Road extraction using object oriented classification

by Paidamwoyo Mhangara, Linda Kleyn and Hardly Remas, SANSA Space Operations, and John Odindi, University of Fort Hare

Rapid mapping of tarred roads has become an important requirement for most urban planners. To match the dynamic nature of urban developments, use of high resolution multispectral aerial photography and object-oriented classification have recently become the most viable options for mapping roads and updating urban maps. Using the city of Cape Town as a case study, this article presents an objected-oriented method for classifying roads using high resolution aerial photography.

Roads are an important part of the transportation infrastructure and their role is central to a variety of applications that support urban planning, traffic management and vehicle navigation. Accurate and up to date information on road networks is a critical requirement for any country. A number of urban planning and management system applications use road networks as an important input (Bong et al., 2009). Roads facilitate efficient transportation of people, goods and services and thus perform a fundamental function in the socio-economic activities of any nation [1]. The importance of road networks is also evident in image geometric rectification procedures where they are used to georeference satellite imagery and aerial photography.

Road mapping is an important but problematic task in photogrammetry and remote sensing. The need for efficient methods that accurately extract road networks is particularly evident in rapidly expanding urban areas where road maps need to be updated regularly. Manual digitising using GIS software is the conventional road extraction method for updating road maps. This method is tedious and time consuming and is usually unable to match the rates of urban developments and the timely provision of up to date road maps. Road extraction is a conventional problem of remote sensing that has not been completely resolved. Mena [2] outlines several methods that have been used to automate the extraction of roads. The extraction of roads is particularly problematic in urban areas due to the spectral similarity of roads and other impervious surfaces such as buildings.
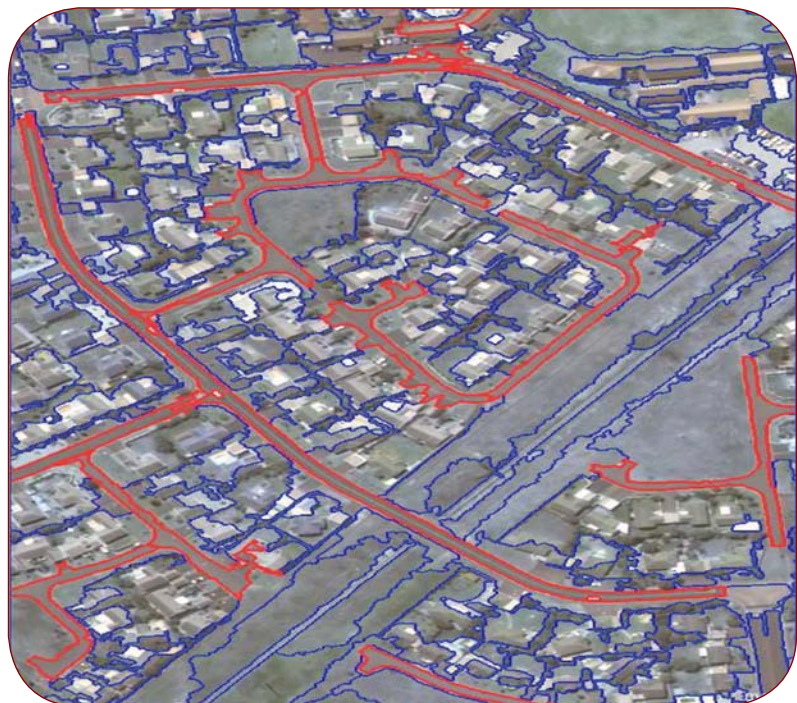

Fig. 1: Multi-resolution segmentation results delineating roads.

Indeed, the spectral separability of asphalt road surfaces and bituminous roofs is still not easy since they tend to share similar spectral properties. This problem is further exacerbated by issues related to the occlusion of road surfaces by trees, shadow and the presence of vehicles.

