URBAN LAND USE CLASSES WITH FUZZY MEMBERSHIP AND CLASSIFICATION BASED ON INTEGRATION OF REMOTE SENSING AND GIS

Qingming Zhan, Martien Molenaar, Ben Gorte
International Institute for Aerospace Survey and Earth Sciences (ITC)
P. O. Box 6, 7500 AA Enschede, The Netherlands
Zhan@itc.nl, Molenaar@itc.nl, Gorte@itc.nl

Working Group IV/III.2

KEY WORDS: Land Use, Classification, Fuzzy Modeling, Image Analysis, GIS

ABSTRACT

Urban land use classification from remotely sensed images has drawn great attention in the past decades. Most researchers derive land use data from remotely sensed images alone, but the results are not quite satisfying for detecting detailed land use classes in urban areas. Fuzzy urban land use classes proposed here consist of a number of fuzzy memberships that offer direct links to findings from remote sensed images and GIS data. When we compare those indicators derived from remote sensed images and GIS data with the fuzzy class memberships using fuzzy criteria or rules, we would be able to evaluate the possibilities of fuzzy membership functions of an area with respect to predefined classes. The indicators will be developed for the case that such a classification should be based on the combination of remotely sensed images and GIS data. Rules and parameters will be presented for classification with the related uncertainty levels. The final result of such a classification process will consist of a land use map and a corresponding map indicating uncertainty levels of the assignment of area to the relevant classes. The proposed approach will be based on estimating the real cover of an area by features like green space, water body, built up area etc. From these cover compositions the major functions will be inferred in a land use map.

1 INTRODUCTION

Urban land use data play a most important role in urban planning and management. They require frequent and timely updating to support decision making in planning offices at various levels. Urban land classification from remote sensed images has drawn great attention in the past decades. Most researchers derive land use data from remote sensed images alone, but the results are not quite satisfying for detailed land use detection in urban areas. This is due to the nature of mixed features for which the classification is developed mainly according to social economic functionality. Features derived from remote sensed images are not directly corresponding to the required land use classes.

Early attempts to derive information of this kind often failed to produce the levels of accuracy and detail required for town planning purposes. At the time, this was ascribed to the relatively low spatial resolution of satellite sensors. However, the use of data from sensors with improved resolving power in more recent years has not always yielded the expected increases in classification accuracy. This has generally been referred to in literature as a problem of increasing 'scene noise' – i.e. increasing spatial heterogeneity in the spectral response of urban areas. In fact, the problem is more properly expressed in terms of the image analysis techniques used and, in particular, to the inappropriateness of standard, per-pixel parametric algorithms for segmenting images of urban areas. Simply stated, it is extremely difficult to derive consistent, representative training statistics for urban land use categories since these are comprised of many land cover types, each with different spectral reflectance properties (Townshend 1980, Barnsley *et al.* 1991, Barr 1992, Zhang 1999).

Alternative approaches to urban land-use mapping have been explored. The most successful of these make use of other sources of spatially-referenced ancillary data (Sadler and Barnsley 1990, Sadler *et al.*1991) or image texture measures (Sadler *et al.* 1991), Baraldi and Parmiggiani 1990). The ancillary data set normally forms an additional data plane in the standard algorithms. Increased classification accuracy has been reported using both of these techniques (Sadler and Barnsley 1990, Sadler *et al.* 1991, Gorte 1999), although improvement has not been consistent across all land use categories.

The fundamental problem involved in producing accurate land use maps of urban areas arises form the fact that urban areas are complex assemblages of a disparate set of land cover types - including man-made structures, vegetation types and water bodies – each of which has different spectral reflectance characteristics. In visual analyses of remotely sensed images, particularly aerial photography, the spatial pattern of these land cover types is often used to distinguish between categories of urban land use. For example, residential areas can often be recognized by their particular mixture of buildings, roads, grass and trees; by contrast, parkland is primarily composed of grass and trees (Barr 1992, Tonjes 1999).

Recently, several studies have attempted to use the spatial mixing of land cover types within urban areas as a means of mapping land use. The studies by Wharton (1982), Whitehouse (1990) and Barnsley *et al.* (1991) have utilized various forms of (per-pixel) spatial re-classification techniques applied to an initial (land cover) segmentation of urban areas. The fundamental basis of these techniques is that it is possible to obtain some measure of the density and distribution of land cover types that is characteristic of a particular urban land use (Barnsley *et al.* 1991). Areas of similar land use can therefore be delineated by grouping pixels with different class (land cover) labels on the basis of these measures. All three studies attempt this through the use of a type of convolution kernel which either sums the density distributions of the constituent cover types (Wharton 1982), whitehouse 1990) or measures their spatial arrangement (Barnsley *et al.* 1991) within the kernel.

