INT. J. REMOTE SENSING, 2002, VOL. 23, NO. 4, 767–776

Taylor & Francis
Taylor & Francis Group

Texture analysis and data fusion in the extraction of topographic objects from satellite imagery

J. B. K. KIEMA

Institute for Photogrammetry and Remote Sensing (IPF), University of Karlsruhe, Englerstrasse 7, 76128 Karlsruhe, Germany;
e-mail: kiema@ipf.uni-karlsruhe.de

(*Received 5 July 2000; in final form 17 November 2000*)

Abstract. This paper examines the influence of multisensor data fusion on the automatic extraction of topographic objects from SPOT panchromatic imagery. The suitability of various grey level co-occurrence based texture measures, as well as different pixel windows is also investigated. It is observed that best results are obtained with a 3×3 pixel window and the texture measure homogeneity. The synthetic texture image derived together with a Landsat Thematic Mapper (TM) imagery are then fused to the SPOT data using the additional channel concept. The object feature base is expanded to include both spectral and spatial features. A maximum likelihood classification approach is then applied. It is demonstrated that the segmentation of topographic objects is significantly improved by fusing the multispectral and texture information.

1. Introduction

Conventional methods used in the classification of multispectral imagery basically employ the image spectral signature. This is largely acceptable in the segmentation of spectrally homogeneous object classes, since it is possible to delineate fairly clean and representative training areas. However, results obtained from such methods are unsatisfactory, particularly in the case of applications involving the mapping of man-made structures and other heterogeneous features in complex urban scenes. In general, these results are often characterized by limited accuracy and low reliability (Haala and Brenner 1999).

This is basically because the potential of spectral information is limited in such applications since urban objects are distinguished better through their spatial rather than their spectral reflection properties (Zhang 1999). Moreover, it is not possible to discriminate between object features that display high spectral overlap, for example, building roofs from pavements that are constructed using similar materials or trees from grass-covered areas.

The above conventional approach is further constrained in the case of remotely sensed data that has limited spectral potential. A classical example here is SPOT panchromatic imagery. In the case of such data, it is important to enhance the image value to support the segmentation of various object features. In principle, this may be achieved by either analysing the particular data with a view to expanding the object feature space, or by fusing this with some other external data.

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2002 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160010026005

Against the above background, the main objectives of the study presented here are twofold. First, to investigate the appropriateness of various grey level co-occurrence based texture measures, as well as different pixel windows in the extraction of topographic objects. Secondly, to consider within the perspective of data fusion, the automatic classification of SPOT panchromatic image data.