Some progress has however been made in trying to resolve the road extraction problem. Many a scholar has proposed different methods to crack this issue. Mena [2] provides a detailed literature review of most methods which have been applied to extract roads. Most of the road extraction techniques fall within these categories: dynamic programming and snakes methods, road tracking methods, morphological analysis, stereoscopic and multi-temporal analysis, hyperspectral experiments, and multi-scale and multi-resolution methods [2]. Gong and Wang [3] suggest that most road network extraction methods are based on linear analysis methods. Methods based on linear analysis detect edges through edge filtering, morphological filtering and gradient profile analysis. A number of road extraction techniques use edge detection algorithms to detect road boundaries. Hinz [4] proposed a road extraction model which uses context information analysis. Clustering,

Object class combination	Separability	Direction	Threshold
Roads from bare ground			
NDWI	1,99964	>	0,0902821
BAI	1,99768	>	0,0635474
Roads from buildings			
Density	1,76939	<	1,00462
Mean layer NIR	1,76549	<	85,5418
Roads from water			
GLDV Ang. 2nd moment 4 (all directions)	1,99944	<	1,99944
GLDV Entropy Layer 4 (all directions)	1,99891	<	1,99943
Roads from grasslands			
BAI	2,00000	>	0,0164802
NDWI	1,99998	>	0,0687425
Roads from sparse vegetation			
Maximum difference	1,99089	>	0,256570
NDW1	1,98831	>	0,109264
Roads from dense vegetation			
NDWI	1,99228	>	0,256570
BAI	1,98909	>	0,109264

Table 1: Summarised result of the Seath analysis.

maximum likelihood classification and contextual classification have also been used to classify road network [3]. An algorithm based on wavelet transform was implemented by Gruen and Li [5] for the semi-automatic extraction of roads. Road extraction methods can be classified into two groups based on the use of spectral and geometric information: spectral classification based road extraction and the geometric based road extraction. Road extraction methods can be broadly classified as semi-automatic and automatic extraction (Bong et al., 2009). Initial points or road seeds need to be provided in semi-automatic extraction while road seeds can be detected automatically and linked to complete the road network in automatic road extraction. Techniques such as the Hough Transform have been proposed to detect road lines, while the snake method has been used to reconnect the broken lines [2]. The Hough Transform and snake methods are however not efficient and fail in complex road networks involving curved lines [2].

Methods based on multi-scale and multiresolutions have gained currency

in recent years due to their ability to control the width of the roads and their efficiency in high resolution imagery. Mayer et al. [6] provides a detailed overview of the model pattern in road extraction using multi-scale analysis. An elaborate description of the abstraction concept of multi-scale analysis is provided by Mayer and Steger [7]. Further literature on road extraction using multi-resolution segmentation and analysis is presented by Baumgartner et al. [8]. High resolution aerial photography and object extraction techniques provide a feasible means of road extraction. The major challenge confronting road extraction is the spectral similarity of roads and other built-up features which makes their separation problematic [9]. Mena [2] indicates that indeed the complete automatic extraction of roads is still unsolved. Manual digitising is still widely used for road extraction. This traditional method is tedious and fraught with error.

In recent years, high resolution multispectral aerial and satellite imagery have proved to be useful data sources for road extraction

(Li and Briggs, 2009). Object-oriented rule-based classification offers a number of new possibilities for road extraction using spectral, shape, and textural parameters. The feasibility of using rule based classification has been demonstrated by Baumgartner et al. [8]. Successful extraction of road features is inevitably dependent on optimum selection of thresholds and image layers which can adequately separate the features of interest. Considering that a large number of possible features describe objects in terms of shape, spectral and texture characteristics, it becomes imperative that effective classification rules can only be successfully developed using a comprehensive feature analysis methodology that can identify the most salient layers for classification and the best thresholds for feature extraction. The objective of this paper is to present an object-oriented rule-based road extraction methodology that uses spectral, shape and textural characteristics. The classification method is based on multiresolution segmentation of multispectral aerial photography and uses a comprehensive features analysis tool to determine the thresholds and most salient features for classification.

Study area

The study was done in Cape Town, South Africa. Cape Town is one of the economic hubs of South Africa and it is located at the border of the Indian and Atlantic Ocean. The parliamentary seat of South Africa is held in Cape Town and is a major tourist destination.

Methodology

This section provides a description of the methods used to map urban land cover features and their automation thereof. The aerial photographs were orthorectified in PCI Geomatica 10.3 software using digital aerial photography math modelling method and achieved an RMS error of less than 0,3 pixels.