Although promising results have been obtained using these algorithms, the use of a pre-defined kernel places an undesirable restriction on the nature of the spatial searching employed. In particular, it is doubtful whether a single kernel of any size can adequately characterize the complex spatial distribution of the cover types contained in all of the land use categories likely to be found within a typical urban scene (Barr 1992, Tonjes 1999).

Besides using pixel-based approaches for image segmentation, we could use road and other linear features that can be derived from GIS data in image segmentation to support land use classification. The conventional pixel-based approaches for land use classification can be applied in a first step. The results will be used as indicators to determine the possibilities that an area, which is surrounding by linear feature from GIS data, can be classified as a certain class. Further measurements will be made based on types of land cover, derived by a pixel-based approach, as well as the proportions and compositions of each type. Since a number of those indicators are not fit the normal distribution, a fuzzy approach can be deployed to assess classification reliability (Molenaar 1996, Hootsmans 1996, Cheng 1999a).

In this case, it is necessary to define fuzzy membership functions for each land use class. Fuzzy criteria or rules have to be set up based on sample data sets. In principle an area will be assigned to a class with the highest membership function value. Those fuzzy membership function values derived from different sensors (different resolutions) will be checked for their consistency. In addition, it is possible to incorporate indicators that are not derived from images, such as location, spatial relation etc. For instance, an area, that has high possibility values for both commercial use and residential use, Based on its compositional, proportional or spectral characteristics, could be classified as ‘commercial’ if it is located near the city center or near the main streets. It could be classified as ‘residential’, as major function, if it is located in a place surrounding by residential areas and it could be classified as ‘commercial”, as secondary function. Uncertainty levels can be evaluated according to the degree that indicators fit the fuzzy membership functions (Hootsmans 1996, Smits *et al.* 1999, Cheng and Molenaar 1999b).

The final result of such a classification will consist of a land use map and a corresponding map indicating uncertainty levels of the assignment of each area to the relevant classes. The proposed approach will be based on estimating the real cover of an area by features like green space, water body, built up area etc. From these cover compositions the major functions will be inferred in a land use map. This approach offers the opportunity to adapt the decision rules for this inference to the local conditions of cities by adjusting the parameters based on local knowledge. It has potential value in monitoring land use changes and their trends by comparing major and secondary functions according to their fuzzy membership values and compositions during different periods.

2 HIERACHIES OF SPATIAL DATA

The required land use classification distinguishes between social-economic functions in addition to natural features. For instance, residential areas consist of certain types of buildings, footpaths, small gardens, public open space and facilities etc. Obviously, it is impossible to automatically delineate the boundary of residential area only with spectral information from remote sensing images, including large-scale aerial photographs. Therefore it is necessary to look for additional solutions.

It is believed that the traditional concept of the "thematic map" derived from the classification of satellite imagery requires considerable improvement, and that greater user participation at all stages is required (FLIERS, 1999). In additional to conventional image classification, following factors can be kept in consideration:

- Multi-sensor and multi-resolution images work together for in-depth image analysis in order to determine the boundaries of natural features as well as conceptual features such as residential, industrial etc. Quite some questions remain unsolved especially in determination and delineation of the boundaries of conceptual features like residential, industrial etc., which are often more important than the natural features in urban planning and management context (Tilton 1999).

- To use existing data GIS data and knowledge of urban pattern and form. Knowledge of urban pattern and form such as location, distance from the city center, spatial relation etc. can be useful in identification of land use. Relevant knowledge has to be extracted from existing data and expertise (Molenaar 1993, Timpf 1998). Proper methodologies for knowledge extraction and representation of existing data and expertise will be needed to enable their application in image interpretation.

- It is necessary to develop algorithms to assign areas to certain land use classes and to indicate the confidence or reliability of those decisions.

2.1 Multi-sensor (image resolution)

Fuzzy membership function derived from several images of different sensors can be checked for consistency, to increase confidence in classification reliability. Different images with different resolution can be selected on what detail is required and in what scale the land use map is produced (Fig.1). i.e. To classify an urban built-up area at map scale of 1:50,000 or smaller, TM or MSS images can be used. For major land use classes in urban built-up area at a scale of 1:10,000 or 1:25,000, SPOT image are needed. To classify more detailed land use classes in urban built-up area or to produce land use map in scale 1:5,000 or larger, higher resolution images have to be applied such as IKONOS images or aerial photographs.