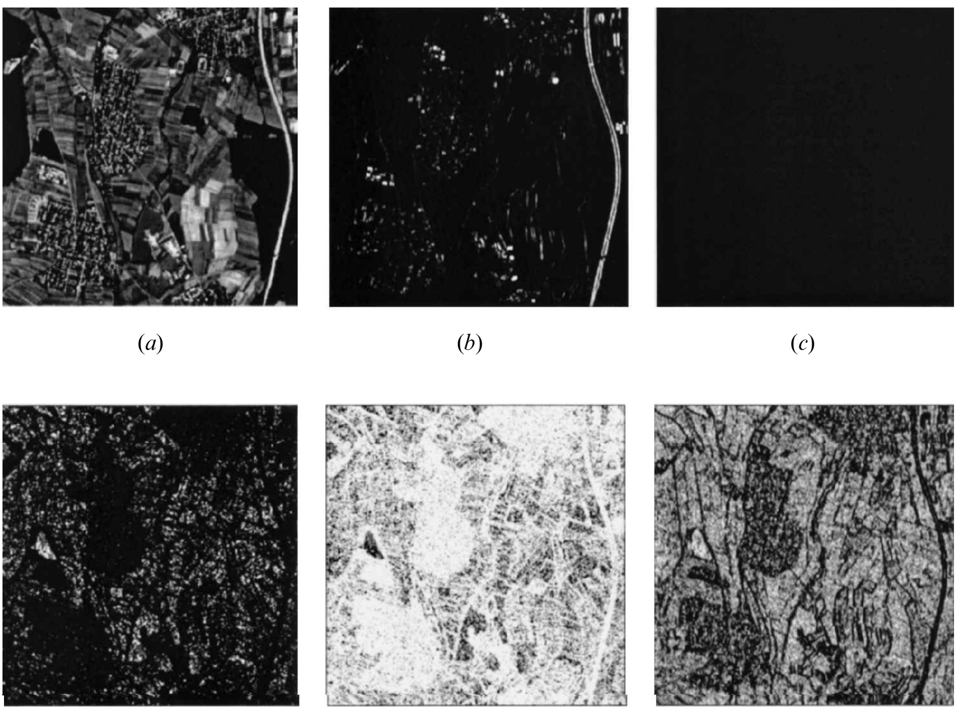
2. Test area and data employed

As shown in figure 1(*a*), the selected test area for this study covers a SPOT panchromatic scene for the area Stutensee–Karlsdorf. This is located about 16 km north-east of the city of Karlsruhe and lies in the south-western part of the Federal Republic of Germany, near the border with France. A Landsat Thematic Mapper (TM) image encompassing the same test area is also employed as secondary data.

3. Extraction of topographic objects

In the past, the multispectral classification of urban environments has achieved only limited success. Some of the reasons for this include the fact that:

- Urban environments are characterized by many different (sub)-objects (e.g. buildings, streets, gardens, water fountains) that exhibit a diverse range of spectral reflectance values.
- With the exception of the new generation high-resolution commercial satellites,



(*a*) (*b*) (*c*)

(*d*) (*e*) (*f*)

Figure 1. Texture analysis with different co-occurrence-based texture measures. (*a*) Original SPOT image, (*b*) contrast, (*c*) correlation, (*d*) energy, (*e*) entropy, (*f*) homogeneity.

the spatial resolution of most previous satellite data has been too coarse to enable effective interpretation of urban objects.
- The digital classification of spatial phenomena, especially in the case of urban environments, defines a relatively arduous task.

In applications where high quality segmentation results are demanded, or high-resolution imagery is employed, it is often necessary to enhance the object feature base. This needs to be expanded to incorporate spectral, spatial, as well as context features. Table 1 highlights some of the feature characteristics that need to be considered within the context of an expanded object feature base.

The feature characteristics that may be applied for a particular segmentation case are influenced by several factors including; the geometric and spectral resolution of the remotely sensed data, the particular object feature(s) of interest, the level of segmentation aspired, among other factors. In this study, both spectral and spatial features are adopted. However, for the purposes of this presentation, and because of the very nature of the SPOT panchromatic image, more emphasis is focused on texture as a spatial feature characteristic.

4. Data fusion

4.1. *Concept of data fusion*

The quantity of geospatial data has continued to grow significantly over the last couple of years. This can be attributed in part, to the increase in the number of spaceborne missions devoted to the observation of the Earth and other planetary bodies. Faced with the prospect of an ever-increasing amount of multisource data, it is important that the best information for a particular application be extracted from the available data. In this regard, data fusion is defined as a formal framework in which are expressed means and tools for the alliance of data originating from different sources (Wald 1998). Conceptually, this aims at obtaining information of greater quality.

The phrase 'greater quality' is used here in a generic sense and hence, its exact meaning will vary depending on the application under consideration. This may denote for instance, an increase in accuracy, or in the production of more relevant information. Within the context of the above definition, spectral channels of the same sensor, as well as images captured at different instants of time are considered as constituting different sources of data.

Despite the fact that the concept of data fusion is relatively easy to understand, its exact meaning and use often varies from one scientist to another. In principle, this may be performed at different levels for example, at measurement, attribute, rule or even decision levels. Different approaches to this may theoretically be employed in the fusion of multisource data including, RGB colour composites, IHS (Intensity–

Table 1. Expanded object feature base (Schilling and Vögtle 1996).

