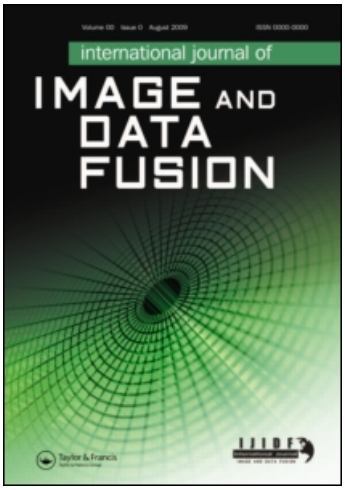

International Journal of Image and Data Fusion
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t911470591

Fusing high-resolution SAR and optical imagery for improved urban land cover study and classification

D. Amarsaikhan [a]; H. H. Blotevogel [b]; J. L. van Genderen [c]; M. Ganzorig [a]; R. Gantuya [a]; B. Nergui [a]
[a] Institute of Informatics and RS, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia [b] Faculty of Spatial Planning Institute of Spatial Planning, Dortmund University of Technology, Dortmund, Germany [c] International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, The Netherlands

Online publication date: 17 February 2010

To cite this Article Amarsaikhan, D., Blotevogel, H. H., van Genderen, J. L., Ganzorig, M., Gantuya, R. and Nergui, B.(2010) 'Fusing high-resolution SAR and optical imagery for improved urban land cover study and classification', International Journal of Image and Data Fusion, 1: 1, 83 — 97
To link to this Article: DOI: 10.1080/19479830903562041
URL: http://dx.doi.org/10.1080/19479830903562041

International Journal of Image and Data Fusion
Vol. 1, No. 1, March 2010, 83–97

Taylor & Francis
Taylor & Francis Group

Fusing high-resolution SAR and optical imagery for improved urban land cover study and classification

D. Amarsaikhan[a]*, H.H. Blotevogel[b], J.L. van Genderen[c], M. Ganzorig[a], R. Gantuya[a] and B. Nergui[a]

[a]*Institute of Informatics and RS, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia;* [b]*Faculty of Spatial Planning Institute of Spatial Planning, Dortmund University of Technology, Dortmund, Germany;* [c]*International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, The Netherlands*

(*Received 28 May 2009; final version received 13 October 2009*)

The two objectives of this study are to compare the performances of different data fusion techniques for the enhancement of urban features and subsequently to improve urban land cover types classification using a refined Bayesian classification. For the data fusion, wavelet-based fusion, Brovey transform, Elhers fusion and principal component analysis are used and the results are compared. The refined Bayesian classification uses spatial thresholds defined from local knowledge and different features obtained through a feature derivation process. The result of the refined classification is compared with the results of a standard method and it demonstrates a higher accuracy. Overall, the research indicates that multi-source information can significantly improves the interpretation and classification of land cover types and the refined Bayesian classification is a powerful tool to increase the classification accuracy.

Keywords: data fusion; refined Bayesian classification; multi-source; urban; feature derivation

1. Introduction

In recent years, very high spatial resolution optical and microwave remote sensing (RS) data have become increasingly available from space platforms and this makes it possible to extract detailed real-time urban land cover information from such data (van Genderen 1989). As it is well known, optical data contain information on the reflective and emissive characteristics of the Earth surface features, while the synthetic aperture radar (SAR) data contain information on the surface roughness, texture and dielectric properties of natural and man-made objects. Over the past decade, the integrated features of these multi-source data sets have been efficiently used for an improved land-cover classification. It is evident that a combined use of the optical and SAR images will have a number of advantages because a specific 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*. 2004, 2007). In addition, the fusion of SAR and optical imagery has a number of additional benefits such as to sharpen images, improve geometric corrections, provide stereo-viewing capabilities for

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

ISSN 1947–9832 print/ISSN 1947–9824 online

DOI: 10.1080/19479830903562041
http://www.informaworld.com

stereophotogrammetry, enhance certain features not visible in either of the single data alone, complement data sets for improved classification, detect changes using multi-temporal data, substitute missing information (e.g. clouds-VIR, shadows-SAR) in one image with signals from another sensor image, replace defective data and for urban change detection (Zeng *et al.* 2010).

Since the review paper on image fusion by Pohl and van Genderen (1998), image data fusion has become a much valuable and popular approach for the integration of multi-source satellite RS data sets. It has been found that the images acquired at different portions of electro-magnetic spectrum provide unique information when they are integrated. Now, image fusion based on the integration of optical and microwave data sets is being efficiently used for the interpretation, enhancement and analysis of different land surface features. Many authors have proposed and applied different techniques to combine optical and SAR images in order to enhance various features and they all judged that the results from the fused images were better than the results obtained from the individual images (Harris *et al.* 1990, Yesou *et al.* 1993, Wang *et al.* 1995, Pohl and van Genderen 1998, Ricchetti 2001, Amarsaikhan and Douglas 2004, Ehlers *et al.* 2008).