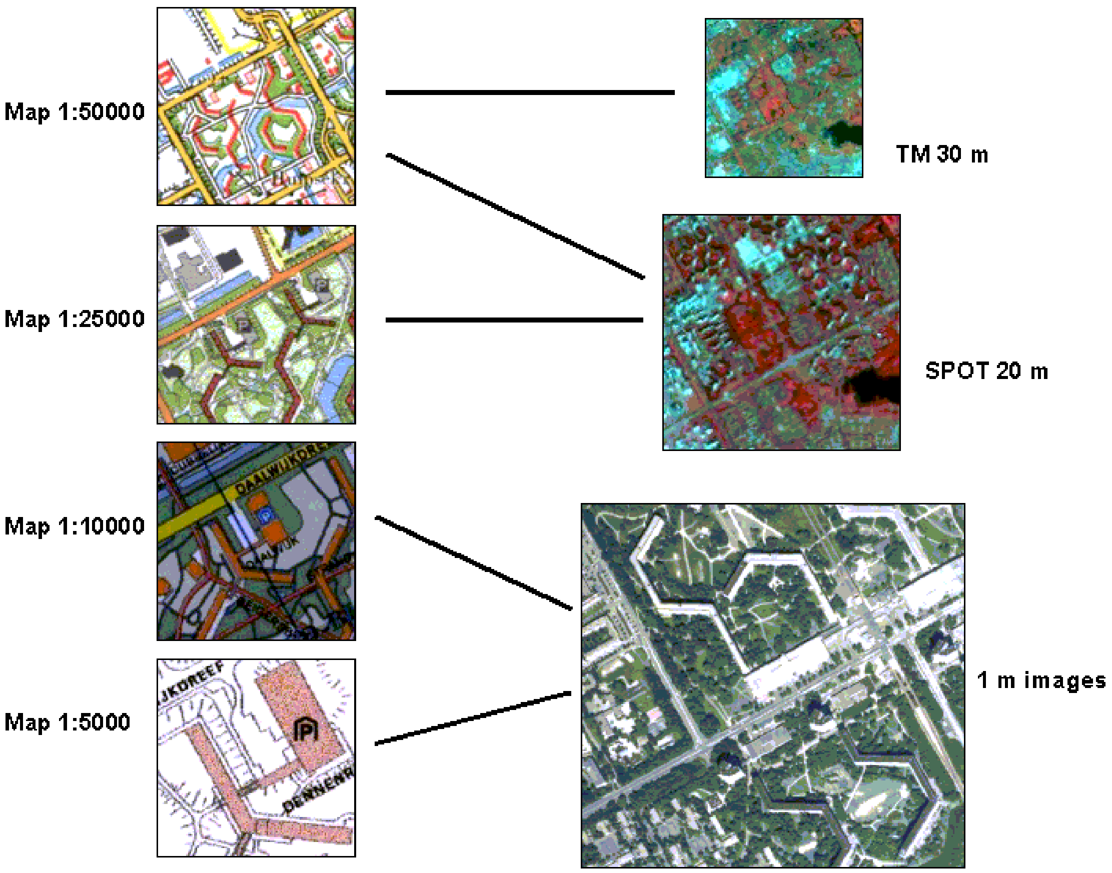


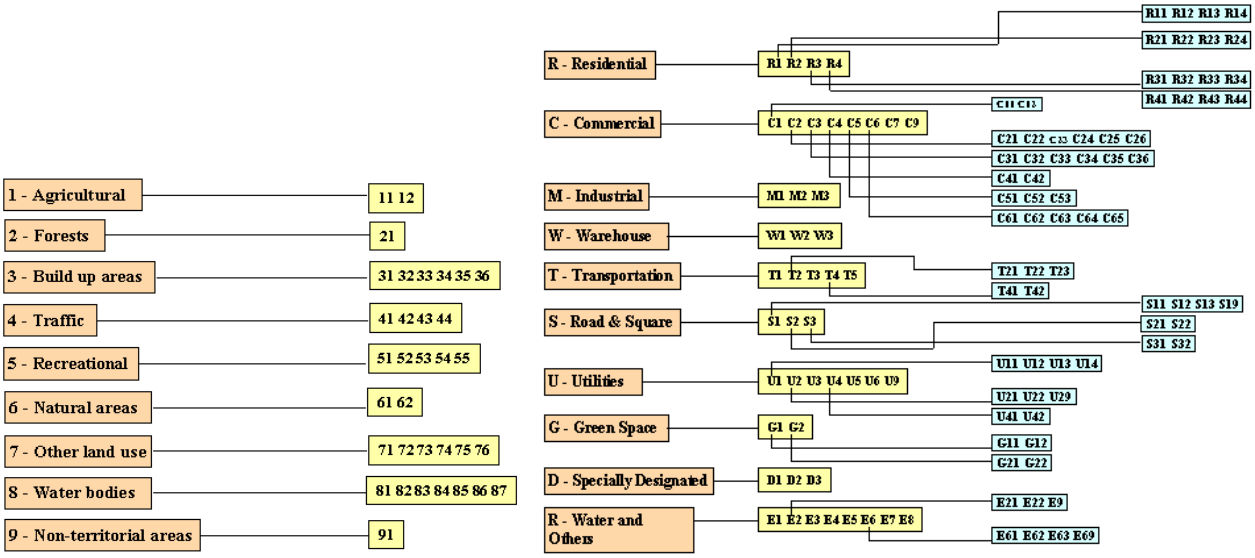
Fig. 1 Hierarchy of Spatial Data (Map scale and resolution) and their relations