Feature domain	Feature characteristic
Spectral	spectral signature
Spatial	texture
	structure/form
	size
	shape/contour
	topology

Hue–Saturation) transformation, multiresolution analysis, among others (Pohl 1999). The particular method adopted is influenced by several factors including, the type of application under study, the structure of the data to be fused, as well as the image characteristics that need to be enhanced or preserved.

4.2. *Fusing the different data sources*

The fundamental ideas behind the combination of multisource data for scene labelling are outlined in Hahn and Stätter (1998). In the application considered in this study, Landsat TM imagery is fused with a SPOT panchromatic scene, to which several synthetic texture imagery are also synergized. In this regard, several different approaches to data fusion may be adopted. For instance, it is possible to make use of the hierarchical classification approach, the additional channel concept or even a knowledge-based strategy.

The hierarchical or layered classification approach is essentially a structured technique through which the different datasets to be fused are applied in such a way as to successively divide the working area into more detailed object classes (Savian and Landgrebe 1991). In principle, this begins with basic object classes before progressively zeroing in on more detailed ones. On the other hand, the different data sources are introduced as separate channels in an integrated fashion, within an expanded data framework for the additional channel method (Haala and Brenner 1999). This, of course, necessitates the co-registration of the different data to a uniform georeferenced system.

For the knowledge-based procedure, the data to be fused needs to be appropriately modelled and structured with respect to the particular knowledge representation adopted. Different formalisms for the representation of knowledge have been identified including, predicate logic, rule-based systems, semantic networks, among others (Kunz *et al.* 1997). An example of the use of semantic networks in the interpretation of digital aerial photographs using a map-supported approach is discussed in Quint (1997).

A comparative assessment of the above methods identifies the main disadvantage of the hierarchical classification approach as the propagation of the classification errors in the subsequent classification steps. Conversely, the main advantage of the additional channel procedure is its simplicity, as well as the enhanced flexibility in the data processing. Further, although knowledge-based methods for the integration of multisource data are fairly rigorous, these are nevertheless relatively complicated as the entire data system needs to be restructured accordingly. This can be a fairly complex task. From the foregoing, the additional channel concept is adopted for the data fusion in this study.

5. Co-occurrence matrix based texture analysis

Within a broad sense, texture can be defined as the spatial variation of the grey level in an image. Several methods have been developed to describe, classify and segment texture (Rosenfeld 1998). In general, it is possible to distinguish between the regular texture manifested by man-made objects from the irregular manner that natural objects exhibit texture. Hence, the texture characteristic can be used to discriminate between different objects and therefore, support their segmentation from remotely sensed data.

For instance, Busch (1998) describes a method through which built-up areas can be extracted using the high spatial density of short linear features in panchromatic satellite imagery. Built-up and non built-up areas are separated by simply adopting a threshold. This is estimated from the calculated feature density in training sites stored in an existing Geographical Information System (GIS). In addition, texture can also be employed to discriminate between spectrally similar object categories (Schilling and Vögtle 1996). In this regard, it is possible to distinguish between the object classes settlement and agriculture.

Nonetheless, simple statistical approaches using mean or standard deviation do not fully take into account the spatial distribution of pixels. These methods are therefore inadequate for most practical image texture analysis. Moreover, in contrast to artificial objects, it is not always easy to model the texture characteristic for different naturally occurring object features. In order to incorporate both the spectral, as well as the spatial distribution of image grey values, use is often made of the grey level co-occurrence matrices described in Haralick and Shapiro (1992) and Bässmann and Besslich (1993).