Multispectral RS data sets have been widely used for land cover mapping. For the generation of land cover information, diverse classification methods have been applied. The traditional methods mainly involved supervised and unsupervised methods and hence, a great number of techniques have been developed (Amarsaikhan and Ganzorig 1997). Unlike single-source data, data sets from multiple sources have proved to offer better potential for discriminating between different land cover types. Many authors have assessed the potential of multi-source images for the classification of different land cover classes (Munechika *et al.* 1993, Serpico and Roli 1995, Benediktsson *et al.* 1997, Hegarat-Mascle *et al.* 2000, Amarsaikhan and Douglas 2004, Amarsaikhan *et al.* 2007). In RS applications, the most widely used multi-source classification techniques are statistical methods, Dempster–Shafer theory of evidence, neural networks, decision tree classifier and knowledge-based methods (Solberg *et al.* 1996, Franklin *et al.* 2002, Amarsaikhan *et al.* 2007).

Urban areas are complex and diverse in nature, and many features have similar spectral characteristics. In order to separate urban classes successfully, reliable features derived from different sources as well as an efficient classification technique should be selected. The main objectives of this study are as follows:

(1) to compare different data fusion techniques for the enhancement of spectral variations of different urban features, later to be used for training sample selection, and
(2) to classify urban land cover types subsequently using a refined Bayesian classification. For the actual analysis, high-resolution TerraSAR and QuickBird images of the urban area in Mongolia were used. The analysis was carried out using ERDAS Imagine 9.1 and ENVI 4.3.

2. Test site

As a test site, Ulaanbaatar, the capital city of Mongolia has been selected. Ulaanbaatar is situated in the central part of Mongolia, on the Tuul River, at an average height of 1350 m above sea level and currently has about one million inhabitants. Although, the city is extended from the west to the east about 30 km, and from the north to the south about

Figure 1. 2006 QuickBird image of the selected part of Ulaanbaatar. (1) built-up area; (2) ger area; (3) open area; (4) central squire; (5) roads; and (6) snow-ice.
Note: The size of the displayed area is about 3.15 km × 215 km.

20 km, the study area chosen for the present study covers a very small portion (it is about 3.15 km from the west to the east and about 2.15 km from the north to the south). As the images have very high spatial resolution, it is possible to define such classes as built-up area, ger area (Mongolian traditional dwelling), open area, road, central squire and snow-ice. The built-up area includes buildings of different sizes, while ger area includes mainly gers surrounded by fences. Open area includes bare soil, pedestrian walking areas and sparsely distributed non-green vegetation (because the image was acquired in March 2008 and in this time of the year vegetation is not yet green in Mongolia). The road class includes various asphalt roads of different sizes and scales. The central squire represents the squire located in the heart of Ulaanbaatar city. The snow-ice class includes water bodies in a solid form. Figure 1 shows the QuickBird image of the test site and some examples of its land cover.

3. Data sources

In this study, for the urban land cover studies, a QuickBird image of 16 March 2006 and a TerraSAR-X image of 20 March 2008 have been used. The QuickBird data have four multispectral bands (B1: 0.45–0.52 μm, B2: 0.52–0.60 μm, B3: 0.63–0.69 μm and B4: 0.76–0.90 μm) and one panchromatic band (Pan: 0.45–0.9 μm). The spatial resolution is 0.61 m for the panchromatic image, while it is 2.4 m for the multispectral bands. In this study, green, red and near infrared bands have been used. TerraSAR-X is a German Earth observation satellite carrying a cloud-piercing, night-vision radar, which is designed

Table 1. The characteristics of the TerraSAR-X data.

Parameter	X-band
Polarisation	HH and VV
Frequency	9.6 GHz
Wavelength	3.1 cm
Spatial resolution	1.0 m

to create the most precise maps and images ever produced by a civilian space radar system. It images the Earth's surface at a rate of one million square kilometres a day and provides information at various spatial resolutions. The characteristics of the TerraSAR-X data used in the current study are shown in Table 1.

4. Co-registration of optical and SAR images

In this study, the intensity images of QuickBird and TerraSAR-X have been used. Therefore, there was no need to apply special procedure to extract grey-scale images from the SAR images. Generally, in order to perform accurate data fusion, high geometric accuracy between the images is needed. As a first step, the QuickBird image was georeferenced to a Gauss–Kruger map projection using nine ground control points (GCPs) defined from a field survey. 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 error (RMSE) was 1.06 pixels. Then, the TerraSAR image was geometrically corrected based on the georeferenced QuickBird image. In order to correct the SAR image, 15 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 RMSE was 1.48 pixels. As both optical and microwave images had a very high spatial resolution, the errors of less than 1.5 m were considered as acceptable for further studies.

5. Speckle suppression of the TerraSAR 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 (Ulaby *et al*. 1986, Amarsaikhan and Douglas 2004, Serkan *et al*. 2008). In this study, five different speckle suppression techniques such as local region, median, lee-sigma, frost and gamma map filters (ERDAS 1999) of 5×5 and 7×7 sizes were compared in terms of delineation of urban features and texture information. After visual inspection of each image, it was found that the 5×5 gamma map 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.