2.2 Multi-scale (GIS Data)

In GIS database, graphic data input from various scale have different spatial extents and different detail levels (Fig.1). They can play an important role in land use classification from images different resolutions. Additional attention has been paid to linear features such as roads, railways and rivers, for which may guide image segmentation for later use in land use classification. In addition, hierarchies defined within linear features may support different levels of spatial data aggregation and image segmentation. On the other hand, it is believed that spectral features multi-resolution images can be applied to understand the spatial extent of those spatial features by checking land cover types surrounding them. Land cover information can be involved in hierarchical spatial data aggregation from large-scale map to smaller scale or less detail level. Image segmentation has been made based on various linear features from GIS data in the case study including roads, real ways, metro lines and other linear feature automatically and manually (Fig. 7).

2.3 Hierarchy of land use classes

Land use classes have been structured hierarchically according to their application fields and levels. For general use, urban built-up area and non-urban area will be classified. For town planning use, residential, commercial, industrial, road, park and other public green space etc. will have to be distinguished. For intensive analysis of urban land use study, however, more detail land use information will have to be explored. i.e. in residential area, residential houses, gardens, footpaths and facilities etc. will have to be indicated. Obviously, those land use classes have their hierarchical structure. Their hierarchical structure can be founded as well in a number of national standards (Fig. 2). Therefore we can link detail level of required land use classification in association with map scale as well as resolution of images to be applied in classification. Moreover, lower level of classes can be defined as membership functions for classification at higher levels.



2-1 Land Use Classification of the Netherlands **2-2 Urban Land Use Classification of China**

Fig. 2 Hierarchy of Land Use Classification

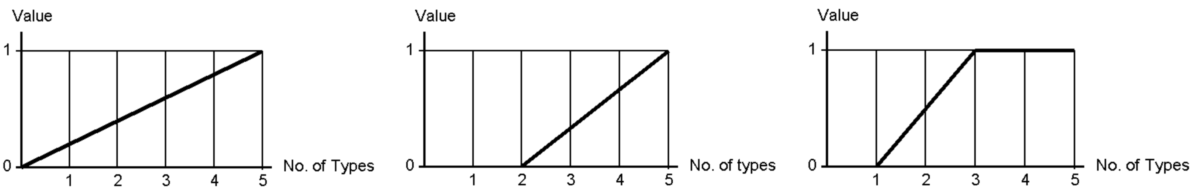
3 URBAN LAND USE CLASS WITH FUZZY MEMBERSHIP

The aim of a ‘hard’ classification approach is to place each pixel into a single land cover class. More recently it has been recognized the spatial varying character of land cover is better described by probability surfaces. Each pixel is allowed to have a 'class membership' probability rather than a single class label and the result of this operation as a 'soft' classification (Mather, 1999; Foody, 1999). Fuzzy classification offers a better choice in urban land use classification. It can indicate the primary and secondary land use functions at the same time. This offers more meaningful information for planners in better understanding of the land use patterns. It can be converted to crisp classification by simply labeling the primary functions to land use classes. On the other hand, it enables land use classification to be based on the degree of findings in matching with their memberships. Therefore it allows for additional indicators to be involved, in order to reach classification accuracy by accommodating those indicators as their membership.

Among new development of classifiers, the maximum likelihood classifier (ML) can reach high accuracy of image classification by incorporating prior knowledge using several probability estimations in image segmentation and classification (Gorte, 1998 and 1999). But it relies upon the assumption that the populations from which these training samples are drawn from a multivatiate normal. To increase or decrease the reliability of competing interpretations, structural relationships of the objects can be exploited (Tonjes *et al*., 1999). Handing objects with fuzzy spatial extent shows more possibilities in this regard (Cheng and Molenaar, 1999a, 199b). Class membership values can be used to estimate the area of different land-cover classes as well as to assess of the uncertainty of these area estimators (Canters, 1997). It incidentally precludes the use of GIS data.

3.1 Fuzzy Membership Functions

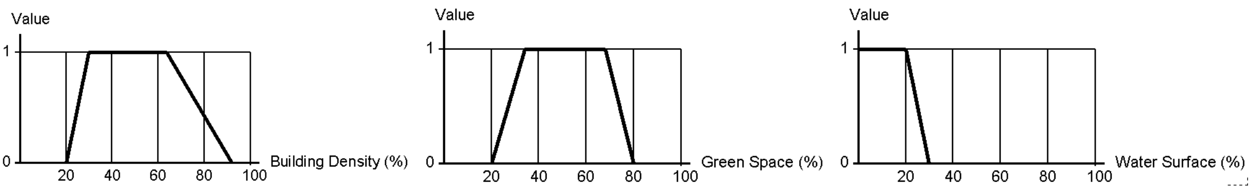
Urban land use classes can be defined according to their definition in urban planning and their actual situations. For instance, 'residential' consists of houses, green space, footpaths and small wa ter bodies etc. So the fuzzy membership function is defined according to a number of compulsory land cover types (Fig. 3). Three major membership functions can be applied according to different situations: 1) When a number of land cover types shows that they will play a more or less equally important role in determination of a land use class and the more land cover types can be found the higher possibility can be reach, then their membership function can be defined as Fig.3-1. 2) When one or two types of land cover are not compulsory to be found, then their membership function can be defined as Fig.3-2. 3) When one or two types of land cover are not compulsory to be found and it will be fine if we could find some of land types, then their membership function can be defined as Fig.3-3. Formulation of actual membership function will be based on actual situation of each land use class and knowledge. It will be better to define the membership functions according to sample data of study area since they might be different from place to place.



