AUTOMATED RECOGNITION OF URBAN OBJECTS AND THEIR MORPHOLOGICAL ATTRIBUTES USING GIS

A. Peeters *, Y. Etzion

Desert Architecture and Urban Planning Unit Jacob Blaustein Institutes for Desert Research Ben-Gurion University of the Negev, Sede Boqer Campus, Israel, 84990 apeeters@bgu.ac.il

KEY WORDS: Contextual, Classification, GIS, Object Recognition, Segmentation, Urban

ABSTRACT:

Urban geometry is one of the most important aspects that influence urban microclimatic conditions. Developing and updating databases of urban geometry is, therefore, important for studying climatic aspects of urban form, especially where no town plans or updated surveillance data is available.

The current paper presents a parametric model which enables automated recognition of urban objects, in particular open spaces, from very-high spatial resolution remotely sensed data. Once objects are recognized, morphological attributes are automatically extracted. The developed model combines *spatial*, *spectral* and *context* – based recognition. It adopts a different approach to the challenge of automated object recognition, in which Geographic Information Systems (GIS) play a major role in the recognition of objects. While the segmentation is performed with image processing software, the classification is performed in GIS using a rule-based reasoning model.

A major challenge in the automated recognition of urban objects is that urban objects often do not adhere to the basic assumptions that automated recognition systems are based on, such as consistent pixel intensity. To overcome this problem, the presented methodology, makes use of (a) the variety of generic *context*-based relations between objects in urban form, and (b) the ability of GIS to recognize *contextual* relations.

The model was applied to a case study and statistically tested for its accuracy. Results are promising and demonstrate the potential of the model as a quantitative and systematic tool. Being a parametric model, it can be modified and applied on a large number of case studies. Recognized objects and extracted attributes can be used for constructing and updating GIS databases of urban form.

1. INTRODUCTION

At the turn of the new millennium urbanization has become one of the most important processes of human civilization, to a large degree, determining the future of mankind and his environment. Urban surfaces are rapidly replacing natural land cover, affecting the energy and water balances of the city and creating unique local conditions known as the "urban climate" (Oke, 1987). Urban geometry in particular has been considered as one of the most important generators of the urban climate (Oke, 1981).

To understand climatic aspects of the urban geometry, a common practice in architectural and urban planning research is to analyze the urban geometry of traditional (vernacular) urban settlements. This type of urban form has developed through a process of trial and error. The underlying rationale is to recognize patterns in urban geometry, which might have developed as a response to prevailing climatic conditions. However, these types of settlement are often lacking updated surveillance data or accurate town plans. This limits the development of comprehensive GIS databases of urban form and consequently impedes subsequent analysis.

Automated recognition of urban objects, such as buildings, and roads, from remotely-sensed data, has been gaining increasing interest and popularity, primarily because of its potential to

extract fast and accurately urban objects, while reducing time and labour intensive tasks associated with manual digitizing and field surveillance. In addition, extracted objects and associated data can be integrated into GIS databases for further analysis, modelling, visualization and mapping (Lillesand and Kiefer, 2000; Mayer, 1999).

A major challenge in the automated recognition of urban objects is that these objects often do not adhere to the basic assumptions that automated recognition systems are based on, such as consistent pixel intensity, predictable shapes and well-defined edges (Irvin and McKeown, 1989). Additional challenges are a high degree of spatial and spectral heterogeneity (Zhang, 1999; Mayunga, Zhang et al., 2005), abundance of urban details which introduces "noise" into the recognition process, the extraction of 3D information from vertical images and converting the raster output of the object recognition into a vector topology for GIS vector-based analysis. To overcome these problems several solutions were integrated and a hybrid model was developed.

The following presents the development, application and verification of a model, which enables automated recognition of urban objects and their morphological attributes from remotely-sensed data. This particular model focuses on the recognition of urban open spaces and on the extraction of morphological

-

^{*} Corresponding author.

attributes essential for analyzing the relation between climatic conditions and urban form.

2. EXISTING APPROACHES FOR RECOGNIZING URBAN OBJECTS

Automated recognition can be based either on a set of predefined classes known as supervised classification or on intrinsic groupings within the dataset termed unsupervised classification. The model presented here adopts the method of supervised classification in which prior knowledge on the objects is used to develop the classes.