In principle, both the conventional texture analysis and the grey level co-occurrence matrix (GLCM) methods describe the grey value relationships in the neighbourhood of the current pixel. However, in the GLCM method, this is analysed within the GLCM space and not from the original grey values, as is the case in the former method. In this regard, the GLCM can be viewed as a two-dimensional histogram of the frequency with which pairs of grey level pixels occur in a given spatial relationship, defined by a specific inter-pixel distance and a given pixel orientation. Hence, in the segmentation of urban objects, texture analysis is usually performed within a GLCM matrix space as outlined previously (Steinnocher 1997, Zhang 1999, Kourgli and Belhadj-Aissa 2000). This is basically because of the significant role that the context parameter orientation plays in the spatial delineation of urban objects.

In this study, the suitability of different GLCM-based texture measures and varying pixel windows in the segmentation of topographic objects is examined. The different texture measures investigated are expressed as follows:

$$contrast = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} [g(i) - g(j)]^2 w(i,j) \tag{1}$$

$$correlation = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \left[\frac{(i-\mu)(j-\mu) w(i,j)}{\sigma^2} \right] \tag{2}$$

$$energy = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} w(i,j) \tag{3}$$

$$entropy = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} w(i,j) \log[w(i,j)] \tag{4}$$

$$homogeneity = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \left\{ \frac{w(i,j)}{1 + [g(i) - g(j)]^2} \right\} \tag{5}$$

where $w(i,j)$ are elements of the co-occurrence matrix space; $g(i)$, $g(j)$ are the row and column values of the co-occurrence matrix, respectively; $\mu = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} i w(i,j)$; $n \times m$ is the dimension of the co-occurrence matrix. Figure 1 compares the original SPOT panchromatic image with several GLCM-based texture measure imagery acquired using a unit inter-pixel distance for a 0° angular relationship between neighbouring pixels, denoted in this case by P_0, and a window size of (3×3) pixels. From the different texture measures examined, the worst results are obtained using the parameter *correlation*. The texture measure *contrast* is clearly able to filter out the outline of the major road in this particular SPOT scene. However, most of the other object features are hardly interpretable.

On the other hand, the texture measures *energy* and *entropy* are essentially able to filter out the homogeneous object features like vegetation cover, with the logarithmic parameter *entropy* clearly capturing this more accurately. In addition, the texture measure *energy* depicts, more or less, a 'salt and pepper' effect on the filtered objects. Nevertheless, the performance of these two texture parameters on heterogeneous object features such as buildings is comparatively poor.

With regard to the overall contours of the different topographic objects, it is observed that best results are obtained with a (3×3) pixel window and the texture measure *homogeneity*. This is in line with previous observations as noted in Zhang (1999). Clearly, the texture measure *homogeneity* is able to filter out both the homogeneous and heterogeneous objects. However, this has the drawback of filtering out virtually all objects and could therefore introduce noise in the filtered data.

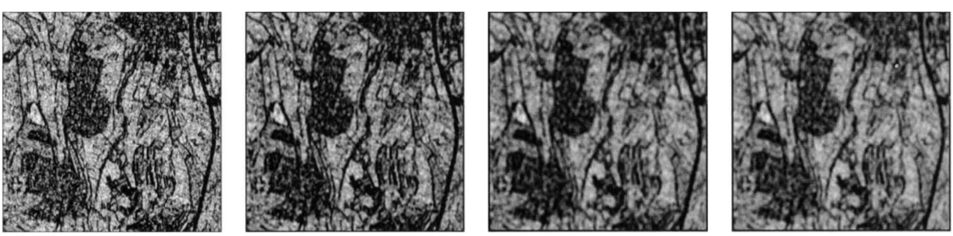
The analysis of different pixel window sizes reveals that the effect of increasing the pixel window size, for instance, from (3×3) pixels to (5×5) pixels, is to propagate the overall smoothing of the image. In principle, this increases the maximum distance to a neighbouring pixel, for example, from 1 to 2 in both image directions. Thus, pixels that are not immediate neighbours of a particular image pixel end up influencing the value of that pixel. This results through some 'averaging operation' depending on the particular texture measure adopted. The effect of this smoothing phenomenon is demonstrated in figure 2 for the texture measure *homogeneity* and pixel windows varying between 3×3 and 9×9.