3-1 **3-2** **3-3**

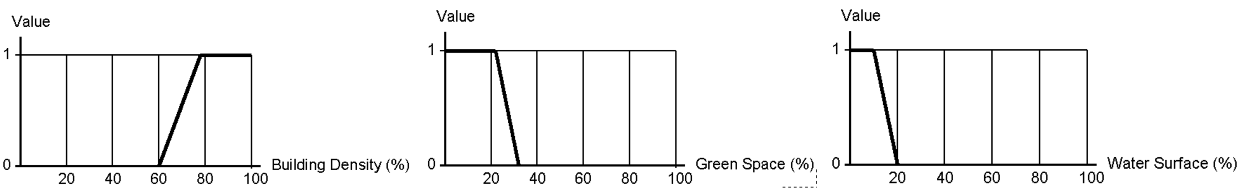
Fig. 3 Fuzzy membership function for compulsory land cover types
(i.e. Residential should include following types: Buildings, Green Space, Footpaths, Water Surface etc.)

In addition, to identify compulsory land cover types. It is necessary to check the proportional composition of each land cover type for each area for their potential land use classes. Their membership functions can be defined based on sample data set as well. The fuzzy membership functions for classifying as residential use are defined as Fig. 4. Their boundary parameters are derived from sample data set according to the mean and the standard deviation as well as from local knowledge. Fuzzy membership functions for commercial and office, parks and green space are showed in Figure 5 and Figure 6.



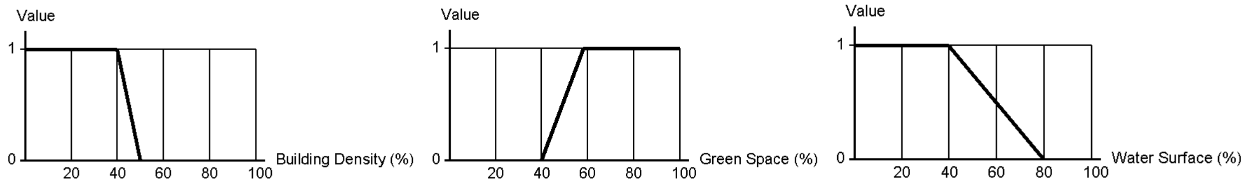
4-1 **4-2** **4-3**

Fig. 4 Fuzzy membership function for classifying as 'residential' use (Proportion)



5-1 **5-2** **5-3**

Fig. 5 Fuzzy membership function for classifying as 'commercial and office' use (Proportion)



6-1 **6-2** **6-3**

Fig. 6 Fuzzy membership function for classifying as 'parks and green space' (Proportion)

3.2 Fuzzy Land Use Classification

Fuzzy membership functions for a number of land cover types to determine of residential, commercial and green Space, are defined as $MF_{R\text{-}LC}$, $MF_{C\text{-}LC}$, $MF_{G\text{-}LC}$ and $MF_{R\text{-}P}$, $MF_{C\text{-}P}$, $MF_{G\text{-}P}$, Their fuzzy membership function values are determined according to their membership function derived from samples data set. The combined membership functions for classification are calculated using different weights based on expert knowledge for their importance in classification and local knowledge. i.e. $MF_R=W_iMF_{Ri}$. The possibility for classifying as certain class (P_i) is accessed by their membership function values for each land use class. Finally, the land use class for certain area is determined with the highest value among possibilities (P_i). i.e. $Class_i=Max_i(P_i)$. By sorting those possibilities, their secondary and third land use functions can be found as well.