6. Image fusion

The concept of image fusion refers to a process that integrates different images from different sources to obtain more information from a single, more complete image, considering a minimum loss or distortion of the original data. In other words, the image fusion is the integration of different digital images in order to create a new image and obtain more information than can be separately derived from any of them (Pohl and van Genderen 1998, Ricchetti 2001, Amarsaikhan *et al*. 2009). In the present study of urban areas, the SAR image provides structural information about buildings and street alignment due to the double bounce effect, while the optical image provides the information about the spectral variations of different urban features. Image fusion can be performed at pixel, feature and decision levels (Abidi and Gonzalez 1992, Pohl and van Genderen 1998). In this study, data fusion has been performed at a pixel level and the following techniques were applied:

(1) wavelet-based fusion,
(2) Brovey transform,
(3) Elhers fusion,
(4) principal component analysis (PCA).

Each of these techniques is briefly discussed below.

Wavelet-based fusion: The wavelet transform decomposes the signal based on elementary functions so called the wavelets. By using this, an image is decomposed into a set of multi-resolution images with wavelet coefficients. For each level, the coefficients contain spatial differences between two successive resolution levels. In general, a wavelet-based image fusion can be performed by either replacing some wavelet coefficients of the low-resolution image with the corresponding coefficients of the high-resolution image or by adding high resolution coefficients to the low-resolution data (Pajares and Cruz 2004). In this study, the first approach, which is based on bi-orthogonal transforms, has been applied.

Brovey transform: This is a simple numerical method used to merge different digital data sets. The algorithm based on a Brovey transform uses a formula that normalises multispectral bands used for a red, green, blue colour display and multiplies the result by high-resolution data to add the intensity or brightness component of the image (Vrabel 1996). For the Brovey transform, the bands of QuickBird data were considered as the multispectral bands, while the HH-polarisation of TerraSAR image was considered as the multiplying panchromatic band.

Elhers fusion: This is a fusion technique used for the spectral characteristics preservation of multi-temporal and multi-sensor data sets. The fusion is based on an IHS transformation combined with filtering in the Fourier domain and the IHS transform is used for optimal colour separation. As the spectral characteristics of the multispectral bands are preserved during the fusion process, there is no dependency on the selection or order of bands for the IHS transform (Ehlers 2004, Ehlers *et al*. 2008).

PCA: The most common understanding of the PCA is that it is a data compression technique used to reduce the dimensionality of the multidimensional data sets (Richards and Xia 1999). It is also helpful for image encoding, enhancement, change detection and multi-temporal dimensionality (Pohl and van Genderen 1998). PCA is a statistical

Table 2. Principal component coefficients from TerraSAR and QuickBird images.

	PC1	PC2	PC3	PC4	PC5
TerraSAR HH	0.07	0.07	0.57	0.58	0.56
TerraSAR VV	0.66	0.73	−0.06	−0.05	−0.05
QuickBird B2	0.00	0.00	−0.35	0.8	−0.47
QuickBird B3	−0.04	0.04	0.73	−0.06	−0.67
QuickBird B4	−0.73	0.67	−0.04	0.01	0.04
Eigenvalue	11,892.9	6156.2	532.3	297.1	33.4
Variance (%)	62.88	32.55	2.81	1.57	0.19

technique that transforms a multivariate data set of inter-correlated variables into a set of new uncorrelated linear combinations of the original variables, thus generating a new set of orthogonal axes. In this study, the PCA has been performed using all available bands and the results are shown in Table 2.

As can be seen from Table 2, PC1 and PC2 are dominated by the variance of VV polarisation of TerraSAR, whereas infrared band of QuickBird has a high negative loading in PC1 and the second highest loading in PC2. Although, PC3 contained 2.81% of the overall variance and had high loadings of red band of QuickBird and HH polarisation of TerraSAR, visual inspection revealed that it contained less information related to the selected classes. However, visual inspection of PC4 that contained only 1.57% of the overall variance, in which green band had a very high loading, revealed that this feature contained useful information related to the urban texture. The inspection of the PC5 indicated that it contained noise from the total data set.

In order to obtain a good colour image that can illustrate spectral and spatial variations of the classes in the selected image frame, different band combinations have been used. Most of the methods created good images on the basis of visual interpretation. The images created by the wavelet-based fusion, Brovey transform and Elhers fusion looked much similar to the original QuickBird image, but they somehow reflected the characteristics of the SAR image. Although, these images looked much similar to one another, detailed analysis of each image revealed that the Brovey transformed image gave a superior image in terms of the spatial separation between different objects and classes. Compared to the Brovey transformed image, the images obtained by the other two methods contained speckle noise of the SAR image though they clearly described individual objects. The PC image demonstrated very much the characteristics of both optical and radar images because of their high loadings in the selected PCs. Figure 2 shows the comparison of the images obtained by different fusion methods.

