OBJECT BASED IMAGE ANALYSIS FOR THE OPERATIONAL PRODUCTION OF HIGH RESOLUTION LAND COVER: THE DEVELOPMENT OF LANDBASE.

TEWKESBURY A., HINCHLIFFE G.
Infoterra Ltd., LEICESTER, UNITED KINGDOM

ABSTRACT

This paper follows the development of LandBase through the operational mapping of land cover from very high resolution (VHR) imagery and digital surface models (DSM). The approach utilised object based image analysis (OBIA) to achieve high levels of accuracy and automation. The variety of input data and end user requirements encountered provided many conceptual and technical challenges. With these in mind a thematically simple nomenclature was devised which may objectively describe the image content while remaining flexible to image variability. It is proposed that a ‘simple’ baseline dataset may form the framework for user specific analysis and thematic upgrades. The use of OBIA facilitated the semi-automatic extraction of large scale mapping with an MMU of 50m² to an overall accuracy of 86%.

1 BACKGROUND AND OBJECTIVES

Imagery is now established as a commodity mapping layer by both professional and casual users alike. The resolutions achieved by recent satellite and airborne systems allow the visual interpretation of a high level of information relevant to a range of different applications and users. Harnessing this information at large scales and making it available on demand is the ultimate goal of remote sensing and photogrammetry researchers.

The automatic analysis of images at a pixel level assumes that the spectral response and contextual relationships are representative of the target. This is rarely the case for pixel abstraction especially when dealing with higher resolution imagery where targets are represented by multiple pixels. To reduce this problem OBIA uses image segmentation to approximate natural units which in turn may give an improved representation of the spectral response, texture and contextual relationships. By attempting to model the contextual relationships of image objects with a knowledge base; an attempt is made to mimic the cognitive process of a human image interpreter. A review of the use of OBIA in remote sensing is given by Blaschke (2010).

Although many studies use this approach to classify VHR imagery alone (Stow et al. 2007; Li et al. 2010; Cleve et al. 2008) there is a distinct advantage achieved by integrating high resolution lidar DSM data; as can be seen in studies by Ke et al. (2010), Zhou & Troy (2008) & Aubrecht et al. (2009). For logistical and cost reasons lidar is often unavailable, especially outside of urban areas. However the latest generation of airborne digital sensors such as the Leica ADS40 and Vexcel UltraCAM have the capability to generate high quality photogrammetric digital surface data (Dazhao et al. 2007; Zhang et al. 2007; Craig 2005) with the advantage of a single capture source. For this reason this study aims to utilise imagery and digital surface data from VHR airborne imagery as a mandatory requirement and seamlessly integrate higher accuracy lidar data where available.

A nomenclature and MMU must be chosen that addresses wide ranging user needs and remain flexible and compatible with existing schemes. A balance must be made so that the selected nomenclature is operationally feasible and easily transferable and repeatable.

2 APPROACH AND METHODS

2.1 Input Data

Airborne imagery for much of the UK has been made available courtesy of the GeoPerspectives program (2010). GeoPerspectives is a joint initiative between Infoterra Ltd and Bluesky International Ltd. and builds on the existing UK Perspectives (2005) national aerial survey. GeoPerspectives utilises the latest digital image sensors to provide nationwide coverage of the following image products:

- 25cm RGB Ortho-image
- 50cm CIR Ortho-image
- 2m DSM
- 5m DTM (Digital Terrain Model)

The data capture is shared between Leica ADS40 II and Vexcel UltraCAM sensors with key urban centres also supplemented by Lidar DSM & DTM typically at 1m grid spacing.
The aim of this study is to develop an operational and flexible land cover classification from such data. The classification must be both cost effective and scalable up to national coverage.