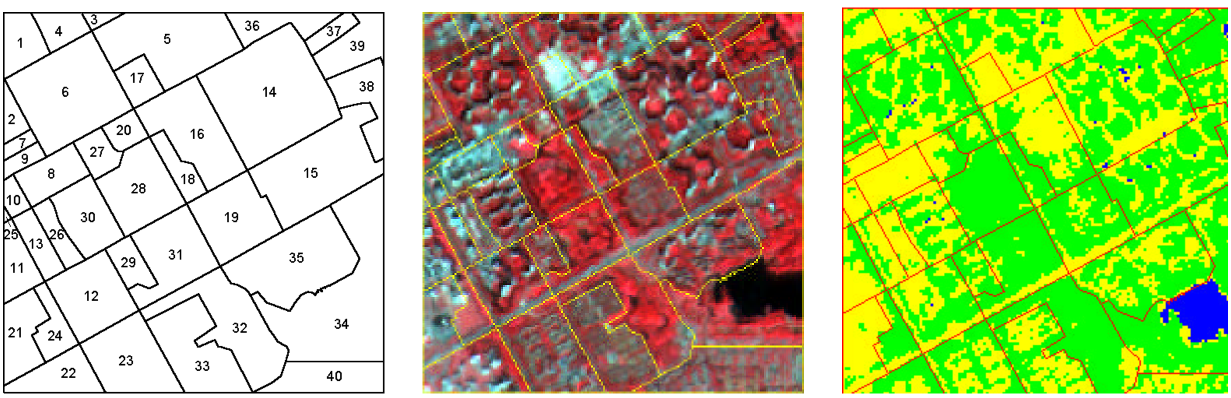
A number of rules can be applied in an early stage of classification i.e. if one or more of those member function values to class 'residential' get 0, then the area cab not be classified as class 'residential' in order to speed up the process.

$$\mathbf{MF_{Residential} = Min[MF_{R\text{-}Building}, MF_{R\text{-}Green\ Space}, MF_{R\text{-}Water}, \ldots] = 0}$$

$$\textbf{or} \quad \mathbf{MF_{Residential} = MF_{R\text{-}Building} \cdot MF_{R\text{-}Green\ Space} \cdot MF_{R\text{-}Water} \ldots = 0}$$

4 CASE STUDY

A study area of 9 square kilometers in the Bijlmer, Amsterdam was selected. Linear features such as roads, real ways, metro lines etc. have been derived from various scale maps in GIS (Fig. 7-1). They are used in image segmentation to find homogeneous areas for land use classification. A SPOT image of the study area has been used for land cover classification (Fig. 7-2). The maximum likelihood classifier (ML) has been applied in land cover classification based on SPOT image (Fig. 7-3). Land cover composition and their proportion have been derived from land cover classification based on SPOT image (Fig. 8). Based on the proposed approach, fuzzy urban land use classification has been implemented. The output land use map by fuzzy approach is produced as Figure 9 and the map for uncertainty assessment is produced as well in Figure 10. The land use map produced by air photo interpretation is shown in Figure 12. Figure 11 shows the adjusted land use map for criteria adjustment to change the secondary land use function to the major land use function.



7-1 Linear Features & Polygons **7-2 Image Segmentation** **7-3 Land Cover Classification**

Fig. 7 Image Segmentation and Land Cover Classification

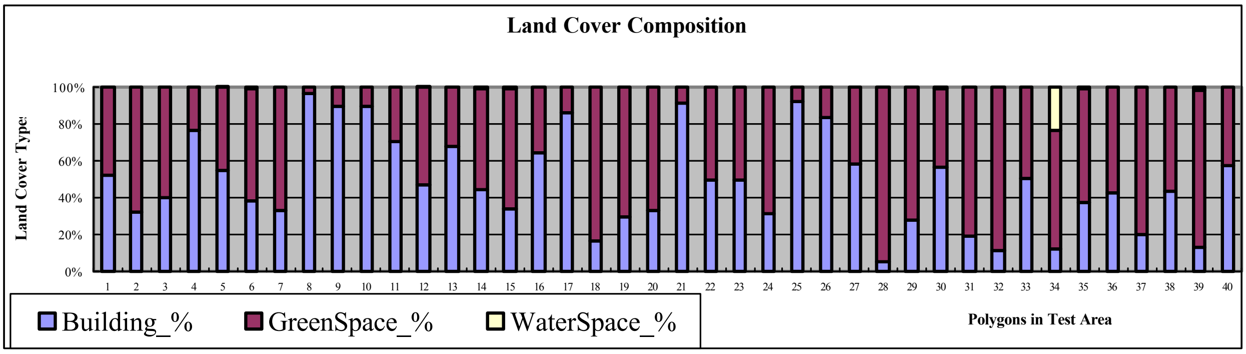


Fig. 8 Land cover composition and their proportion derived from land cover classification