7. Classification of the images

7.1 *Derivation of features and standard Bayesian classification*

Initially, in order to increase the spatial homogeneity of the data, to the TerraSAR image, a 3×3 average filtering was applied. Then, to derive texture features from the multi-source images, contrast, entropy and dissimilarity measures (using a 15×15 window size) have been applied and the results were compared. The bases for these measures are the co-occurrence measures that use a grey-tone spatial dependence matrix to calculate texture

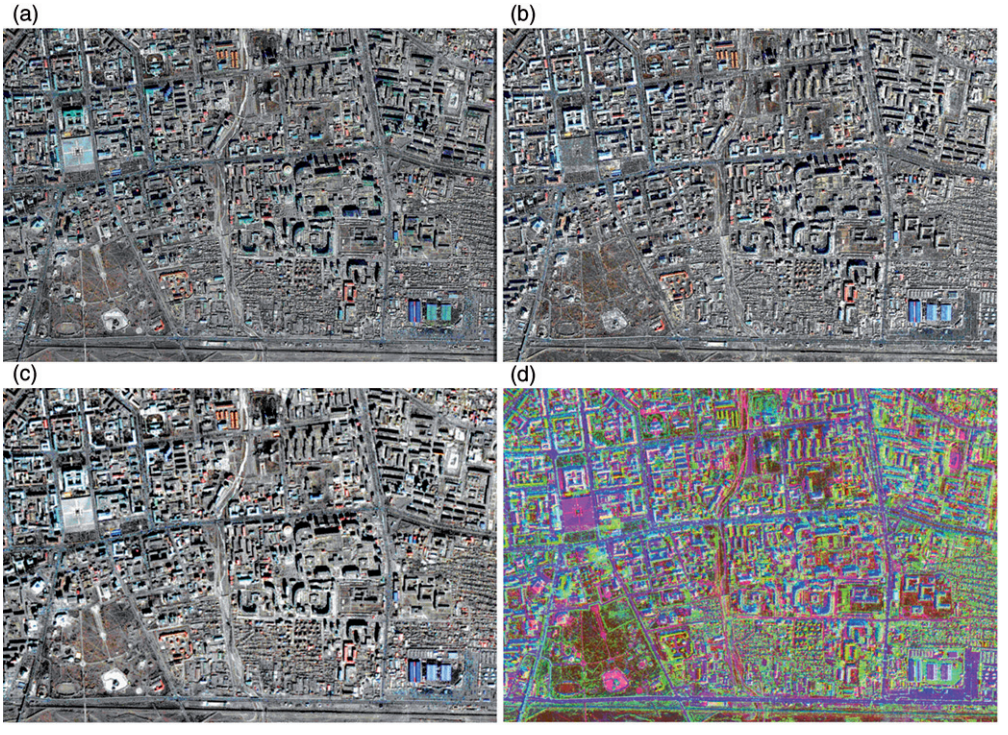

Figure 2. Comparison of the fused images: (a) the image obtained by wavelet-based fusion; (b) Brovey transformed image; (c) the image obtained by Elhers fusion; and (d) PC image (red = PC1, green = PC2, blue = PC3).

values, and the matrix shows the number of occurrences of the relationship between a pixel and its specified neighbour (ENVI 1999). The contrast measure indicates how most elements do not lie on the main diagonal, whereas, the entropy measures the randomness and it will have its maximum when all elements of the co-occurrence matrix are the same. The dissimilarity measure indicates how different the elements of the co-occurrence matrix are from each other (Lee *et al*. 2004). By applying these measures, initially 15 features have been derived, but after thorough checking of each individual feature only four features, including the results of the entropy measure applied to HH polarisation image of TerreaSAR and infrared band of QuickBird, and the results of the dissimilarity measure applied to VV polarisation image of TerreaSAR and infrared band of QuickBird, were selected.

To define the sites for the training signature selection from the multi-sensor images, two to four areas of interest (AOI) representing the available six classes (built-up area, ger area, open area, road, central squire and ice) have been selected through the analysis of fused images. As the data sources included both optical and SAR features, the fused images were very useful for the determination of the homogeneous AOI as well as for the initial intelligent guess of the training sites. The separability of the training signatures was first checked in feature space and then evaluated using Jeffries–Matusita distance (Table 3). 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

Table 3. The separabilities measured by Jeffries–Matusita distance.

	Built-up area	Ger area	Open area	Roads	Squire	Snow
Built-up area	0.000	0.823	1.308	1.023	1.697	1.901
Ger area	0.823	0.000	0.967	1.659	1.960	1.904
Open area	1.308	0.967	0.000	1.801	1.968	1.992
Roads	1.023	1.659	1.801	0.000	1.431	1.968
Squire	1.697	1.960	1.968	1.431	0.000	1.975
Snow-ice	1.901	1.904	1.992	1.968	1.975	0.000

that the pairs have good separability (ENVI 1999, ERDAS 1999). After the investigation, the samples that demonstrated the greatest separability were chosen to form the final signatures. The final signatures included about 3415–10,5343 pixels. For the classification, the following feature combinations were used:

(1) The original spectral bands of the QuickBird data.
(2) The HH and VV polarisation components of TerreaSAR and original spectral bands of the QuickBird data.
(3) Multiple bands including the original TerreaSAR and QuickBird images as well as four other derivative bands obtained from texture measures.
(4) The PC1, PC2, PC3 and PC4 of the PCA (PCA was performed using nine bands including the original TerreaSAR and QuickBird images as well as four texture features and, the first four PCs included 99.9% of the overall variance).