Datasets and software

High resolution multispectral aerial photography acquired from the Chief Directorate National Geo-spatial Information (South Africa) was used for the classification. The spectral bands for the aerial photographs are blue, green, red and near infrared. Orthorectification of the aerial photographs was done using PCI Geomatica 10.3 software. Image
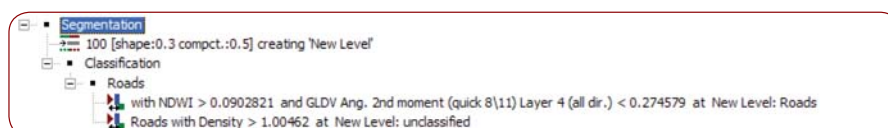
Fig. 2: Roads classification ruleset

Fig. 3: Separation of roads from buildings and non-roads land cover classes.

Fig. 4: Final road extraction results.

segmentation and classification was done in eCognition Developer 8.0 software. The feature selection and identification of thresholds was done using Seath software.

Image segmentation

The rationale of this procedure is to generate image objects that closely mirror tarred roads. To achieve this several tests were executed to segment the aerial photographs using the multiresolution segmentation algorithm. The multiresolution segmentation algorithm consecutively merges pixels or existing image objects. This procedure identifies single image objects of one pixel size and merges them with their neighbours, based on a relative homogeneity criterion. The homogeneity criterion is a combination of spectral and shape criteria [10]. Higher values for the scale parameter result in larger image objects, smaller values in smaller ones. The multiresolution segmentation uses an optimisation route that minimises the average heterogeneity of image objects and maximises their respective homogeneity for a given resolution. The testing involved applying different segmentation settings and changing the composition of homogeneity criterion. Key parameters that were tested are the layer weights as well as the shape and compactness criterion. The trials indicate that roads are well segmented at a scale parameter of 100. The image layer weights for all the four bands – blue, green, red and NIR were assigned the same weight. Settings for the composition of homogeneity criterion were assigned as 0,3 for shape and 0,5 for compactness. These parameters were found suitable to delineate tarred roads.

Classification

The classification was achieved by developing a ruleset that uses thresholds on the spectral bands, spectral indices, geometric properties, texture related features to classify tarred roads. Band ratios such the Normalised Difference Vegetation Index (NDVI) and Normalised Difference Water Index (NDWI) are computed as customised features in eCognition Developer, this approach permits the transferability of the ruleset. The brightness was also incorporated into the ruleset to facilitate the extraction of very bright surfaces such as beach sand and bare surfaces. All Haralick textural features such as the Gray Level Co-occurrence Matrix (GLCM) and Gray Level Difference Vector (GLDV) for all directions were used as input to the Seath program. The separation of roads was found to be most effective using the NDWI, GLDV. 2nd moment and density.

$$NDVI = \frac{NIR - R}{NIR + R} \qquad (1)$$

The Normalised Difference Vegetation Index (NDVI) was used to enhance vegetation information in the aerial imagery by ratioing a near-infrared (NIR) band to a red. This takes advantage of the high

Reference/ Classification	Roads	Non-roads	Sum
Roads	315	12	327
Non-roads	28	266	294
Sum	343	278	621
Overall statistics			
Producer's accuracy	0,918	0,957	
User's accuracy	0,963	0,905	
Specificity	0,957	0,918	
Overall accuracy	0,936		
Khat	0,870		

Table 2: Error matrix for road classification.

vegetation reflectance in NIR spectral range and high pigment adsorption of red light.

$$NDWI = \frac{G - NIR}{G + NIR} \qquad (2)$$

The Normalised Difference Water Index (NDWI) was used to delineate water features and tarred roads. This index maximises reflectance of water by using green light wavelengths and minimises low reflectance of NIR by water features while taking advantages of the high reflectance of NIR by vegetation and soil features. As a result, water features are enhanced due to positive values and vegetation and soil are suppressed due to zero or negative values. The spectral signature features of tarred roads in green and NIR band are similar with those of water, i.e., they both reflect green light more than reflected near infrared light. Consequently, the computation of the NDWI also produces positive values for tarred roads and built-up even though they are much lower than those of water.

$$BAI = \frac{B - NIR}{B + NIR} \qquad (3)$$

The Built-Up Areas Index is considered useful in detecting asphalt and concrete surfaces. Built-up Area Index (BAI) was calculated because it is a very robust index that remains relevant even if it is computed on an image with some vagueness.

