Combining multispectral aerial imagery and digital surface models to extract urban buildings

ADRIAN J.D. McNALLY[1] and SAMUEL J.P. McKENZIE[2]

[1] Fugro-BKS, 47 Ballycairn Road, Coleraine, County Londonderry, BT51 3HZ, UK; amcnally@fugro-bks.com.

[2] School of Environmental Sciences, University of Ulster, Cromore Road, Coleraine, County Londonderry, BT52 1SA, UK.

Abstract

This paper presents an automated classification of buildings in Coleraine, Northern Ireland. The classification was generated using very high spatial resolution data (10 cm) from a Digital Mapping Camera (DMC) for March 2009. The visible to near infrared (VNIR) bands of the DMC enabled a supervised classification to be performed to extract buildings from vegetation. A Digital Surface Model (DSM) was also created from the image to differentiate between buildings and other land classes with similar spectral profiles, such as roads. The supervised classification had the lowest classification accuracy (50%) while the DSM had an accuracy of 81%. The combination of the DSM and the supervised classification achieved an overall classification accuracy of 95%. Two spatial metrics (percentage of the landscape and number of patches) were also used to test the level of agreement between the classification and digitised building data. The results suggest that fine resolution multispectral aerial imagery can automatically detect buildings to a very high level of accuracy. Current space borne sensors, such as IKONOS and QuickBird, lag behind airborne sensors with VNIR bands provided at a much coarser spatial resolution (4m and 2.4m respectively). Techniques must be developed from current airborne sensors that can be applied to new space borne sensors in the future. The ability to generate DSMs from high resolution aerial imagery will afford new insights into the three-dimensional aspects of urban areas which will in turn inform future urban planning.

1.   Introduction

Urban areas currently contain around 50% of the world's population with the figure expected to rise to 70% by 2050 (Mesev, 2003). It is expected that 93% of this increase will occur in developing countries (Baudot, 2001), thus exerting pressure on resources and leading to substantial urban growth. The speed of change in urban areas demands regular monitoring to identify areas of high population density and plan for sustainable urban development.

Satellite imagery has been widely used for monitoring urban change over time with varying degrees of success (Rashed et al., 2010). Increases in the spatial resolution of satellite imagery have long been considered to be a panacea for urban remote sensing to differentiate between various forms of land use, for example between residential buildings and road networks (Welch, 1982; Donnay et al., 2001; Ke and Im, 2010). Other work has suggested that increases in the spectral resolution of satellite imagery will improve urban mapping (Heiden et al., 2007; Cavalli et al., 2008). However, there has generally been a trade-off between spectral and spatial resolution. While airborne sensors have usually had a substantially higher spatial resolution they were generally restricted to the visible bands of the electromagnetic spectrum. With recent improvements in the spatial resolution of space borne sensors there has been a concomitant loss of spectral resolution. For instance, very high spatial resolution satellite sensors, such as IKONOS and QuickBird, offer imagery with a spatial resolution of 4 m and 2.4 m respectively yet are restricted to the visible to near infrared (VNIR) bands of the electromagnetic spectrum. Recent advances in airborne imagery have led to the provision of VNIR bands on airborne sensors, thereby offering multispectral imagery with a very high spatial resolution. It is necessary therefore to determine the capabilities of such aerial imagery to automatically detect buildings and other land classes in urban environments. Only by identifying the intricate assemblage of urban land classes can planners prepare for sustainable urban growth. For instance, the substantial amount of biodiversity in private urban gardens (Loram and Gaston, 2007) is largely overlooked due to the coarse resolution of some satellite sensors.