For the actual classification, standard Bayesian maximum likelihood classification (MLC) has been used assuming that the training samples have the Gaussian distribution (Mather 1999). The MLC is the most widely used statistical classification technique, because a pixel classified by this method has the maximum probability of correct assignment (Erbek *et al*. 2004). The decision rule assuming Bayes' rule can be written as follows:

$$P(Ci|x) = P(x|Ci) \times P(Ci)/P(x)$$

where $P(Ci|x)$ is the posterior probability, $P(x|Ci)$ the conditional probability, $P(Ci)$ the prior probability and $P(x)$ the probability of finding a pixel from any class. The actual classification is performed according to $P(Ci|x) > P(Cj|x)$ for all $j \neq i$.

In this study, for the available classes, as each of them has different areas, the following prior probabilities have been used:

$P(\text{built-up area}) = 0.25$,
$P(\text{ger area}) = 0.2$,
$P(\text{open area}) = 0.2$,
$P(\text{road}) = 0.15$,
$P(\text{central squire}) = 0.1$,
$P(\text{ice}) = 0.1$.

As the built-up area, ger area and open area have larger areas in the image frame than the other classes, to these classes higher prior probabilities were assigned. To increase the reliability of the classification, to the initially classified images, a fuzzy convolution with a 5×5 size window was applied. The fuzzy convolution creates a thematic layer by

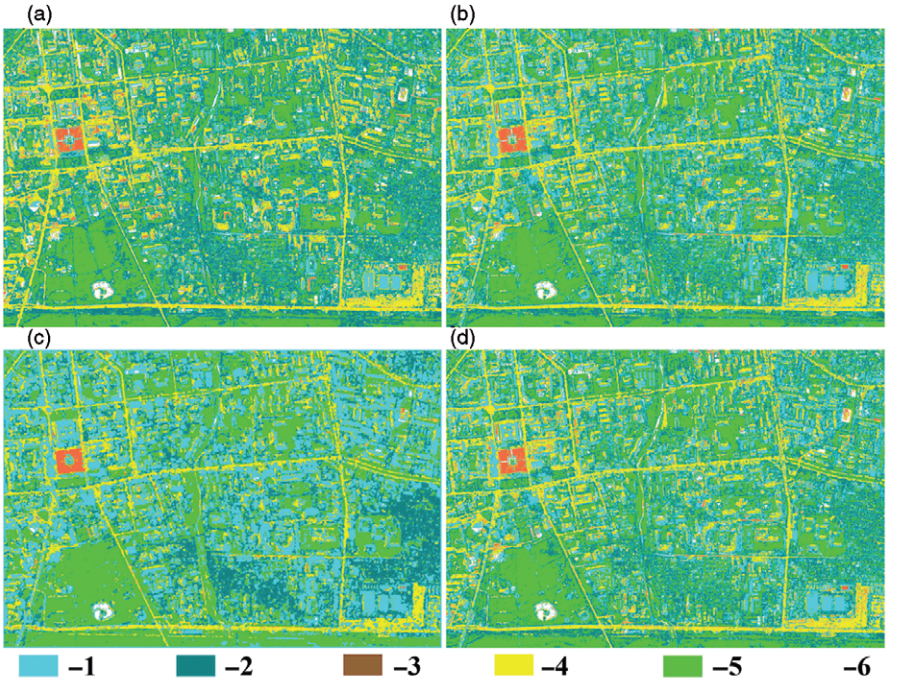

Figure 3. Comparison of the standard classification results for the selected classes: (1) built-up area; (2) ger area; (3) central squire; (4) roads; (5) open area; (6) snow-ice.
Note: Classified images (a) using QuickBird bands, (b) using QuickBird and TerraSAR bands, (c) using multiple bands and (d) using the PCs.

calculating the total weighted inverse distance of all the classes in a determined window of pixels and assigning the centre pixel the class with the largest total inverse distance summed over the entire set of fuzzy classification layers, i.e. classes with a very small distance value will remain unchanged while the classes with higher distance values might change to a neighbouring value if there are a sufficient number of neighbouring pixels with class values and small corresponding distance values (ERDAS 1999). The visual inspection of the fuzzy convolved images indicated that there are some improvements on the borders of the neighbouring classes that significantly influence the separation of the decision boundaries in multidimensional feature space. The final classified images are shown in Figure 3(a)–(d). As seen from Figure 3(a)–(d), the classification result of the QuickBird image gives the worst result, because there are high overlaps among classes: built-up area, ger area and open area. However, these overlaps decrease on other images for the classification of which SAR and optical bands as well as other derivative features have been used. As could be seen from the overall classification results (Table 4), although the combined use of optical and microwave data sets produced a better result than the single-source image, it is still very difficult to obtain a reliable land cover map by the use of the standard technique, specifically on decision boundaries of the statistically overlapping 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

Table 4. The overall classification accuracy of the classified images.