The primary concern when undertaking this scale of classification is the data volume and variability. A complete data stack for the whole of England and Wales is approximately 10 Terabytes. With space required for interchange formatting, processing and classification output the total footprint of online storage required would be closer to 40 Terabytes. Further complications arise with respect to image stability especially in terms of phenological dynamics. Due to highly variable UK weather conditions a national coverage will consist of an amalgamation of varying capture dates (both season and year), sensor hardware and atmospheric and illumination conditions. With this in mind a robust classification strategy was required over theoretical data maximisation.

2.2 Nomenclature selection

An initial review of existing schemes such as LCM2000 (Fuller 2003), NLUD Land Cover (Harrison 2006), CORINE (Commission of the European Communities 1994), GMES CSL (Schrage 2005) and FAO LCCS (Gregorio & Jansen 2000) yields both significant overlap and a very large combined class list. Many of the classes contained require either highly skilled manual interpretation or multi-temporal imagery within a complex knowledgebase. The use of multi-temporal data has been demonstrated with high resolution imagery at a national scale (Fuller 2003; Smith & Fuller 2001) but is beyond the scope of this study. Relatively early on in this investigation it was clear that the adoption of an overriding, all inclusive nomenclature was not feasible for the initial development without significant manual interpretation.

In order to successfully achieve a robust, transferable classification a thematically simple, generic nomenclature captured at a very high resolution was investigated. The logic behind this is to aim towards a generic baseline that could be thematically upgraded to specific user requirements; be it complex land cover, land use or habitat maps. This baseline approach would allow wide area coverage while facilitating future higher order requirements.

A generic, baseline of available imagery supports a two step approach of image classification whereby primitive land cover objects are the focus of step one. Step two involves the applications of a domain specific knowledge base focusing on spectral, textural or contextual relationships. Barr and Barnsley (1997) propose this as a method of land use extraction via the graph theoretic relations of land cover objects.

Table 1 shows the selected nomenclature which was derived iteratively by selecting key, classes and then testing the feasibility of automatic extraction using a selection of images. Objectivity was a large consideration with ‘pure’ land cover remaining the focus to form building blocks for further analysis.

The ten ‘core’ level 2 classes extracted to a high resolution of 50 m² were the initial focus of the investigation with the aim of accurately quantifying the image content. The initial draft did not include the classes ‘sub-shrubs’ and ‘Tall Shrubs & Small Trees’ so that the vegetation classes ranged from herbaceous vegetation to the ‘woody’ shrubs and trees. This left potential gaps in the future compatibility, for example it was envisioned that shrubs would be classified based on texture, height signature and variability from their surroundings. Heath land dwarf shrubs do not conform to these criteria and remain ecologically significant and so were introduced as a separate class; Sub-shrubs. In some phonological states the identification of dwarf shrubs may be subjective in aerial imagery and as such an increased MMU of 1 hectare (ha) is applied to reduce interpretation redundancy. The inclusion of ‘Tall Shrubs & Small Trees’ give more information on woodland dynamics and scope to identify features such as orchards and hedgerows.

The three level 1 classes give context to level 2 classes for instance, herbaceous vegetation could be either aquatic (reed bed, marsh land), cultivated & built up area (pasture, cereal crop, amenity grassland) or natural & semi-natural (calcareous grassland, bracken).
2.3 Data pre-processing

The standard GeoPerspectives image product is an 8-bit radiometrically balanced, seamless mosaic for each acquisition block, typically a UK county. A traditional approach would be to classify atmospherically corrected raw flight lines using the same classification parameters and then compile the results. Initial testing indicated that the quality of the image calibration and seasonal and illumination variation would not support this technique.

The mosaiced image product was used allowing classification calibration that can be extended across an entire acquisition block. Due to memory limitations segmentation cannot be performed across very large images so the Ordnance Survey of Great Britain 10 km tiling scheme was adopted. With the 50cm CIR imagery acting as the primary information source the addition of 25cm RGB imagery adds very little additional detail while greatly increasing the processing logistics. As a result of this the RGB imagery is sub sampled to 50cm for segmentation and classification.