While availing of the multispectral capabilities of high resolution aerial imagery to detect buildings, it is also necessary to appraise their potential in creating Digital Surface Models (DSM) of urban environments. The high spatial resolution of aerial imagery is likely to lead to significant improvements in modelling building elevation. Despite the potential of digital aerial imagery, few results on their use to create DSMs have been forthcoming for urban areas (Baltsavias and Gruen, 2003). Holland et al. (2008) used imagery from a Digital Mapping Camera (DMC) with a multispectral spatial resolution of 75 cm to classify a small area of Heathrow airport. The per-pixel classification approach alone achieved an accuracy of between 70% and 73%. However the classification accuracy increased to between 85% and 91% when combining the classification with a DSM

generated from the imagery. While the study was restricted to a small study area dominated by industrial buildings it highlights the potential of combining these techniques to automatically extract urban buildings. Both LIDAR (Laser-Induced Detection and Ranging) and Radar have been used to model the urban landscape. As active sensors both LIDAR and Radar are unaffected by shadow and are able to penetrate vegetation. Barnsley and Barr (2003) integrated LIDAR data with IKONOS imagery to detect urban landscape types in Cardiff, UK. The study identified that the combination of both datasets could achieve accuracies of between 88% and 95% when compared to Ordnance Survey data. The ability to couple high resolution remotely sensed data with precise elevation models clearly has a high potential for mapping and monitoring urban form and structure over time.

High resolution imagery from satellites, such as IKONOS, has been used to create DSMs for urban areas for many years. Ridley and Dowman (1997) investigated the potential of high-resolution satellite imagery (1 m) to update national maps in the UK. The models required grid spacing between 1 m and 3 m to extract heights from small buildings. The 1 m imagery had lower accuracy with root mean square error (RMSE) for maximum building height ranging from 1.5 m to 3 m and RMSE for mean building height ranging from 3.5 m to 6 m. Further work by Toutin et al. (2001) found that an automatically produced DSM from 1 m IKONOS imagery had errors of 5 m. However, the project used imagery for a small area of Reno, USA, that contained a low amount of urban land. Greater confidence in the models generated from remotely sensed sources will come from applications in denser urban areas.

While satellite imagery has a high temporal frequency and lower cost than aerial imagery it is imperative to investigate the potential of digital aerial imagery for extracting buildings. Applying techniques to very high resolution aerial imagery will provide insights into the processes and products that will inevitably emerge from new very high resolution satellite imagery.

In this paper we use multispectral aerial imagery with a spatial resolution of 10 cm to identify buildings in three ways. Firstly we create a DSM from the aerial imagery to differentiate features based on their height. Secondly we perform a supervised classification of the aerial imagery to differentiate between buildings and other land classes. Finally the DSM and the classification are combined to form a model of buildings as identified from the aerial photography. Each of these three models is tested for accuracy against digitised building polygons for the same study area.

2.   Methods

The aerial imagery was available for Coleraine, Northern Ireland. Coleraine is a large town with approximately 24,000 people and 9,700 household spaces (Northern Ireland Statistics and Research Agency, 2007). A digital aerial image was captured by light aircraft flying at a height of 1,100 m. The aerial image was captured by Fugro-BKS using a DMC in March 2009.

The imagery was collected with four spectral bands (Blue, Green, Red and Near Infrared) with a ground sample distance of 10 cm and a RMSE of ± 30 cm. To capture the complexity of the urban environment, fifteen study areas (600 m$\times$600 m) were distributed across the image to represent different urban landscapes including dense residential (DR), industrial (IN), rural (RU) and suburban (SU) land (Table 2). This study area size was used to minimise the amount of computer processing needed to perform a pixel-based classification of very high resolution imagery while investigating the models across a range of land cover types.

MATCH-T DSM software (inpho) was used to create a DSM from the aerial image with a nominal accuracy of ± 7 cm. A Digital Terrain Model (DTM) was used to represent the ground surface. The DTM was generated by Ordnance Survey Northern Ireland (OSNI) in 2003 and had a nominal accuracy of ± 1 m. Above ground features were identified by subtracting the DTM from the DSM to create a normalised DSM (Baltsavias and Zhang, 2005). The nDSM was set with a 3m height tolerance which caused both buildings and vegetation greater than 3m in height to be present on the surface.