	Reference data					
Classified data	Built-up area	Ger area	Open area	Road	Central squire	Ice
Panel A: The original spectral bands						
Built-up area	9183	1967	1492	409	295	497
Ger area	2498	9125	558	381	243	186
Open area	1956	1433	9126	1247	228	565
Road	956	59	398	7126	387	162
Central squire	223	36	302	107	2578	0
Ice	127	48	561	84	86	2235
Total	14943	12,668	12,437	9354	3817	3645
Panel B: TerreaSAR and QuickBird data						
Built-up area	10,421	1828	1109	211	203	498
Ger area	2241	9738	412	192	191	270
Open area	1039	898	9859	621	278	567
Road	793	86	276	8031	436	151
Central squire	259	67	298	198	2651	0
Ice	190	51	483	101	58	2159
Total	14,943	12,668	12,437	9354	3817	3645
Panel C: Multiple bands						
Built-up area	12,192	798	774	248	568	982
Ger area	1985	11026	399	183	56	561
Open area	211	693	10,740	509	129	474
Road	329	79	297	8227	272	0
Central squire	226	72	130	187	2792	0
Ice	0	0	97	0	0	1628
Total	14,943	12,668	12,437	9354	3817	3645
Panel D: PC1, PC2, PC3 and PC4						
Built-up area	10,533	1811	1123	219	221	423
Ger area	2293	9891	454	201	187	234
Open area	1021	789	9882	478	179	554
Road	787	57	301	8311	386	124
Central squire	203	62	312	85	2812	0
Ice	106	58	365	60	32	2310
Total	14,943	12,668	12,437	9354	3817	3645

Note: Overall accuracy for Panel A: 69.24% (39,373/56,864); Panel B: 75.37% (42,859/56,864); Panel C: 81.96% (46,605/56,864); Panel D: 76.92% (43,739/56,864).

indicating the percentages of the overall accuracy (ERDAS 1999). As ground truth information, different AOIs containing 56,864 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 built-up area and open area, than the smaller classes, such as central squire and snow-ice. The overall classification accuracies for the selected classes are shown in Table 4.

7.2 *The refined Bayesian classification*

For several decades, single-source multispectral data sets have been effectively used for a land cover mapping. Unlike single-source data, multi-source data sets have proved

Figure 4. Classification result obtained by the refined method.

Table 5. The overall classification accuracy of the classified image using the refined method.

	Reference data					
Classified data	Built-up area	Ger area	Open area	Road	Central squire	Ice
Built-up area	14,563	76	298	284	671	1291
Ger area	0	12,318	0	0	0	0
Open area	234	274	11,954	754	0	106
Road	146	0	152	8316	430	391
Central squire	0	0	33	0	2716	0
Ice	0	0	0	0	0	1857
Total	14,943	12,668	12,437	9354	3817	3645

Note: Overall accuracy = 90.96% (51,724/56,864).

to offer better potential for discriminating between different land cover types. Generally, it is very important to design a suitable image processing procedure in order to classify any RS data successfully 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 this study, for the classification of urban land cover types, a refined Bayesian statistical