6. Classification of SPOT panchromatic imagery

The segmentation of object features from SPOT panchromatic imagery is greatly constrained by the fact that this data contains only one panchromatic channel. This significantly limits the ability to distinguish between most natural and artificial object features using this data source. Figure 3 clearly highlights this limitation for a typical maximum likelihood classification of SPOT panchromatic imagery.

In order to enhance the value of this data, particularly in the segmentation of urban objects, two different approaches may be adopted. Firstly, texture analysis may be performed on the panchromatic data. As mentioned above, this is normally performed within the GLCM space in practice. Alternatively, the panchromatic data may be fused with other multispectral data (e.g. SPOT XS, Landsat TM, etc.). The idea behind this is to enhance the image spectral value as outlined previously (Carper *et al.* 1990, Yesou 1993).

To support the segmentation process 'synthetic' texture images are generated for a



(a) 3×3 (b) 5×5 (c) 7×7 (d) 9×9

Figure 2. Effect of pixel window size on the texture measure *homogeneity*.

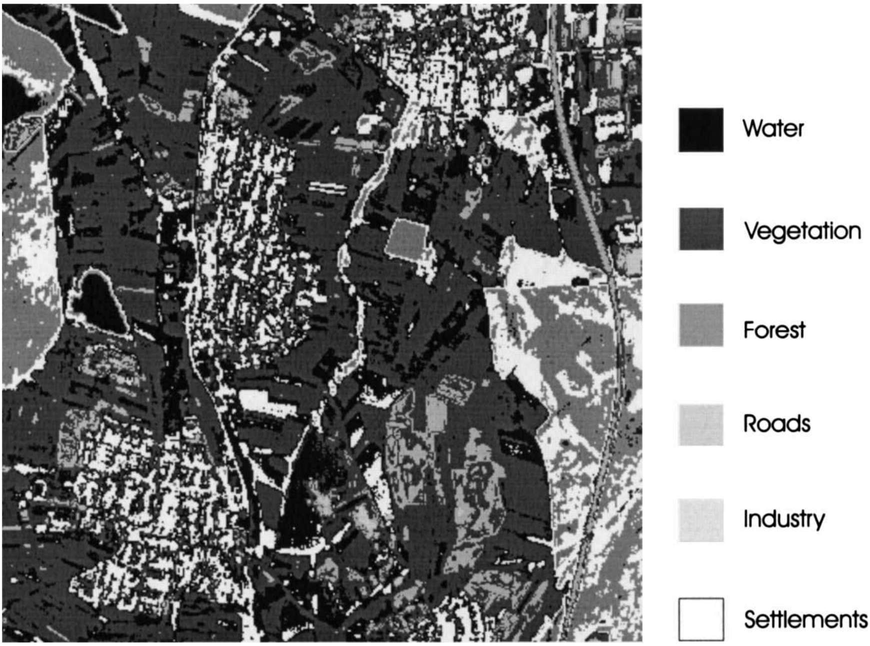


Figure 3. Classification results for panchromatic data alone.

window of 3×3 pixels using the texture measure homogeneity. The four principle co-occurrence orientations P_0, P_{45}, P_{90} and P_{135} and are employed. Resulting texture imagery are then fused with the panchromatic channel using the additional channel concept.

A simple maximum likelihood classification as opposed to the more complicated contextual segmentation methods is then carried out. The selection of the training data is basically done using manual digitizing. Six basic urban object classes are identified: Settlement, Vegetation, Industry, Water, Forest and Roads. Figure 4 illustrates the classification results obtained for the fused data.

A Landsat TM scene (bands 3, 4 and 7) encompassing the same area is then re-sampled to a pixel size of 10 m. Since both the Landsat TM and SPOT panchromatic imagery are already georeferenced, an affine transformation is used to transform the re-sampled Landsat TM imagery to the co-ordinate system of the SPOT panchromatic data. The transformed imagery is then fused with the SPOT panchromatic image through the additional channel concept. To this integrated image is also fused the texture measure *homogeneity* (P_0). Classification results obtained for this dataset are shown in figure 5.