<table>
<thead>
<tr>
<th>Level</th>
<th>Class</th>
<th>Examples</th>
<th>MMU</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aquatic</td>
<td>Open ocean, inland water, marsh, fen, intertidal zone</td>
<td>1 ha</td>
</tr>
<tr>
<td></td>
<td>Cultivated &amp; Built Up Area (BUA)</td>
<td>Agricultural land, towns, villages and cities, transportation networks, extraction sites</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Natural &amp; Semi-natural</td>
<td>Ancient forest, moorland, mountainous regions, coastal margin</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sea &amp; Estuary</td>
<td>Open ocean, estuary</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Inland Water</td>
<td>Rivers, lakes, canals and ponds</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Artificial Surface</td>
<td>Tarmac, concrete, paving, gravels, construction and extraction sites</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bare Ground</td>
<td>Inland rock, coastal rock, sand, exposed soil and very sparse vegetation not part of a cultivation cycle</td>
<td>50 m²</td>
</tr>
<tr>
<td></td>
<td>Buildings</td>
<td>Buildings, monuments, bridges, raised transport, large pylons</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Herbaceous Vegetation</td>
<td>Grasses, ferns, cereals, brassicas, root crops, legumes, horticulture, exposed soil and very sparse vegetation clearly part of a cultivation cycle</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sub-shrubs</td>
<td>Dwarf shrubs and heathers</td>
<td>1 ha</td>
</tr>
<tr>
<td></td>
<td>Shrubs</td>
<td>Bushes, hedges, gorse and immature trees, mean height typically below 2m</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tall Shrubs &amp; Small Trees</td>
<td>Tall bushes, small trees such as apple, new plantations, mean height typically 2-5 m</td>
<td>50 m²</td>
</tr>
<tr>
<td></td>
<td>Trees</td>
<td>Mature broadleaf and coniferous trees, mean height typically greater than 5 m</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. LandBase nomenclature.
2.4 Image Classification

Image classification was performed within Trimble eCognition with the key attributes being the extraction of image objects via segmentation and the classification of those objects based on overlay and contextual operations.

The two classification levels of the nomenclature are represented in the classification by two image object levels, with the small, level 2 object ‘primitives’ depicting almost all visible delineations and the level 1 objects as broad areas or ‘pseudo’ parcels. Achieving such simulated land parcels requires a segmentation ruleset comprising of processes such as small object removal, very large object splitting and the fusion of highly textured regions. By Qualitative analysis it was found that no one standard segmentation algorithm or parameter could achieve the desired objects. Instead the objects were the result of a segmentation ruleset that creates initial objects and then systematically applies new segmentation and merging routines to targeted objects. Effectively a knowledge based segmentation is applied.

![Image object hierarchy used within LandBase](image)

Figure 1. Image object hierarchy used within LandBase

Post classification the image objects are re-segmented to create intelligent intra-class boundaries. This new image object level (Product Objects) forms the basis of a transferable, upgradable layer whereby features that fall within a LandBase class may be represented by an appropriate object. The complete object level hierarchy is shown in Figure 1 above.

A dichotomous approach to the classification was preferred as this enabled much easier calibration and trouble shooting. Where possible crisp classification decisions are made based upon single criteria for easier calibration. However multiple criteria are integrated in situations of low confidence. For instance vegetated objects at the border of an NDVI threshold are tested again against other image ratios and intensity.

The utilisation of photogrammetric height information adds significantly to the identification of elevated features. However due to resolution, feature smoothing, photogrammetric omissions and noise the classification cannot be made by height alone to the desired precision. For elevated features height is used as a guide and then refined with image content; for instance woodland may often ‘bleed’ into herbaceous fields due to DSM smoothing. This ‘bleeding’ can be reduced significantly by utilising object texture to help separate rough woodland from relatively smooth grassland. An example of such a correction can be seen in figure 2 below, where had height been the only factor the selected object would have been classified as shrubs. The inclusion of shrub features without any significant height signature can also be seen. The process of object ‘seed and grow’ proves useful especially for building roofs where the growing is constrained by both height and spectral difference to surrounding objects.
Where available, lidar data is used to further refine the classification. Buildings in particular are improved as the stronger footprint integrity allows the application of a squaring algorithm giving a visually more appealing result.