The aerial imagery was classified into built and non-built land classes (Bahr, 2001) using a supervised classification. The maximum likelihood classification rule was used in the classification. Land classes such as grassland, trees, water bodies, impervious surfaces and bare soil were grouped into the non-built class while the built land class consisted of commercial, residential and industrial buildings.

The nDSM and classified aerial imagery were finally combined to create the final building model (nDSM_AER, Table 2). Buildings in the model were included based upon both their height and spectral values.

To test the accuracy of the model, building polygons were digitised from the aerial image and were used as the control dataset. Classification accuracy was performed by generating 150 random control points across the study areas to determine the level of accuracy between the aerial imagery and the three models. Additionally, two spatial metrics were used to assess the degree of correspondence between the models and the digitised buildings. The percentage of the landscape (PLAND) measured the amount of a target class, namely buildings, within a study area while number of patches (NP)

identified the count of building patches in both datasets. The spatial metrics were gene-rated using Fragstats spatial pattern software (McGarigal and Marks, 1994). A bivariate correlation was carried out to assess the level of agreement between the three models and the digitised building data using SPSS statistical analysis software. For full details of the procedures see McNally (2010).

3.   Results

The map identifies a high level of agreement between the digitised data and the combi-ned model generated from the nDSM and classified aerial imagery. A standard accuracy assessment determined an accuracy of 50% for the classified aerial imagery, 81% for the nDSM and 95% for the combined model (Table 1).

Classification	Bivariate correlation		Overall accuracy
	% landscape	Number of patches	
Supervised classification	92%	41%	50%
nDSM	93%	95%	81%
Spectral + DSM	94%	86%	95%

Table 1. Classification accuracies of the three building models.

The greatest correlation for PLAND was achieved for the combined model ($r > 0.94$, $p < 0.001$). The correlation between the combined model and the digitised data for NP weakened but remained strong and positive ($r = 0.86$, $p < 0.001$). There was greatest correlation between the digitised data and the nDSM for NP ($r = 0.95$, $p < 0.001$). The classified aerial imagery was subject to a large number of misclassified pixels which caused a non-significant and weak correlation ($r = 0.40$, $p = 0.13$) with the digitised data for NP. Table 2 identifies the values for PLAND and NP for each of the 15 sample areas.

Errors in the classification were apparent when vegetation was over a building's roo-fline. The digitised dataset ignored tree lines and maintained straight lines around the roof line while the supervised classification removed the vegetation class from the built class. This accounted for the loss of building between the datasets in some instances. Furthermore, pylons, garages, car parks, roads, air conditioning units and ducting were also included in the built class of the supervised classification due to a similar spectral response with buildings. These features were included in the classification due to the ability of the camera to detect very small features. Indeed, the very high spatial resolu-tion of aerial imagery is likely to cause significant difficulties in automatically extracting buildings when using pixel-based classifiers.

Study area	Percentage of the Landscape (PLAND)				Number of Patches (NP)			
	Digitised	nDSM	Aerial	nDSM_AER	Digitised	nDSM	Aerial	nDSM_AER
DR1	7.75	10.33	14.14	4.93	153	278	1338	473
DR2	11.75	14.75	31.22	6.35	285	468	929	822
DR3	12.72	14.92	22.47	9.25	179	274	881	673
DR4	13.89	20.19	39.50	13.29	268	447	975	641
IN1	4.31	13.78	16.67	6.81	37	324	903	181
IN2	13.03	21.08	33.53	14.76	112	276	1867	294
IN3	22.06	26.11	46.64	18.53	234	391	615	1017
IN4	10.31	11.25	30.36	8.75	93	249	729	188
RU1	0.06	0.42	14.56	0.04	2	22	1106	18
RU2	0.06	0.06	0.39	0.04	2	7	205	3
RU3	0.22	3.03	0.14	0.02	6	219	61	11
SU1	14.83	14.67	30.31	10.43	487	656	954	682
SU2	11.83	16.61	28.17	9.92	425	683	1104	1255
SU3	9.61	10.56	21.31	6.46	261	499	1825	475
SU4	15.97	20.53	33.03	11.62	555	882	1563	1080

Table 2. Spatial metrics between digitised data and the three building models.