Current automated recognition of urban objects is based on two main approaches of classification: (a) a spectral or pixel-based classification, which is based on the spectral characteristics of the pixels, and (b) a spatial or object-based classification, which is based on the spatial characteristics of the objects. While the first ignores spatial characteristics such as object size and shape, the second consists of methods which categorize pixels based on the spatial relationship between them and surrounding pixels.

A more recent approach is the context-based classification. While the previous two approaches extract the information required for the recognition from the pixels or from the objects in the image, the context-based method operates at the level of image understanding, looking at the whole image to draw the required information. This approach is well-suited to recognize objects in urban form since a variety of generic relations between objects can be found, that are independent of lighting conditions, building materials, image rotation, object's scale and size. For example, a courtyard will always be located within a building or confining walls and a building and its shadow will always share an edge segment.

A model based only on one type of recognition might be insufficient for producing accurate results, especially in urban environments. For example, pixel value might vary within one type of object, while different object classes might have the same pixel value e.g. a stone house and a stone paved road. In addition previous research indicates that a combination of approaches improves the accuracy of the classification (Mayer, 1999; Jing, Qiming et al., 2007). The tendency in recent studies is to develop systems that combine algorithms from different approaches (Zhang, 1999; Straub, Wiedemann et al., 2000; Zhu and Blumberg, 2002; Mueller, Segl et al., 2004; Jing, Qiming et al., 2007).

To enhance recognition and to develop a more generic rather than a specific model, a hybrid approach was adopted, that combines techniques from spectral, spatial and context based recognition.

3. GIS FOR OBJECT RECOGNITION

The common practice in systems which combine remote sensing and GIS is to perform the object recognition using external image processing software. Only after objects have been recognized (classified) they are vectorized and integrated into a GIS database for further analysis.

This paper presents a different methodology in which only the first part of the object recognition - the segmentation - is performed using image processing software, while the actual object recognition is performed in GIS. Once objects are represented in GIS using a vector topology, morphological attributes can be extracted from the objects using geoprocessing tools. When the GIS database is completed, the data can be analyzed to identify trends, patterns and relationships.

Spatial analysis is considered the core of GIS and consists of the processes, methods and tools for analyzing spatial data for developing spatial information (Longley, Goodchild et al., 2005). There are various types of spatial analysis. Two types: (a) queries and reasoning and (b) transformations, are combined in the object recognition model presented here. Queries and reasoning refers to methods that allow the user to interrogate the data and transformations refer to methods in which data is combined to obtain new information. Both are used to derive information about the spatial context of the urban objects for the context-based recognition. Objects are selected and recognized based on their location in relation to other objects, for example, whether the objects in one layer intersect with the objects of another layer or are completely contained within the objects of another layer.

4. CALCULATING HEIGHT BASED ON SHADOWS

Shadows provide a good method for extracting 3D information from a 2D image (Irvin and McKeown, 1989; Mayer, 1999). The length of a shadow cast by an object protruding from the ground (a pole, a building, a tree or a wall) depends on the height of the object, the sun's position (i.e. the date and the solar time of the day) and the slope of the ground. On flat terrain the height of an object can be calculated from the shadows length and the sun's altitude angle by using simple trigonometry. To extract the length of the shadows they must first be automatically extracted. This is achieved within the segmentation process described in the methodology section, in which the shadowed areas are first defined as an additional class.

It is important to note that for the task of shadow recognition images must contain clear shadows. However, shadows in the image might interfere with the process of open space recognition, since objects that are partly obstructed by cast shadows cannot be considered as homogenous regions. Ideally two images of the same site are needed, one image, captured on early morning or late afternoon hours when shadows are clear, for recognizing shadowed areas and the other captured on midday, with minimum shadows, for recognizing open spaces.

5. METHODOLOGY

5.1 Developing a Hierarchical Structure of Classification

The first step in designing a system for object recognition that is based on supervised classification consists of pre-defining the object classes. This is a subjective non-computational process which is based on prior understanding of the objects. A visual analysis of remotely-sensed images and figure-ground maps was carried out combined with a literature review to identify a number of key urban form components that generally characterize the morphology of traditional urban form. In this study only locations situated in drylands were considered. Drylands are ideal areas for remote sensing of urban form since (a) open spaces have generally well-defined borders (b) vegetation cover is low and confined (c) images tend to be clearer due to low cloud cover, and (d) high reflectance values due to arid conditions result in a higher signal to noise ratio.

Fig.1 is an example taken from the visual analysis illustrating the major urban elements which were identified.