In order to achieve robust, transferable results; restraint was exercised in the ruleset development so as not to create scene specific correction rules. Hugely improved automatic results could be achieved for confined areas such as those achieved by Ehlers et al. (2006) but a balance has been made to simplify as much as possible giving minimal parameters and quick calibration.

To check classification consistency and to ensure quality all results are checked and manually edited for significant errors present at 1:5000 scale. Due to data volumes only a small percentage of errors will result in an un-manageable level of edits and so smart editing was devised to utilise the classification technology. If for instance a river is missed the objects need not be manually re-classified; instead the region is marked and re-classified for water only; using more liberal parameters. To avoid confusion, all figures (except figure 4) show fully automated results prior to manual editing.

In order to maximise the scope for further classification the output ‘product objects’ are attributed with several intrinsic contextual properties. Intrinsic features include class labels at two levels, area and height statistics. Contextual measures are made by calculating land cover proportions over a local area based upon a 50m radius and zonal area based upon the enclosing parent level 1 objects. The local land cover proportion calculation is illustrated in figure 3.
Figure 3. Local cover statistics calculation within LandBase.
This level of object attribution allows the creation of ‘instant’ thematic maps such as those depicting building and tree density in figure 4. Furthermore they may form the basis of complex queries aiming to extract specific features such as urban green space.
3 RESULTS

The classification process has been successfully applied to approximately 10 000 km² throughout the UK with the image calibration process now reduced to only a few hours per acquisition block. Examples of unedited classifications performed in Chester and Newcastle can be seen in figures 5 & 6.

A combined desk based accuracy assessment of Chester and Maidstone yielded an overall accuracy of 86%, see table 2 for the complete error matrix.
Table 2. LandBase Error Matrix.

4 CONCLUSION AND FUTURE PLANS

An operational, widespread classification process has been achieved that can give a highly accurate description of the landscape essential for monitoring and planning purposes. The derivation of a high resolution, thematically simple base map such as LandBase could form the framework for further analysis. By capturing entities that approximate simple, objective features it is proposed that further, user specific requirements may be derived by modelling the human cognitive processes used in image interpretation. There is much work to be done to realise such a prospect where there is a shift away from spectral classification towards one based largely upon spatial analysis. Figure 7 shows an example using LandBase to automatically extract gardens from the spatial relationships alone. Such a process is uncoupled from sensor characteristics as only the intrinsic spatial properties and neighbourhood relationships are taken into account. Combining spatial analysis routines with auxiliary information such as satellite data could prove very powerful indeed.
Figure 7. Gardens extracted automatically using LandBase.

REFERENCES


Schrage, T., 2005. Integrated Project geoland - Core Service Land Cover, Service Portfolio.


Zhang, B. et al., 2007. NEXT GENERATION AUTOMATIC TERRAIN EXTRACTION USING MICROSOFT ULTRACAM IMAGERY. ASPRS Proceedings.


The standard GeoPerspectives image product is an 8-bit radiometrically balanced, seamless mosaic for each acquisition block, typically a UK county. A traditional approach would be to classify atmospherically corrected raw flight lines using the same classification parameters and then compile the results. Initial testing indicated that the quality of the image calibration and seasonal and illumination variation would not support this technique. The mosaiced image product was used allowing classification calibration that can be extended across an entire acquisition block. Due to memory limitations segmentation cannot be performed across very large images so the Ordnance Survey of Great Britain 10 km tiling scheme was adopted. With the 50cm CIR imagery acting as the primary information source the addition of 25cm RGB imagery adds very little additional detail while greatly increasing the processing logistics. As a result of this the RGB imagery is sub sampled to 50cm for segmentation and classification.