Texture after Haralick

Texture has been one of the most important characteristics which has been used to classify and recognise objects and scenes. Different image textures manifest themselves by dissimilarity in both the property values and the spatial interrelationships of their component texture primitives.

Texture features are used to evaluate the texture of image objects and subsequently discriminate between them. Features based upon the gray level co-occurrence matrix Haralick are used in this study. Texture after Haralick based on the gray level co-occurrence matrix (GLCM), which is a tabulation of how often different combinations of pixel gray levels occur in an image [10]. In order to use the information contained in the gray level co-occurrence matrices. A detailed description of Haralick features is provided by Haralick and Shapiro [11]. Since many distances and orientations result in a large amount of computation, only the variance was used. The gray level co-occurrence matrix (GLCM) is a tabulation of how different combinations of pixel gray levels occur in a scene. A different co-occurrence matrix exists for each spatial relationship. Another approach to measure texture is to use a Gray Level Difference Vector (GLDV) instead of the GLCM. The GLDV is the sum of the diagonals of the GLCM and counts the occurrence of references to the neighbour pixels' absolute differences. To receive direction invariance, the sum of all four directions (0°, 45°, 90°, 135°) are calculated before texture calculation [10].

GLDV Angular 2nd Moment measures the local homogeneity. The value is high if some elements are large and the remaining ones are small.

GLDV Angular 2nd Moment=

$$\sum_{k=0}^{N-1} V_k^2 \qquad (4)$$

Where: N is the number of rows or columns

V_k is image object level, $k=1,\dots n$

Density for 2D image objects

The density feature describes the distribution in space of the pixels of an image object. In Definiens eCognition Developer 8, the most dense shape is a square; the more an object is shaped like a filament the lower the density. The density is calculated by the number of pixels forming the image object divided by its approximated radius, based on the covariance matrix [10].

$$Density = \frac{\sqrt{\#P_V}}{1 + \sqrt{VarX + VarY}} \qquad (5)$$

Where $\sqrt{\#P_V}$ is diameter of a square object with $\#Pv$ pixels.

$\sqrt{VarX + VarY}$ is diameter of the ellipse

Threshold and feature determination using Seath

Seath was used to statistically compute the separability and the corresponding thresholds of object classes for any number of given features. Seath calculates the thresholds which allow the best separability in the selected features using the *Jeffries-Matusita distance J* [see 12]. This allows a better comparison of the feature analysis results to identify that feature which has the best separability. The Jeffries-Matusita distance measures the separability of two classes on a scale $[0 - 2]$ in terms of B:

$$J = 2(1 - e^{-B}) \qquad (6)$$

$J = 0$ implies that the two distributions are completely correlated and

$J = 2$ implies that the distributions are completely uncorrelated. For every feature, we can calculate the separability between the two classes using J. The features which have very high J value are the optimum features which characterise the classes [12]. An interpretation of the Seath results leads to the compilation of a classification ruleset.

Validation

A total of 621 road and non-road reference points collected from a 2,5 m pansharpened SPOT imagery and ground validated reference points were used for the accuracy assessment. A GPS was used for the collection of ground reference points.

Results

The results of the techniques

implemented in this study are presented in this section.

Multiresolution segmentation results

The multiresolution segmentation procedure was able to delineate tarred roads effectively. The outlines for road boundaries are illustrated in red in Fig. 1.

The summarised results of the Seath analysis are presented in Table 1. These results were used to develop the rule for the classification of roads. The rule involves the multiresolution segmentation of the aerial photography at a scale parameter of 100. A shape parameter of 0,3 and compactness parameter of 0,5 was used. This was subsequently followed by assigning thresholds to the NDWI, GLDV Ang. 2nd Moment Layer 4 (all dir.) and density features.

Roads classification ruleset

See Fig. 2 for an illustration of the roads classification ruleset.