7. Analysis and discussion of results

A comparison of figures 3 and 4 confirms the importance of incorporating texture information in distinguishing man-made urban objects (e.g. settlements and major roads) from natural object features (e.g. vegetation and forests). However, it is still not possible to effectively interpret certain object features. For instance, it is not possible to recognize the lake in the southern part of the study area. This problem is solved after introducing the multispectral information through the Landsat TM imagery as shown in figure 5.

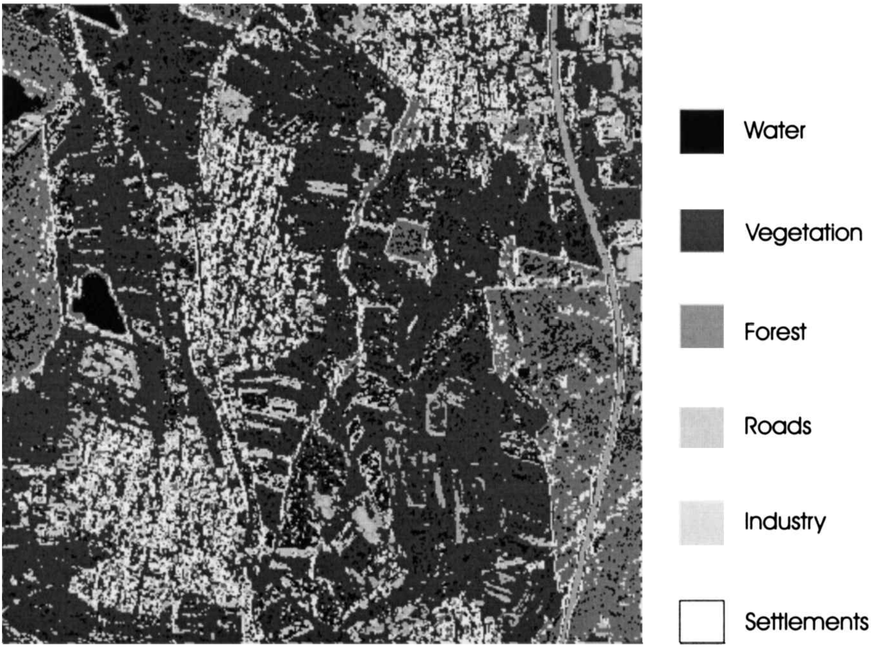


Figure 4. Classification results for panchromatic and texture data.