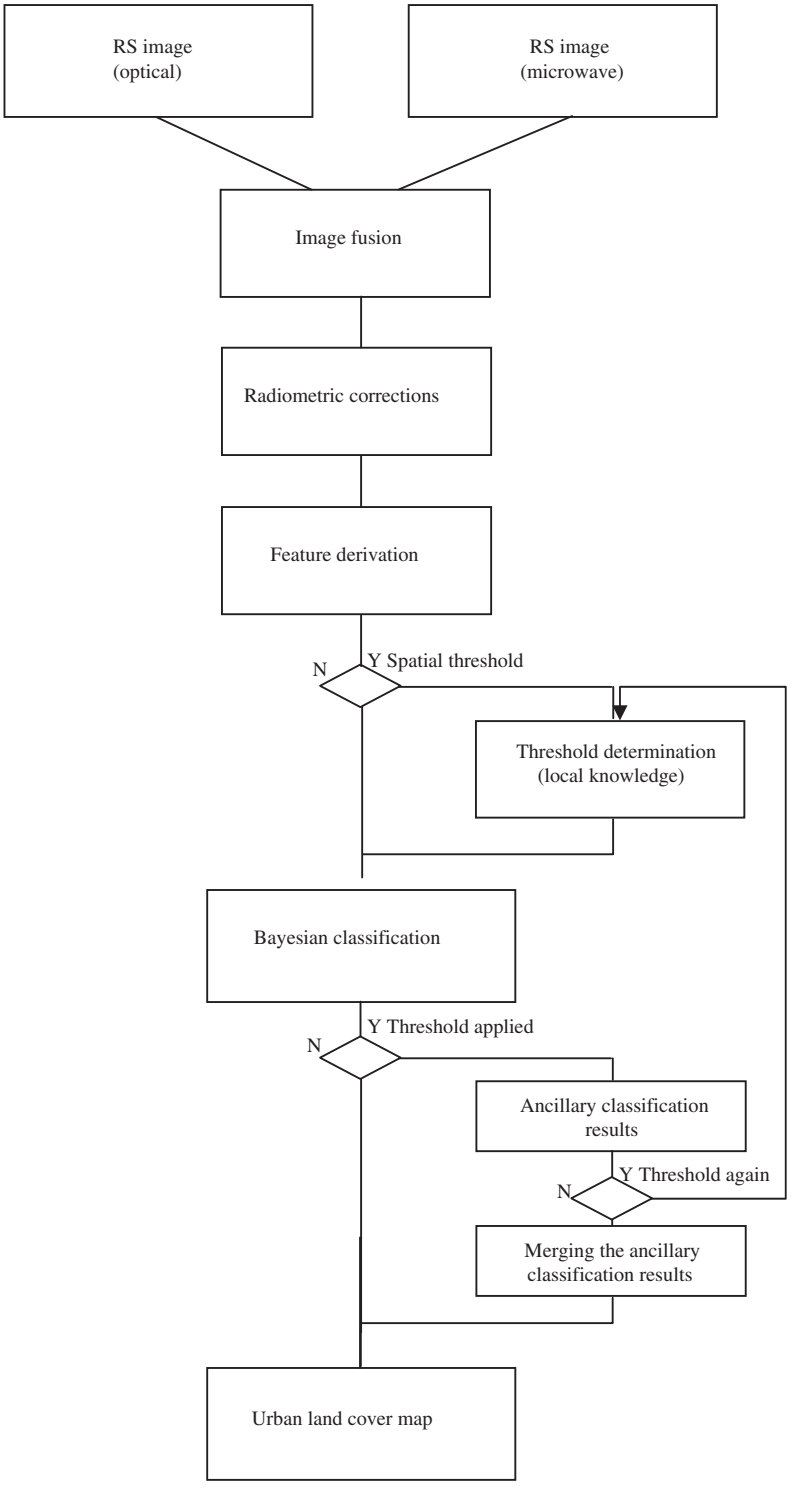

Figure 5. A general diagram for the refined Bayesian classification.

MLC algorithm has been constructed. As features, multiple bands that include the original TerraSAR and QuickBird images, as well as four other derivative bands obtained from texture measures have been used.

Unlike the traditional Bayesian classification, the constructed classification algorithm uses spatial thresholds defined from the local knowledge. The local knowledge was defined on the basis of the spectral variations of the land surface features on the fused images as well as the texture information delineated on the dissimilarity images. 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. For such spectrally mixed classes, classification accuracies should be improved if the spatial properties of the classes of objects could be incorporated into the classification criteria. The idea of the spatial threshold is that it uses a polygon boundary to separate the overlapping classes and only the pixels falling within the threshold boundary are used for the classification. In that case, the likelihood of the pixels to be correctly classified will significantly increase, because the pixels belonging to the class that overlaps with the class to be classified using the threshold boundary are temporarily excluded from the decision-making process. In such a way, the image can be classified several times using different threshold boundaries and the results can be merged (Amarsaikhan and Sato 2004). As prior probabilities, the probabilities used in the standard method (i.e., built-up area = 0.25, ger area = 0.2 open area = 0.2, road = 0.15, central squire = 0.1 and ice = 0.1) have been used.

The result of the classification using the refined method is shown in Figure 4. For the accuracy assessment of the classification result, the overall performance has been used, taking the same number of sample points as in the previous classifications. The confusion matrix produced for the refined classification method showed overall accuracy of 90.96% (Table 5). As could be seen from Figure 4, the result of the classification using the refined Bayesian classification is much better than result of the standard method. The classification accuracy would have been still higher if images of the same year had been used. As it was, there was a time difference of two years between the optical and SAR scene. Hence, some misclassifications may have been due to changes in the urban area because of infrastructure developments. A general diagram of the refined Bayesian classification is shown in Figure 5.

8. Conclusions

The overall idea of the research was to compare the performances of different data fusion techniques for the enhancement of different urban features to be used for the training site selection and subsequent enhanced classification of urban land cover classes of Ulaanbaatar city, Mongolia using a refined classification methodology. For the data fusion, wavelet-based fusion, Brovey transform, Elhers fusion and PCA were used. Although, the fused images looked very similar, detailed analysis of each image revealed that the Brovey transformed image gave a superior image in terms of the spatial separation among different urban features. To extract the reliable urban land cover information from the selected RS data sets, a refined Bayesian classification algorithm that uses spatial thresholds defined from the local knowledge was constructed. Overall, the study demonstrated that multi-source information can significantly improve the interpretation and classification of land cover types and the refined Bayesian classification is a powerful

tool to produce a reliable land cover map. In addition, it could be seen that the method can produce a map that is ready-to-use for decision-making, when there is a set of proper spatial thresholds applied to multi-source data sets in which optical and SAR images as well as historical thematic information are integrated.

Acknowledgements

A part of this research was conducted under the sponsorship for Humboldt Fellows. The authors also gratefully acknowledge the free TerraSAR data provided by Infoterra GmbH as part of a joint research project between ITC and Mongolia.