The results of the road classification ruleset are presented in Fig. 3.

The final road extraction results are presented in Fig. 4.

The accuracy assessment results are shown in Table 2.

Discussion

The multiresolution segmentation implemented in this study enabled the effective delineation of roads. This segmentation procedure facilitated the accurate demarcation of roads which is a fundamental step before image classification. The features identified with Seath provided a fundamental step for feature selection and threshold identification. Seath is a very useful feature analysis tool that is able to statistically evaluate any number of features. Feature analysis and threshold identification is a precondition to good land cover road classification. The results from the Seath feature analysis tool optimised the development of a rule based object-oriented classification ruleset which minimises the misclassifications. These results show potential for accurate road classification. The approach used in this study provides a robust strategy for the development of successful road extraction rulesets. Furthermore, such a procedure permits ruleset transferability if the imagery is properly radiometrically normalised. The results also indicate that Haralick texture, spectral indices and shape parameters can be used to successfully classify roads. This is illustrated by the use of GLDV Ang. 2nd Moment, NDWI and density in this study. These features were statistically considered to be the best for road classification using the Seath software. The integration of object oriented classification and feature analysis proved useful in the extraction of roads. Some of the errors occurred in the classification as a result of obstructions caused by tree canopies covers on roads. These errors can be resolved using the pixel connecting function available in eCognition software. An overall accuracy of 93,6% was achieved in this study. This high level of accuracy is attributed to the techniques used in this study.

Conclusion

The study demonstrates that multiresolution segmentation is an effective means of delineating road networks. This segmentation technique is fundamental to successful object oriented road classification. The use of statistical tools for feature analysis and threshold identification enabled the development of a successful classification ruleset. It is also evident in this study that Haralick texture, shape density and spectral indices are important for effective road classification. This study showcases the efficacy of object-oriented classification in road extraction. The ability of object based classification in road classification is revealed by high accuracy of 93,6% obtained in this study.

References

[1] J I Igbokwe: October 3-7, 2004, 'Mapping of Regional Transportation Network with Medium Resolution Satellite Imagery' Paper presented at the 3rd FIG Regional Conference, Jakarta, Indonesia, (2004).

[2] J B Mena: 'State of the art on automatic road extraction for GIS update: a novel classification', Pattern Recognition Letters, 24, 3037–3058, (2003).

[3] P Gong and J Wang: 'Road Network Extraction from High Resolution Airborne Digital Camera Data', Annals of GIS, 3(1), 51-59, (1997).

[4] S Hinz, A Baumgartner and H Ebner: 'Modeling Contextual Knowledge for Controlling Road Extraction in Urban Areas.' Paper presented at the IEEE / ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, Rome, (2001).

[5] A Gruen and H Li: 'Road extraction from aerial and satellite images by dynamic programming'. ISPRS Journal of Photogrammetry and Remote Sensing, 50(4), 11-20, (1995).

[6] H Mayer, I Laptev and A Baumgartner: 'Multi-scale and snakes for automatic road extraction'., 5th Europ. Conf. on Computer Vision (Vol. 1, pp. 720-733), (1998).

[7] H Mayer and C Steger: 'Scale space events and their link to abstraction for road extraction' ISPRS J. Photogrammetry & Remote Sensing, 53(2), 62–75, (1998).

[8] A Baumgartner, C Steger, H Mayer, W Eckstein and H Ebner: 'Automatic road extraction based on multi-scale, grouping and context', Photogrammetry Engineering & Remote Sensing, 65, 777-786, (1999).

[9] Y Chen, W Su, J Li and Z Sun: 'Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas', Advances in Space Research, 43(2009), 1101–1110, (2009).

[10] Definiens: "Definiens Developer Reference Guide', Definiens, (2010).

[11] R Haralick and L Shapiro: 'Computer and Robot Vision' (Vol. 2). Reading: Addison-Wesley, (1993).

[12] S Nussbaum and G Menz: 'Object-based image analysis and treaty verification. New approaches in remote sensing- Applied to nuclear facilities in Iran': Springer, (2008).

Contact Paidamwoyo Mhangara, SANSA Space Operations, pmangara@sansa.org.za ◎