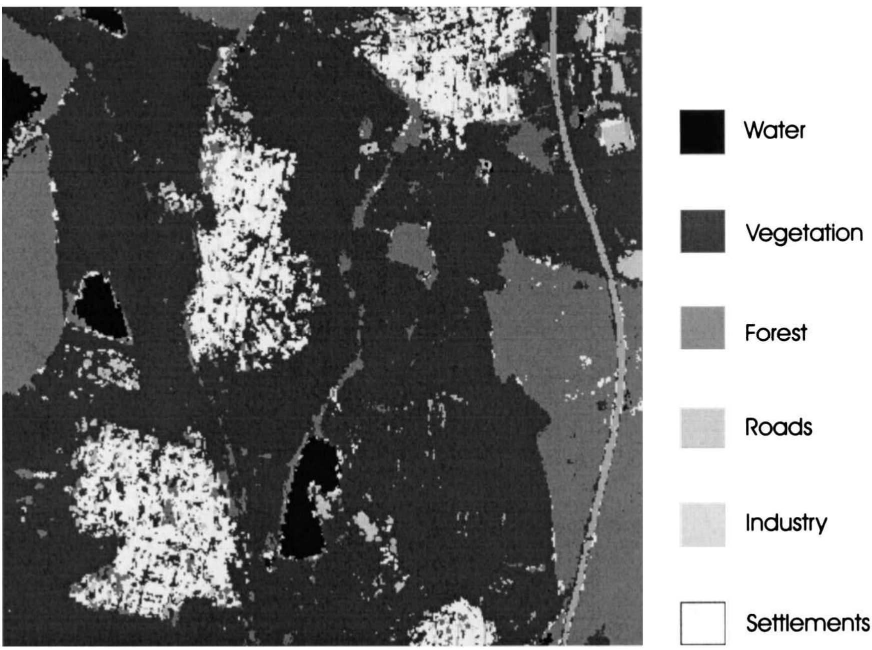


Figure 5. Classification results for panchromatic, multispectral and texture data.

Normally, after classifying an image some quality measure is required in order to allow a degree of confidence to be attached to the results. This also serves to indicate whether the analysis objectives have been realized (Richards 1993). Different quality aspects may be examined in the evaluation of spatial data quality depending on the particular analysis domain considered. This may vary from spatial, temporal to thematic domains and may involve the assessment of different quality aspects, including accuracy, resolution, completeness and consistency.

This section reviews the thematic accuracy of the final classification results obtained. In general, different approaches to the assessment of this can be distinguished including, error matrix-based methods, spectral distance-based methods and quantitative methods. Because of its popularity, the error matrix method is adopted here.

Ground truth methods are employed as the reference basis for the evaluation. Stratified random sampling procedures are used to estimate the error (confusion) matrix (Congalton *et al.* 1983, Congalton 1991). Various Kappa related measures, as well as accuracy parameters derived from Kullback–Leibler information on multinormal distributions described in Nishii and Tanaka (1999) are then estimated as shown in table 2. In general, results obtained confirm that a fairly good image classification is finally realized.

Basically, the Kappa statistic together with its related derivative parameters (class-averaged accuracy and overall accuracy) are evaluated using only the diagonal elements, as well as the column and row summations of the error matrix. Thus, the off-diagonal elements of the error matrix are neglected. These elements may, however, be critical in certain applications. The Kullback–Leibler information includes these later elements and constitutes therefore, a more rigorous parameter for assessing the thematic classification accuracy. In particular, the parameters *Jpro* and *Juni* provide a measure for the degree of departure of the actual classification from complete or ideal classification. On the other hand, the parameters $Rpro(X|\tau)$ and $Runi(X|\tau)$ are coefficients for inaccuracy assessment which give an impression of the instances of object misclassification(s). In this particular example, these are estimated based on the assumption that all misclassifications are treated equally for the different object classes identified.

8. Summary and conclusions

This study underlines the importance of texture analysis and data fusion in the automatic extraction of topographic objects. In particular, the need to fuse multispectral data and textural indexes is underscored. The fusion of all the above information sources is necessary in order to facilitate the effective discrimination of the various man-made and natural objects found within urban neighbourhoods. For the texture analysis, it is noted that the best co-occurrence based texture measure in the extraction of topographic objects is *homogeneity*. Similarly, a pixel window of size (3×3) is best suited for this. The need to enhance the object feature base to support feature extraction is also highlighted.

Table 2. Thematic classification accuracy.

Kappa statistic	0.882
Class-averaged accuracy	0.878
Overall accuracy	0.883
Jpro	0.881
Juni	0.877
$Rpro(X\|\tau)$	0.117
$Runi(X\|\tau)$	0.122

Acknowledgments

The author wishes to acknowledge the support of Prof. Dr-Ing. H.-P. Bähr (IPF) and the German Academic Exchange Service (DAAD) in the preparation of this paper.

