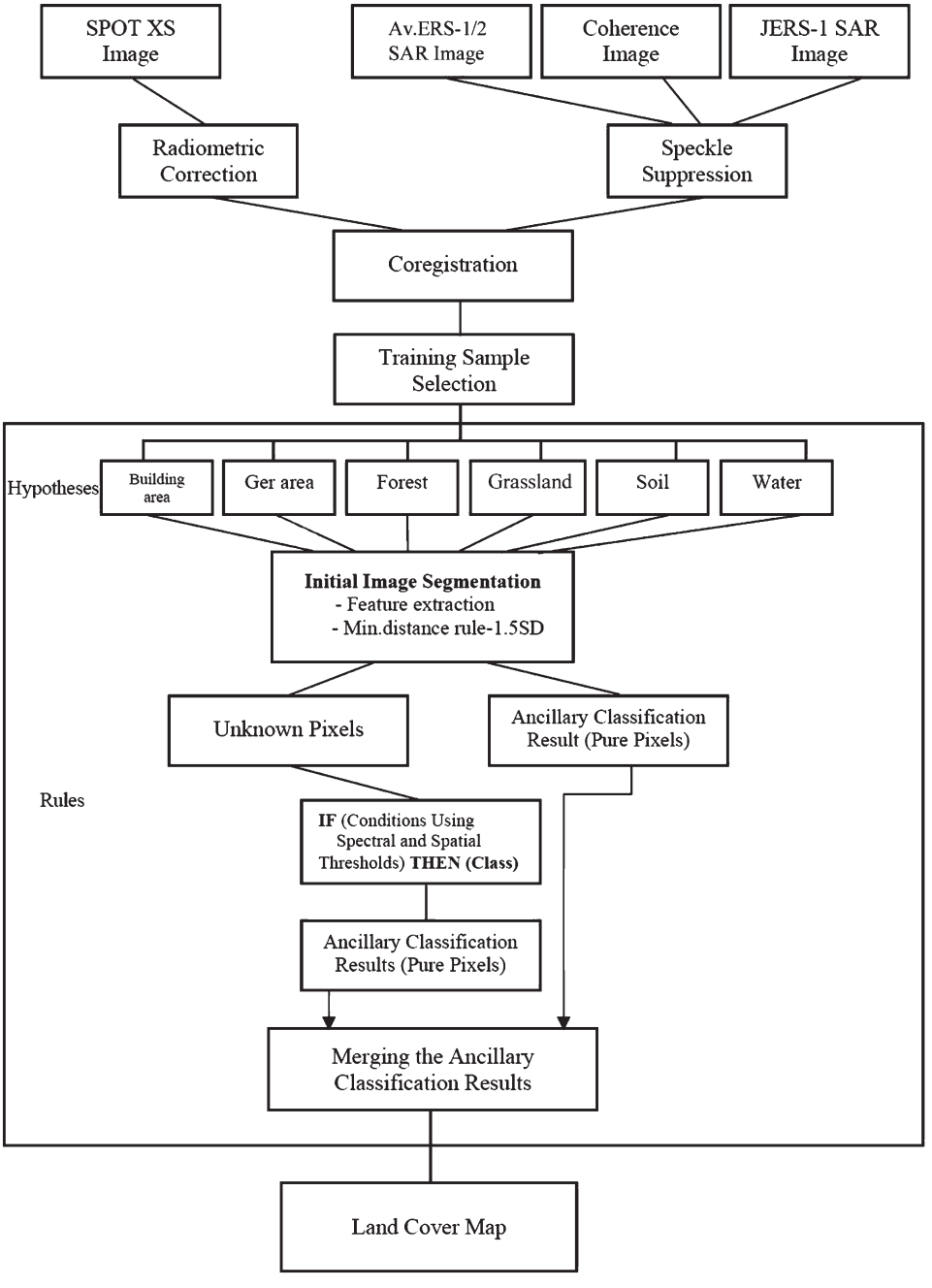

Figure 4. A general diagram for the rule-based classification.

References

AMARSAIKHAN, D. and GANZORIG, M., 1997, Comparison of different pattern recognition techniques. Scientific Papers of Informatics Centre, Mongolian Academy of Sciences, pp. 71–75.

AMARSAIKHAN, D., GANZORIG, M. and ENKHTUVSHIN, B., 2000, Application of spectral and scattering knowledge for interpretation of active and passive sensor data. Institute of Informatics and RS, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia.

AMARSAIKHAN, D., GANZORIG, M., BATBAYAR, G., NARANGEREL, D. and TUMENTSETSEG, S.H., 2004, An integrated approach of optical and SAR images for forest change study. *Asian Journal of Geoinformatics*, **3**, pp. 27–33.

AMARSAIKHAN, D. and DOUGLAS, T., 2004, Data fusion and multisource data classification. *International Journal of Remote Sensing*, **25**, pp. 3529–3539.

BENEDIKTSSON, J.A., SVEINSSON, J.R., ATKINSON, P.M. and TATNALI, A., 1997, Feature extraction for multisource data classification with artificial neural networks. *International Journal of Remote Sensing*, **18**, pp. 727–740.

BENEDIKTSSON, J.A. and SVEINSSON, J.R., 2003, Multisource remote sensing data classification based on consensus and pruning. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 932–936.

BEST, 2005, User Manual, Version 4.0.2.

ENVI, 1999, User's Guide, Research Systems.

ERDAS, 1999, *Field guide*, Fifth edition (Atlanta, Georgia: ERDAS).

GAMBA, P. and HOUSHMAND, B., 2001, An efficient neural classification chain of SAR and optical urban images. *IEEE Transactions on Geoscience and Remote Sensing*, **22**, pp. 1535–1553.

HEGARAT-MASCLE, S.L., QUESNEY, A., VIDAL-MADJAR, D., TACONET, O., NORMAND, M. and LOUMAGNE,, 2000, Land cover discrimination from mutitemporal ERS images and multispectral Landsat images: a study case in an agricultural area in France. *International Journal of Remote Sensing*, **21**, pp. 435–456.

MATHER, P.M., 1999, *Computer Processing of Remotely-Sensed Images: An Introduction*, 2nd edition (Chichester: Wiley).

MUNECHIKA, C.K., WARNICK, J.S., SALVAGGIO, C. and SCHOTT, J.R., 1993, Resolution enhancement of multispectral image data to improve classification accuracy. *Photogrammetric Engineering and Remote Sensing*, **59**, pp. 67–72.

RICHARDS, J.A., 1993, *Remote Sensing Digital Image Analysis: An Introduction*, 2nd edition (Berlin: Springer-Verlag).

SERPICO, S.B. and ROLI, F., 1995, Classification of multisensor remote sensing images by structural neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, pp. 562–578.

SOLBERG, A.H.S., TAXT, T. and JAIN, A.K., 1996, A Markov random field model for classification of multisource satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **34**, pp. 100–112.

SRINIVASAN, A. and RICHARDS, J.A., 1990, Knowledge-based techniques for multisource classification. *International Journal of Remote Sensing*, **11**, pp. 505–525.

WEYDAHL, D.J., 2001, Analysis of ERS SAR coherence images acquired over vegetated areas and urban features. *International Journal of Remote Sensing*, **22**, pp. 2811–2830.