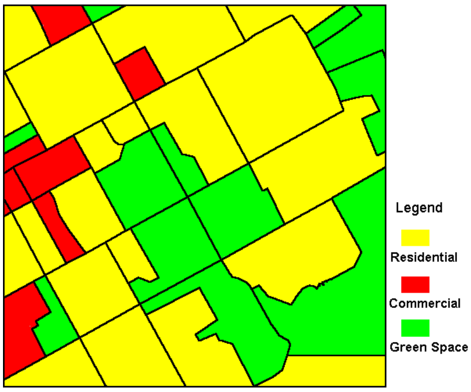


Fig. 9 Fuzzy land use classification

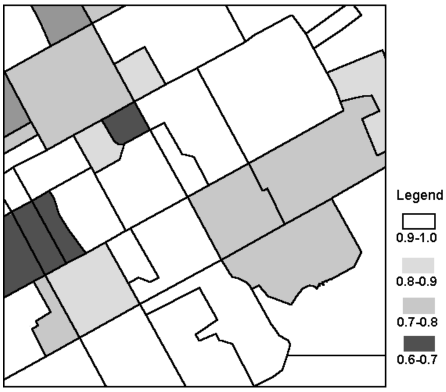


Fig. 10 Uncertainty assessment

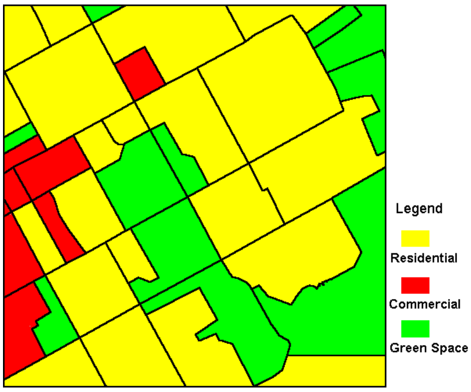


Fig. 11 Land use Map after adjustment

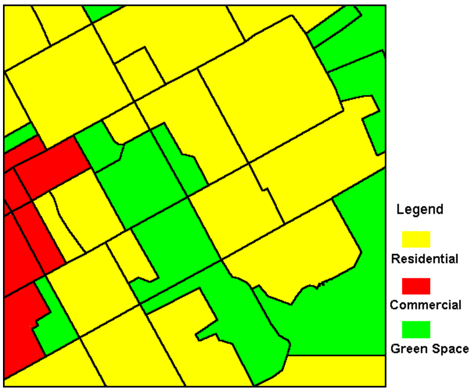


Fig. 12 Visual air photo interpretation

5 UNCERTAINTY ASSESSMENT

Comparing the Land use maps produced based on fuzzy land use classification to the one produced from visual air photo interpretation, 33 areas were correctly classified and 7 were misclassified (Table 1). Among the misclassified, 3 contain the right land use class as secondary land use function. As a result, 96.3 percent of the area has been classified correctly after adjustment for classification criteria. Of the misclassified area, Polygon No. 17 is a newly built-up residential area, and the building density has been overestimated in the land cover classification phase. In polygon No. 13 and 26, the building density or green space has been overestimated mainly due to the problem of image segmentation. Based on the fact that 10% of the polygons have been wrongly classified, compared to only 3.7% in the total area (Table 1), we feel that smaller area will have comparably higher opportunity of being misclassified. In other words, large homogeneous areas can be classified easier. In addition, most of the misclassified areas have a lower value (Fig. 10). This offers the end user an indication of confidence in using produced land use map.

Table 1 Quality Assessment of fuzzy land use classification

	Correct Classification	Correct Class found in Secondary Class	Wrong Classification
Number of Polygons	33	3	4
Percentage	82.5%	7.5%	10%
Area in Hectare	816.24	50.00	33.76
Percentage	90.7%	5.6%	3.7%

6 CONCLUSION

The proposed approach has incorporated conventional pixel-based approaches and fuzzy approach in land use classification in two steps. The result shows that the proposed approach can reach a quite good quality. The final result of such a classification process will consist of a land use map and a corresponding map indicating uncertainty levels of the assignment of each area to the relevant classes. In addition, it is possible to explore additional indicators other than those derived from images such as location, spatial relation etc. For instance, an area that has, based on its compositional, proportional or spectral characteristics, a high possibility values for both commercial use and residential use could be classified as 'commercial' if it is located near the city center or near the main streets. It could be classified as 'residential', as major function, if it is located in a place surrounding by residential areas and it could be classified as 'commercial', as secondary function. In this case study, the approach is mainly based on estimating the real cover of an area by features like green space, water body, built up area etc.

This approach offers the opportunity to adapt the decision rules for this inference to the local conditions of cities by adjusting the parameters based on local knowledge. It has potential value in monitoring land use changes and their tendency by comparing their major function and mineral functions according to their fuzzy membership values and their compositions for different periods.

ACKNOWLEDGEMENTS

Thanks to Drs. Henk Lutchman for his help in data collection. Thanks also to Ms. Yinghui Xiao for her assistance in data processing.

REFERENCES

Baraldi, A., and Parmiggiani, F., 1990. Urban area classification by multispectral SPOT images. *IEEE Trans. On Geoscience and Remote Sensing*, 28(4): 674-680.

Barnsley, M.J., Barr, S.L., and Sadler, G.J., 1991. Spatial Re-classification of Remote Sensed Images for Urban Land Use Monitoring. *Proc. Spatial Data 2000*, 17-20 September 1991, Oxford University, (Nottingham: Remote Sensing Society), 106-117.