References

Abidi, M.A. and Gonzalez, R.C., 1992. *Data fusion in robotics and machine intelligence*. New York: Academic Press.

Amarsaikhan, D. and Ganzorig, M., 1997. Comparison of different pattern recognition techniques. *Scientific Papers of Informatics Centre*, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia, 71–75.

Amarsaikhan, D. and Douglas, T., 2004. Data fusion and multisoure data classification. *International Journal of Remote Sensing*, 25 (17), 3529–3539.

Amarsaikhan, D. and Sato, M., 2004. Validation of the Pi-SAR data for land cover mapping. *Journal of the Remote Sensing Society of Japan*, 24 (2), 133–139.

Amarsaikhan, D., *et al*., 2004. An integrated approach of optical and SAR images for forest change study. *Asian Journal of Geoinformatics*, 3, 27–33.

Amarsaikhan, D., *et al*., 2007. The integrated use of optical and InSAR data for urban land cover mapping. *International Journal of Remote Sensing*, 28, 1161–1171.

Amarsaikhan, D., *et al*., 2009. Applications of remote sensing and geographic information systems for urban land-cover changes studies. *Geocarto International*, 24 (4), 257–271.

Benediktsson, J.A., *et al*., 1997. Feature extraction for multi-source data classification with artificial neural networks. *International Journal of Remote Sensing*, 18, 727–740.

Ehlers, M., 2004. Spectral characteristics preserving image fusion based on Fourier domain filtering. *In: Remote sensing for environmental monitoring, GIS applications, and geology IV, proceedings of SPIE*, Bellingham, WA, 93–116.

Ehlers, M., Klonus, S., and Åstrand, P.J., 2008. Quality Assessment for multi-sensor multi-date image fusion, *CD-ROM proceedings of ISPRS congresses*, 3–11 July 2008, Beijing, China.

ENVI, 1999, User's guide, Research systems.

Erbek, F.S., Zkan, C.O., and Taberner, M., 2004. Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. *International Journal of Remote Sensing*, 25 (9), 1733–1748.

ERDAS 1999, *Field guide*. 5th ed. Atlanta, GA: ERDAS, Inc.

Franklin, S.E., *et al*., 2002. Evidential reasoning with Landsat TM, DEM and GIS data for landcover classification in support of grizzly bear habitat mapping. *International Journal of Remote Sensing*, 23 (21), 4633–4652.

Harris, J.R., Murray, R., and Hirose, T., 1990. IHS transform for the integration of radar imagery with other remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 56, 1631–1641.

Hegarat-Mascle, S.L., *et al*., 2000. Land cover discrimination from muti-temporal ERS images and multispectral Landsat images: a study case in an agricultural area in France. *International Journal of Remote Sensing*, 21, 435–456.

Lee, K., Jeon, S.H., and Kwon, B.D., 2004. Urban feature characterization using high-resolution satellite imagery: texture analysis approach, CD ROM *Proceeding of MAPASIA 2004 Conference*, Beijing, China.

Lu, D. and Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, 28 (5), 823–870.

Mather, P.M., 1999. *Computer processing of remotely-sensed images: an introduction*. 2nd ed. Chichester: John Wiley & Sons.

Munechika, C.K., *et al*., 1993. Resolution enhancement of multispectral image data to improve classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 59, 67–72.

Pajares, G. and Cruz, J.M., 2004. A wavelet-based image fusion. *Pattern Recognition*, 37 (9), 1855–1872.

Pohl, C. and van Genderen, J.L., 1998. Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, 19, 823–854.

Ricchetti, E., 2001. Visible-infrared and radar imagery fusion for geological application: a new approach using DEM and sun-illumination model. *International Journal of Remote Sensing*, 22, 2219–2230.

Richards, J.A. and Xia, S., 1999. *Remote sensing digital image analysis – an introduction*. 3rd ed. Berlin: Springer-Verlag.

Serkan, M., *et al*., 2008. Edge and fine detail preservation in SAR images through speckle reduction with an adaptive mean filter. *International Journal of Remote Sensing*, 29 (23), 6727–6738.

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, 562–578.

Solberg, A.H.S., Taxt, T., and Jain, A.K., 1996. A Markov random field model for classification of multi-source satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 34, 100–112.

Ulaby, F.T., *et al*., 1986. Textural information in SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 24, 235–245.

van Genderen, J.L., 1989. High resolution satellite data for urban monitoring. *International Journal of Remote Sensing*, 10, 257–258.

Vrabel, J., 1996. Multispectral imagery band sharpening study. *Photogrammetric Engineering and Remote Sensing*, 62, 1075–1083.

Wang, Y., Koopmans, B.N., and Pohl, C., 1995. The (1995) flood in the Netherlands monitored from space – a multisensor approach. *International Journal of Remote Sensing*, 16, 2735–2739.

Yesou, H., *et al*., 1993. Merging Seasat and SPOT imagery for the study of geologic structures in a temperate agricultural region. *Remote Sensing of Environment*, 43, 265–280.

Zeng,Y., *et al*., 2010. Image fusion for land cover change detection. *International Journal of Image and Data Fusion* (accepted).