4. Conclusions

The classification identified a high level of agreement between the combined model and the digitised data. The nDSM alone achieved a very high level of agreement between the classified buildings and the digitised data. The accuracy of the DSM (± 7 cm) is substantially higher than the accuracies achieved by Ridley and Dowman (1997) and Toutin et al. (2001) who reported RMSE values for building height ranging from 1.5 m to 5 m. The supervised classification differentiated between buildings and vegetation but had a very low classification accuracy. When taken alone, neither the supervised classification nor the nDSM achieved the same accuracy as the combined model. Furthermore, the combined model successfully omitted features such as road networks that are commonly classified as buildings. The accuracy of our combined model (95%) surpasses the accuracy achieved by Holland et al. (2008) who achieved an accuracy of 91% when combining an object-oriented classification with a DSM for Heathrow airport. The results achieved in this paper relate to areas with a more complex mix of different urban land classes than the large industrial buildings around airports. These results indicate the potential of using object-oriented classifiers on imagery with a very high spatial resolution.

Despite the high level of correlation between the combined model and the digitised data there were a number of discrepancies. Classification errors are apparent in some cases. The digitised data represented buildings that appeared to be residential, commercial or industrial. The combined model selected all features that had a strong spectral reflectance value and that were greater than 3 m in height. Garages, electricity pylons, articulated lorries and temporary storage containers were all therefore apparent in the

combined model. Bare trees were also included in some cases due to the lack of leaves during March. Furthermore, shadow represented a source of error especially along roof lines. This is a particular problem associated with remote sensing yet it appears to be compounded due to the high spatial resolution of the aerial imagery (Sanchez-Hernandez and Holland, 2007). The imagery was captured in the early afternoon with a sun angle between 31.5° and 27.7°. Consequently, the sun angle led to a different spectral response on each side of the roof ridge line. In addition to the problem of shadow, Sawaya et al. (2003) also found that similarities in spectral values between land classes can be compounded in very high resolution imagery.

Multispectral aerial imagery with a high spatial resolution, such as that used in this paper, should be tested for accuracy against other phenomenon such as forests, hedgerows and geological features. An accurate control dataset is vital in order to objectively determine the level of fit of the classification.

Software

QuickBird imagery was handled using Erdas Imagine 9.3 software while GIS analysis and map production were performed using ESRI ArcGIS 9.2. Correlations were conducted using Statistical Package for the Social Sciences (SPSS) version 13. MATCH-T DSM (inpho GmbH) software was used to create the correlated DSM. OrthoMaster (inpho GmbH) was used to create the orthophotography.

Acknowledgements

References

BAHR, H.-P. (2001) Remote Sensing and Urban Analysis, chap. Image segmentation for change detection in urban environments, Taylor and Francis: London, pp. 95–114.

BALTSAVIAS, E. and GRUEN, A. (2003) Remotely Sensed Cities, chap. Resolution convergence: A comparison of aerial photos, LIDAR and IKONOS for monitoring cities, Taylor and Francis: London, pp. 47–82.

BALTSAVIAS, E. and ZHANG, C. (2005) Automated updating of road databases from aerial images, International Journal of Applied Earth Observation and Geoinformation, 6, 199–203.

BARNSLEY, S. A. M., M. J. and BARR, S. L. (2003) Remotely Sensed Cities, chap. Determining urban land use through an analysis of the spatial composition of buildings identified in LIDAR and multispectral image data, Taylor and Francis: London, pp. 83–108.