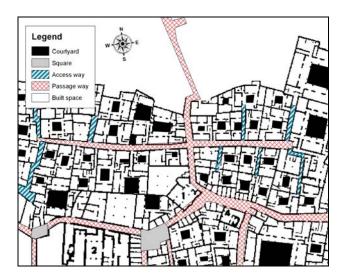


Figure 1. Identifying major urban elements in the residential district of Mokhfiya in Fez, Morocco. Base map from: Urban Form in the Arab World (Bianca, 2000)

The major elements formed the base for predefining four classes: built space, courtyards, squares and streets. These classes were organized in a hierarchical pyramidal structure (Fig.2) in different levels of abstraction; the image being the first level of abstraction whereas the courtyard, for example, being the third level of abstraction (Straub, Gerke et al., 2001). Each class was divided into sub-classes; the open space class, for example, is a sub-class of the image while the polygonal features class is a sub-class of the open space class. Recognition is based on a top-down process which is divided into three levels of scale, from the level of the entire settlement down to the single objects. An additional class – the shadows class – was added to the hierarchical structure to allow the extraction of height attributes.

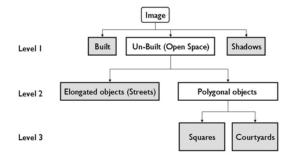


Figure 2. Hierarchy of classification

5.2 Remotely-Sensed Data

The spatial resolution of the image is considered by many as the most crucial technical matter regarding urban remote sensing (Donnay, Barnsley et al., 2001). The recognition system should operate at a spatial resolution that will facilitate recognition of important object details. Based on Mayer's (1999) review on object extraction, the optimal spatial resolution is defined as half of the size of the object which has to be recognized. Konecny and Schiewe (1996) present in their article on mapping from satellite image data a variety of object types and the spatial resolution required to facilitate their recognition in terms of location and object type. For buildings they define a 2 meters or higher ground pixel size.

The often compact dimensions of traditional open spaces require a relatively high spatial resolution, of a magnitude of 1m or higher. For example, traditional open spaces are sometimes defined by confining walls. To optimize the detection of these walls the spatial resolution must be higher than the required spatial resolution for contemporary buildings. To extract morphological attributes for GIS analysis, in addition to location and object type, a high spatial resolution is required. High spatial resolution data can be acquired either from new generation satellites or from digital aerial photography. A study by Toutin and Cheng (2002) has demonstrated that Quickbuird imagery has narrowed down the existing gap between aerial and satellite imagery in terms of spatial resolution. A 0.70m spatial resolution in their standard color imagery makes Quickbird imagery an ideal data source for developing the current model.

5.3 Structure of Parametric Model

The developed model consists of two main components: (a) object recognition and (b) attribute extraction. The approach adopted for the object recognition process (Fig.3) is based on the three level hierarchy which was identified.

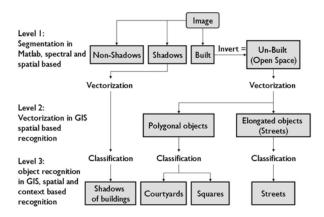


Figure 3. Approach of the object recognition component

Level 1: Differentiation between *built* and *un-built* areas and between *shadows* and *non-shadow* areas.

Recognition in this level is based on segmenting the image into homogenous regions of objects using Matlab® image processing tools. Since two images of the same location, captured in different times of the day, were not available for this research, one image with clear shadows was used and its histogram was manipulated to enhance either the shadows or the buildings. Familiarity with the form of vernacular urban settlements helps in the recognition of the *un-built* class. Vernacular urban settlements are characterized by a continuous built fabric. Therefore, urban space is either *built* or *un-built* and recognition of the *un-built* (open space) class can be based on an invert of the recognized *built* space.

Level 2: Differentiation between *elongated objects* (streets) and *polygonal objects* (courtyards and squares)

Recognition in this level is based on a process of *vectorization* performed in GIS. Since at this level the classes are characterized by similar spectral properties (they are represented in binary images), recognition in this level is based on the spatial characteristics of the objects. The process consists of representing the objects through vectorizing them either by their boundaries using polygons, or by reducing them into a linear representation. This process allows to describe the objects using spatial and contextual descriptors at the following level.

Previous knowledge of vernacular urban form serves for the differentiation between the elongated objects and the polygonal objects class.

Level 3: Differentiation between *squares*, *courtyards*, *streets* and *shadows*.