Barr, S.L., 1992. Object-based Re-classification of High Resolution Digital Imagery for Urban Land-Use Monitoring. *International Archives of Photogrammetry and Remote Sensing* (ISPRS, Washington, D. C. 1992). Vol. XXIX, Part B7, Commision VII, 969-76.

Cheng, T., 1999a. *A Process-Oriented Data Model for Fuzzy Spatial Objects.* Ph.D. Thesis Wageningen Agricultural University (WAU), ITC, Enschede, The Netherlands, ISBN 90-5808-038-2, ITC Publication 68.

Cheng, T. and Molenaar M., 1999b. Objects with Fuzzy Spatial Extent. *Photogrammetric Engineering & Remote Sensing*, Vol. 63, No. 4, pp. 403-414.

FLIERS, (1999). Fuzzy Land Information from Environmental Remote Sensing, from: http://www.geog.le.ac.uk/fliers/details/further.html.

Foody, G. M., 1999. Image Classification with a Neural Network: From Completely-Crisp to Fully-Fuzzy Situation. In: Atkinson, P. M. and Tate, N. J., (eds.). *Advances in Remote Sensing and GIS Analysis*, UK: John Wiley and Sons Ltd, pp. 17-37.

Gorte B., 1998. *Probabilistic Segmentation of Remote Sensed Images.* Ph.D Thesis Wageningen Agricultural University (WAU), ITC, Enschede, The Netherlands, ISBN 90-6164-157-8, ITC Publication 63.

Gorte B., 1999. Supervised Segmentation by Region Merging. In: Kanellopoulos I., Wilkinson G. G. and Moons T. (eds.), *Machine Vision and Advanced Image Processing in Remote Sensing*, Berlin: Springer, pp. 328-335.

Hootsmans, R. 1996. Fuzzy Sets and Series Analysis for Visual Decision Support in Spatial Data Exploration. Ph.D Theisi Utrecht University, The Netherlands, ISBN 90-6266-134-3.

Mather, P. M., 1999. Land Cover Classification Revisited. In: Atkinson. P. M. and Tate, N. J., (eds.). *Advances in Remote Sensing and GIS Analysis*, UK: John Wiley and Sons Ltd, pp. 7-16.

Molenaar M., 1993. Object hierarchies and uncertainty in GIS, or why is standardisation so difficult. *Geo-informations-Systeme*, 6(4), 22-28.

Molenaar M., 1996. A Syntactic Approach for Handling the Semantics of Fuzzy Spatial Objects. in: Burrough P. A. et al. (Eds), *Geographic Objects with Indeterminate Boundaries*, pp. 207-224, London: Taylor & Francis Ltd.

Molenaar M., 1998. *An Introduction to Theory of Spatial Object Modelling for GIS*, London: Taylor & Francis Ltd.

Sadler, G.J, and Barnsley, M.J., 1990. Use of population density data to improve classification accuracies in remotely-sensed images of urban area. *First European GIS Conference (EGIS'90)*, Amsterdam, The Netherlands, 10-13 April 1990, (Utrecht: EGIS Foundation), 968-977.

Sadler, G.J, and Barnsley, M.J., and Barr, S.L., 1991. Information extraction from remotely-sensed images for urban land analysis. *Second European GIS Conference (EGIS'91)*, Brussels, April 1991, (Utrecht: EGIS Foundation), 955-964.

Smits, P. C., Dellepiane S. G., and Schowengerdt R. A., 1999. Quality assessment of image classification algorithms for land-cover mapping: a review and a proposal for a cost-based approach. *Int. J. Remote Sensing*, Vol. 20, No. 8, 1461-1486.

Tilton J. C., 1999. Hierarchical Image Segmentation: as applied to Remotely Sensed Multispectral or Hyperspecral Imagery. From: http://iifs.gsfc.nasa.gov/code935/tilton/index.html#INTRO.

Timpf S., 1998. *Hierarchical Structures in Map Series.* Vienna: GeoInfo Series.

Tonjes, R., Growe S., Bucker J. and Liedtke C. –E., 1999. Knowledge-Based Interpretation of Remote Sensing Images Using Semantic Nets. *Photogrammetric Engineering & Remote Sensing*, Vol. 65, No. 7, pp. 811-821.

Townshend, J. R.G., 1980. Spatial resolving power of earth resources satellites. *Progress in Physical Geography*, 5, 32-55.

Wharton, S.W., 1982. A contextual classification method for recognizing land use patters in high resolution remotely-sensed data. *Pattern Recognition*, 15, 317-324.

Whitehouse, S., 1990. A spatial land use classification of an urban environment using high resolution multispectral satellite data. *Proc. 16th Ann. Conf. Remote Sensing Society, Remote Sensing and Global Change*, University Collage Swansea, 19-21 September 1990, (Nottingham: Remote Sensing Society), 433-437.

Zhang, Y., 1999. Optimisation of building detection in satellite images by combining multispectral classification and texture filtering, *Journal of Photogrammetry & Remote Sensing*, Vol. 54, No. 1, pp. 50-60.