BAUDOT, Y. (2001) Remote Sensing and Urban Analysis, chap. Geographical analysis of the population of fast-growing cities in the Third World, Taylor and Francis: London, pp. 225–241.

CAVALLI, R. M., FUSILLI, L., PASCUCCI, S., PIGNATTI, S. and SANTINI, F. (2008) Hyperspectral Sensor Data Capability for Retrieving Complex Urban Land Cover in Comparison with Multispectral Data: Venice City Case Study (Italy), Sensors, 8, 3299–3320.

DONNAY, J. P., BARNSLEY, M. J. and LONGLEY, P. A. (2001) Remote Sensing and Urban Analysis, chap. Remote Sensing and Urban Analysis, Taylor and Francis: London, pp. 3–18.

HEIDEN, U., SEGL, K., ROESSNER, S. and KAUFMANN, H. (2007) Determination of robust spectral features for identification of urban surface materials in hyperspectral remote sensing data, Remote Sensing of Environment, 111, 537–552.

HOLLAND, D. A., SANCHEZ-HERNANDEZ, C. and GLADSTONE, C. (2008) Detecting changes to topographic features using high resolution imagery, In Proceedings of the ISPRS Conference, Commission VI, WG VI/4, Beijing.

KE, Q. L. J., Y. and IM, J. (2010) Synergistic use of QuickBird multispectral imagery and LIDAR data for object-based forest species classification, Remote Sensing of Environment., 114, 1141–1154.

LORAM, T. J. W. P. H., A. and GASTON, K. (2007) Urban domestic gardens (X): the extent & structure of the resource in five major cities, Landscape Ecology, 22, 601–615.

MCGARIGAL, K. and MARKS, B. J. (1994) Fragstats: Spatial Pattern Analysis Program for Quantifying Landscape Structure. Version 2.0, Oregon State University, Corvallis.

MCNALLY, A. (2010) Comparison of air borne and space borne multispectral digital sensors for automatic building detection, Unpublished MSc Thesis, School of Environmental Sciences, University of Ulster.

MESEV, T. V. (2003) Remotely Sensed Cities, Taylor and Francis: London, 378 pp.

NORTHERN IRELAND STATISTICS AND RESEARCH AGENCY (2007) Towns and Villages Area Profile Report, May 2005 [Online]. Available from: http://www.ninis.nisra.gov.uk/mapxtreme_towns/report.asp?SettlementName=Coleraine&bandName=LargeTown, [Last accessed: 4 January, 2011].

RASHED, T., JÜRGENS, T., C.RASHED and JÜRGENS, C. (2010) Remote Sensing of Urban and Suburban Areas, Springer: Netherlands, 300 pp.

RIDLEY, A. P. M. A. P. M. J.-P., H. M. and DOWMAN, I. (1997) Evaluating the Potential of the Forthcoming Commercial U.S. High-Resolution Satellite Sensor Imagery at the Ordnance Survey, Photogrammetric Engineering and Remote Sensing, 63, 997–1005.

SANCHEZ-HERNANDEZ, G. G., C. and HOLLAND, D. A. (2007) Classification of urban features from Intergraph's Z/I Imaging DMC high resolution images for integration into a change detection flowline within Ordnance Survey, In Proceedings of the Urban Remote Sensing Joint Event, Paris.

SAWAYA, K. E., OLMANSON, L. G., HEINERT, N. J., BREZONIK, P. L. and BAUER, M. (2003) Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery, Remote Sensing of Environment, 88, 144–156.

TOUTIN, T., CHENIER, R. and CARBONNEAU, Y. (2001) 3D geometric modelling of IKONOS GEO images, Conference of the Joint ISPRS Workshop on High Resolution Mapping from Space 2001.

WELCH, R. (1982) Spatial Resolution Requirements for urban studies, International Journal of Remote Sensing, 3, 139–146.