References

Bässmann, H., and Besslich, P. W., 1993, *Bildverarbeitung*(Berlin: Springler Verlag), pp. 117–130.
Busch, A., 1998, Revision of built-up areas in a GIS using satellite imagery and GIS data. *International Archives of Photogrammetry and Remote Sensing*, **32**, Part 4, Commission IV, pp. 91–98.
Carper, W. J., Lillesand, T. M., and Kiefer, R. W., 1990, The use of Intensity–Hue–Saturation transformations for merging SPOT panchromatic and multispectral image data. *Photogrammetric Engineering and Remote Sensing*, **56**, 459–467.
Congalton, R. G., Oderwald, R. G., and Mead, R. A., 1983, Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric Engineering and Remote Sensing*, **49**, 1671–1678.
Congalton, R. G., 1991, A review of assessing the accuracy of classifications of remotely sensed images. *Remote Sensing of Environment*, **37**, 35–46.
Haala, N., and Brenner, C., 1999, Extraction of Buildings and Trees in Urban Environments, *Photogrammetric Engineering and Remote Sensing*, **54**, 130–137.
Hahn, M., and Stätter, C., 1998, A scene labelling strategy for terrain feature extraction using multisource data. *International Archives of Photogrammetry and Remote Sensing*, **32**, Part 3/1, 435–441.
Haralick, R., and Shapiro, L., 1992, *Computer and Robot Vision, 1* (Reading Massachusetts: Addison-Wesley), pp. 157–255.
Kourgli, A., and Belhadj-Aissa, A., 2000, textural classification using textural signatures. In *Proceedings of the European Association of Remote Sensing Laboratories (EARSeL) Symposium: Remote Sensing in the 21st Century, Economic and Environmental Applications*, edited by J. L. Casanova (Rotterdam: A. A. Balkema), pp. 557–561.
Kunz, D., Schilling, K.-J., and Vögtle, T., 1998, A new approach for satellite image analysis by means of a semantic network. In *Semantic Modelling and Extraction of Spatial Objects from Images and Maps (SMATI97)*, edited by W. Förstner and L. Plümer, (Basel, Switzerland: Birkhäuser Verlag), pp. 20–36.
Nishii, R., and Tanaka, S., 1999, Accuracy and inaccuracy assessment in land-cover classification. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, 491–498.
Pohl, C., 1999, Tools and methods for fusion of images of different spatial resolutions. *International Archives of Photogrammetry and Remote Sensing*, **32**, Part 7-4-3, 76–82.
Quint, F., 1997, Kartengestuetzen Interpretation monokuler Luftbilder, Dissertation, Deutsche Geodaetishe Kommission, Reihe C, Heft Nr. 477, Muenchen.
Richards, B., 1993, Remote Sensing and Digital Image Analysis (Berlin: Springer-Verlag), pp. 271–275.
Rosenfeld, A., 1998, Image analysis and computer vision: 1997. *Computer Vision and Image Understanding*, **70**, 239–284.
Savian, F., and Landgrebe, D., 1991, A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man and Cybernetics*, **21**, 660–674.
Schilling, K. -J., and Vögtle, T., 1996, Satellite image analysis using integrated knowledge processing. *International Archives of Photogrammetry and Remote Sensing*, XVIII Congress of the ISPRS, Part B3, Commission III, 752–757.
Steinnocher, K., 1997, Texturanalyse zur Detektion von Siedlungsgebieten in hochauflösenden panchromatischen Satellitendilddaten. In Salzburger Geographische Materialien. Tagungsband zur AGIT 1997, 26, pp. 143–152.
Wald, L., 1998, A European proposal for terms of reference in data fusion. *International Archives of Photogrammetry and Remote Sensing*, **32**, Part 4, Commission VII, pp. 651–654.
Yesou, H., Besnus, Y., and Rolet, J., 1993, Extraction of spectral information from Landsat TM data and merger with SPOT panchromatic imagery—a contribution to the study of geological structures. *ISPRS Journal of Photogrammetry and Remote Sensing*, **48**, 23–36.
Zhang, Y., 1999, Optimisation of building detection in satellite images by combining multi-spectral classification and texture filtering. *ISPRS Journal of Photogrammetry and Remote Sensing*, **54**, 50–60.