Recognition in this level is based on spatial and contextual descriptors. The same contextual relations which were observed between the objects in this study can be found in the majority of vernacular settlements regardless of spectral and spatial characteristics such as building material or size. For example a square will most probably intersect with the street network. This enhances the performance of the model and makes it suitable for the analysis of most vernacular settlements. A list of classification rules based on spatial descriptors (morphological characteristics) and on contextual relations between the objects was defined. These were used to perform queries on the objects for the final classification. The process of differentiating between the final classes was executed in GIS. The following describes the processes involved in the object recognition component and the attribute extraction component.

5.3.1 Segmentation: Following pre-processing to convert images to grey-scales and adjust their intensity to enhance either the buildings or the shadows, images are segmented twice into regions: first into *shadows* and *non-shadow* areas, and then into *built* and *un-built* (open) areas. Segmentation of shadows is based on region segmentation using *Morphological image processing* techniques. The main objective in the *shadows* class recognition was to find "candidates" from which the shadow length could be extracted for the calculation of the object's height. Therefore segmentation focuses on segmenting shadow regions which are significant in terms of their length.

The segmentation of the built class is based on both a region-based and on an edge-based segmentation and uses also *Morphological image processing* techniques. The *Canny edge detector* (Canny, 1986) is used for the edge-based segmentation.

Description and Classification: The outputs of the 5.3.2 segmentation are introduced into the GIS and are registered to enable correct spatial location. To allow a vector representation for morphological attribute extraction, a batch vectorization is applied to the segmented outputs. The final step in the object recognition is the actual classification of objects. Structural analysis is a recognition method which describes the objects based on their spatial structure - composition and arrangement of elements. This approach is particularly suited when objects have an obvious structure and an arrangement that can be defined by a combination of rules for example in the analysis of urban shapes (Barr and Barnsley, 1998; Anil, Robert et al., 2000). A set of spatial descriptors, which define morphological attributes, and contextual descriptors, which define generic relations between objects, was used to develop the classification rules. The relations and classification rules were described using a rule-based reasoning model which is based on a decision tree that consists of an IF / THEN / ELSE sequence of rules. The answer to the IF condition directs the process to the THEN or to the ELSE branch of the tree. Object candidates are evaluated using the predefined classification rules based on thresholds or on a Yes/No answer. Polygon overlay is used to derive information about the context and organization of the urban objects. The information is extracted through queries, which identify objects from one layer based on their shape properties and relations to objects in another layer.

Morphological Attribute Extraction: The outputs of the object recognition are the recognized classes: shadows, courtyards, squares and streets, represented as polygons or polylines. This step consists of building the database by extracting morphological attributes from the objets. At this stage, this study suffices with extracting only the height and the width of objects. These parameters are needed for the calculation of the height-to-width ratio and the solar access to open spaces - used for subsequent bioclimatic analysis of urban form. Calculation of building's height is based on the length of the shadow and is done using a technique developed in GIS and explained in detail in Peeters and Etzion (2009): polygons that represent shadows are queried to identify all lines with a specific azimuth angle (computed according to the date and the solar time in which the data was obtained and the geographic latitude). This returns only the lines that represent the shadow length. The width of open spaces is extracted using GIS methods. Additional morphological attributes can be extracted at more advanced stages in future development of the model.

6. APPLICATION OF MODEL

The process of recognition was automated by developing: a Matlab® script for the segmentation process which can be found in Peeters and Etzion (2009), and two GIS geoprocessing models: one for the classification process and the other for the attribute extraction process. All are parametric and can be modified based on the specific image.

6.1 Case Study

The automated model was applied to a sub-set from a satellite image of a vernacular section in the city of Marrakesh, Morocco (Fig.4). The image was downloaded from GoogleEarth (Google, 2005). It is an image from DigitalGlobe Inc. and was captured by the Quickbird satellite (DigitalGlobe, 2006) at March 24 2006. By choosing an image from an open source – like Google Earth – preference is given to wider usability of the model. The image is a vertical RGB image with a spatial resolution of 70cm and a 0% cloud cover.



Figure 4. Grey-scale image of case study (converted from the original RGB image). Image © 2008 DigitalGlobe © 2008 Europa Technologies © 2008 Google

Fig.5 and Fig.6 illustrate the recognition outputs of three classes in different stages of the recognition process: the *shadows* class, the *built* class and the *courtyards* class.

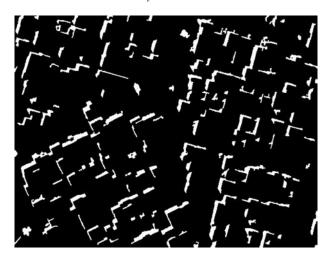


Figure 5. Segmented *shadows* and *non-shadow* areas - output of Matlab® segmentation process



Figure 6. Recognized *buildings* and *courtyards* overlaid over original image - output of GIS classification process

Table 7. presents a section from the developed database for the courtyard class. The height and width of the courtyards was extracted (in meters). The database consists also of the extracted shadow length and of additional climatic parameters. These were extracted in order to analyze the influence of climate on the geometry of the open spaces and are described in detail in Peeters and Etzion (2009).

Shadow length extracted	Height	Width	h/w extracted
3.32	4.75	1.42	3.35
3.32	4.75	1.57	3.03
2.71	3.88	2.27	1.71
3.32	4.75	1.28	3.71
2.71	3.88	2.26	1.72
1.68	2.41	1.94	1.24
2.01	2.88	1.44	2.00

Table 7. Section from the attribute extraction of the courtyard class (all values are in meters)

6.2 Model verification: results and analysis

The accuracy of the object recognition component was evaluated using a confusion matrix (Lilesand and Kiefer, 1994; Jing, Qiming et al., 2007). Object recognition results were compared to a manually digitized dataset. The *stratified random sampling* method (Jing, Qiming et al., 2007) was adopted for the random sampling of pixels to improve the sampling set and its representation of the whole dataset. In addition the *kappa statistic* was computed for each class and for the whole matrix.

Results of the confusion matrix demonstrate an overall accuracy of 80.30% with a kappa coefficient of 0.679. Although the overall accuracy is satisfactory with a rate common to the rates of existing recognition systems, the low kappa coefficient is ambiguous. A closer examination of the matrix shows that the results can be divided into two distinctive groups: one group including the *courtyards* class and the *built* class has quite high values of *users accuracy* with 87.50 % and 90.76 % respectively and kappa coefficients are quite high too, with values of 0.8714 and 0.8021. The other group which includes the *squares* class and the *streets* class has lower values of *users accuracy* with 78.69 % and 66.49 % respectively and kappa coefficients which are also lower with values of 0.7513 and 0.5239.

Several major conclusions can be drawn from the results:

- Confusion between classes occurs mainly among pairs of classes which share edges, for example, between the streets class and the built class or between the streets and the squares. Classes which do not share edges such as the streets class and the courtyards class or the squares class and the courtyards class are not being confused (zero values in the matrix).
- The high confusion between the built class and the street class in which a large number of built pixels were omitted and recognized as streets (errors of omission in the built column) might be due to shadows cast on buildings by adjacent buildings.
 These might be mistakenly recognized by the system as streets (most likely as access paths between buildings).
- The complexity of the image posses a challenge to manual digitizing. It might well be that shadows on buildings, for example, are confused as access ways. Manual digitizing inevitably introduces errors; classified data might be compared to erroneous manually digitized data. This problem could be solved with images of higher spatial resolution and with minimum shadows. Another option is to compare the classified data to field data for example to digitized urban plans which are based on field survey.

7. SUMMARY AND CONCLUSIONS

Urban open spaces are particularly essential in metropolitan and in densely-built areas for their visual, health and climate moderating aspects. To enhance the public use of these open spaces, it is essential to improve their climatic performance by correct planning of their form and relation to surrounding buildings. All this requires comprehensive databases of urban objects and their morphological attributes.

The important feature of the developed model is its ability to extract and analyze urban data off site, minimizing the need for time, labour and capital intensive processes inherent to on location field surveys and manual digitizing. In cases of large bodies of data it may prove impractical.

Automatic processes offer a feasible alternative for constructing and updating databases of urban form.

Although results are promising, future research is required to improve performance of the model by introducing to the model (a) images created by other wavelength bands, for example NIR images for addressing issues of vegetation (b) multiple images for addressing the shadows problem (c) more context-based rules for classification to substitute the spatial-based rules and enhance generalization of the model to support a larger variety of case studies (d) higher spatial resolution images to address the problem of wall recognition and differentiation of narrow streets between buildings, and (e) improvement of segmentation process to enhance segmentation of buildings particularly those which share similar spectral properties with the streets.

ACKNOWLEDGEMENTS

This work was supported and funded by an Eshkol Scholarship from the Ministry of Science & Technology, The State of Israel.

REFERENCES

- Anil, K. J., Robert, P. W., Jianchang, M., 2000. Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Inteligence*, 22(1), pp. 4-37.
- Barr, S. L., Barnsley, M. J., 1998. Application of structural pattern-recognition techniques to infer urban land use from ordnance survey digital map data. *3rd International Conference on GeoComputation*, Bristol.
- Bianca, S., 2000. *Urban Form in the Arab World*. Thames & Hudson, London.
- Canny, J. F., 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Learning*, 8(6), pp. 679-698.
- DigitalGlobe, 2006. QuickBird Scene of Marrakesh ID1010010004E1EB01. © DigitalGlobe, Longmont, Colorado.
- Donnay, J. P., Barnsley, M. J., Longley, P. A., 2001. *Remote Sensing and Urban Analysis*. Taylor&Francis, London.
- Google, 2005. *Google Earth 3.0.0762*. http://earth.google.com/Accessed: March, 2007.
- Irvin, R. B., McKeown, D. M., 1989. Methods for exploiting the relationship between buildings and their shadows in aerial imagery. *IEEE Transactions on Systems, Man and Cybernetics* 19(6), pp. 1564-1575.
- Jing, Q., Qiming, Z., Quan, H., 2007. Comparison of pixel-based and object-oriented classification methods for extracting built-up areas in aridzone. *ISPRS Workshop on Updating Geospatial Databases with Imagery & The 5th ISPRS Workshop on Dynamic and Multi-dimensional GIS Urumchi*, China.
- Konecny, G., Schiewe, J., 1996. Mapping from digital satellite image data with special reference to MOM-02. *ISPRS Journal of Photogrammetry & Remote Sensing*, 51, pp. 173-181.

- Lillesand, T. M., Kiefer, R. W., 2000. *Remote Sensing and Image Interpretation*. New York, John Wiley & Sons, Inc, New York.
- Longley, P. A., Goodchild, M. F., Maguire, D. J., Rhind, D. W., 2005. *Geographic Information Systems and Science*. John Wiley & Sons Inc., West Sussex.
- Mayer, H., 1999. Automatic object extraction from aerial imagery A survey focusing on buildings. *Computer Vision and Image Undestanding* 74(2), pp. 138-149.
- Mayunga, S. D., Zhang, Y., Coleman, D. J., 2005. Semiautomatic building extraction utilizing Quickbird imagery. In: Proceedings of the ISPRS Workshop CMRT05, IAPRS, International Archives of Photogrammetry and Remote Sensing. Vienna, Austria.
- Mueller, M., Segl, K., Kaufman, H., 2004. Edge and region based segmentation technique for the extraction of large, manmade objects in high-resolution satellite imagery. *Pattern Recognition*, 37, pp. 1619-1628.
- Oke, T. R., 1981. Canyon geometry and the nocturnal urban heat island: comparison of scale model and field observations. *Journal of Climatology*, 1, pp. 237-254.
- Oke, T. R., 1987. *Boundary Layer Climates*. Routledge, Taylor & Francis Group, London.
- Peeters, A., Etzion, Y., 2009. A Quantitative Systematic Approach for Evaluating the Relation between Climate and Urban Form, PhD Thesis, Ben-Gurion University of the Negev, Israel
- Straub, B. M., Wiedemann, C., Heipke, C., 2000. Towards the automatic interpretation of images for GIS update. *IC Working Group IV/III.2 Integration of Image Analysis and GIS, International Archieves of Photogrammetry and Remote Sensing*, Amsterdam.
- Straub, B. M., Gerke, M., Koch, A., 2001. Automatic extraction of trees and buildings from image and height data in an urban environment. *International Workshop on Geo-Spatial Knowledge Processing for Natural Resource Management*, Varese, Italy.
- Toutin, T., Cheng, P., 2002. 3D models for high resolution images: examples with QuickBird, IKONOS and EROS. Symposium on Geospatial Theory, Processing and Applications, Ottawa.
- Zhang, Y., 1999. Optimisation of building detection in satellite images by combining multispectral classification and texture filtering. *Photogrammetry & Remote Sensing*, 54, pp. 50-60.
- Zhu, G., Blumberg, D. G., 2002. Classification using ASTER data and SVM algorithms; the case study of Beer Sheva, Israel. *Remote Sensing & Environment*, 80, pp. 233-240.