AN ADAPTIVE TEXTURE SELECTION FRAMEWORK FOR ULTRA-HIGH RESOLUTION UAV IMAGERY

Joshua Kelcey*, Arko Lucieer

* School of Geography and Environmental Studies, University of Tasmania, Private Bag 76, Hobart, TAS 7001, Australia; josh.kelcey@utas.edu.au, arko.lucieer@utas.edu.au

Commission ICWG I/5

KEY WORDS: Texture, Classification, Vegetation

ABSTRACT:

The capacity for additional textural derivatives to compensate for the lack of broader spectral sensitivity of consumer grade digital cameras is established within a UAV context. A texture selection framework utilising random forest machine learning, was developed for application with ultra-high spatial resolution UAV imagery limited to the visible spectrum. The framework represents an adaptive approach, providing a rapid assessment of different texture measures relative to a specific user-defined application. This framework is illustrated within the context of UAV salt marsh mapping. This study highlights the importance of texture selection for improving classification of UAV imagery exhibiting high local spatial variance.

1 INTRODUCTION

1.1 UAV Remote Sensing

The unmanned aerial vehicle (UAV) represents an important, emerging platform within the field of remote sensing. The UAV holds strong appeal to researchers due to both their generated data characteristics and their flexibility in application. The low operational altitude of UAV platforms results in the generation of ultra-high resolution geospatial data (Dunford et al., 2009), while their reduced physical size allows for their rapid deployment, improving their capability to exploit limited windows of opportunity.

The low-cost nature of UAV components offers researchers the opportunity to develop a remote sensing platform to specifically accommodate their own research niche. As UAV technology matures, cost will inevitably fall and accessibility to high performance sensors will improve. As current technology stands however, the miniaturisation necessary to develop high quality, small low weight multispectral or hyperspectral sensors appropriate for UAV systems threatens to place high quality sensors out of reach of budget-limited research groups. The alternative to more advanced, but expensive sensors is to use standard consumer digital cameras.

Standard consumer digital cameras offer UAV researchers an accessible sensor system at the cost of some limitations. Consumer cameras remain capable of generating high quality digital data at a price point that is magnitudes of order lower than more advanced sensors. However the fundamental, potentially crippling disadvantage of consumer digital cameras is their limitation to the visible spectrum.

It is well established that different ground cover classes can be identified by spectral reflectance characteristics. As these spectral characteristics typically fall outside of the visible spectrum, limiting measurements to the visible spectrum effectively excludes complicated spectral analyses.

1.2 Ultra High Resolution Data

Another aspect of the UAV which appeals to researchers is the ultra-high spatial resolution of generated data. The low operational altitude of UAV’s can result in image data at spatial resolutions below 5cm, allowing the clear delineation of fine scale groundcover features. Ultra-high resolution imagery, while vastly increasing resolving power, also requires advanced image processing and remote sensing classification techniques.

Classification techniques rely strongly upon two assumptions for successful class delineation: low intra-class variability and a high inter-class variability (Foody, 1999). Coarser spatial resolution facilitates these assumptions through spectral mixing, thereby generalising the spectral response of groundcover classes. The improved resolving power of UAV imagery, however, allows a clearer separation of spectrally distinct components of groundcover class (e.g. vegetation components including leaves, branches and underlying earth visible through gaps in the canopy). This reduction in spectral mixing also results in a corresponding increase in intra-class spectral variability (Puissant et al., 2005). Relying solely upon the spectral characteristics alone for classification, particularly where groundcover classes have shared components, becomes increasingly unreliable. This has resulted in the adoption of image analysis techniques capable of accounting for the high intra-class variability of ultra-high resolution spatial data. Two readily adopted approaches is the implementation of a geospatial object based image analysis approach and the quantification of local variability through the use of image texture.

1.3 Geospatial Object-based Image Analysis

Geospatial object based image analysis (GEOBIA) is an evolution of the traditional pixel-based approaches that shifts the focus from discrete pixels as the primary analysis unit to primitive image objects. Object images are created from individual pixels grouped together into objects by their degree of homogeneity threshold preset by a user (Jones et al., 2011). The process works by iteratively growing objects by grouping together neighbouring homogeneous pixels. A pixel is added to an adjacent object only if such an addition does not decrease object homogeneity past the set threshold. The higher the threshold for an acceptable level of object heterogeneity is set, the larger primitive image objects are generated. As the emphasis of this paper is upon texture layer selection, GEOBIA is implemented rudimentarily, but consistently across the study.

978-1-4799-1114-1/13/$31.00
©2013 Crown 3883 IGARSS 2013
1.4 Image Texture

Another approach to analysing ultra-high spatial resolution imagery is the adoption of image texture. Image texture may be broadly defined as the replications, symmetries and patterns within the tonal variation that occurs across an image (Woodcock and Strahler, 1987). Texture is typically a generalisation of image data structure derived from local spatial statistical measures, whereby differing properties may represent different aspects of texture.

Studies have demonstrated the fundamental value of additional textural quantification to overcome the problems of high local variance when analysing remote sensing imagery. Tuttle (2006) utilised first and second order texture for habitat mapping. Shinde (2012) implemented local binary pattern operators for crop identification. Chen (2007) investigated wavelet-based texture for groundcover classification.

In addition to generalising ultra-high resolution imagery, texture may also provide potentially important derivatives when analysing imagery with the reduced spectral set of a consumer digital camera. The quantification of local spatial variation provides an additional source of information which may provide some compensation for the lack of spectral information.

From the perspective of remote sensing classification, selection of texture measure is driven by two specific requirements: the improvement of classification accuracy while maintaining a minimal dataset size practical for both modelling and overall processing speed. A comparative analysis of the performance of different textural properties requires an approach capable of comparing a large number of texture measures. Decision trees are a non-parametric technique capable of modelling dimensionally large, complex datasets of this nature.

1.5 Random Forest

Random forest is an ensemble machine learning technique used primarily for data classification and regression modelling, and was first introduced by Breiman (2001). A random forest is an ensemble approach constructed from many individual decision trees. Each decision tree is constructed using bootstrap aggregation, a technique whereby random subsets of input variables are used to construct each individual tree. The use of a subset both reduces correlation between individual trees, as well as increasing processing time. For each tree, two-thirds of input samples are randomly chosen to serve as training data to construct the tree, while the remaining third is used for cross-validation.

As part of its assessment, random forest measures the importance of each variable as an extension of its cross-validation process. This is achieved by substituting the variable with completely random values, in effect modelling its absence, and consequently utilising the resulting change in accuracy as a basis for measuring the relative importance of the variable to the classification. In this way, the importance of specific textural characteristics relative to a classification task may be assessed.

One noted issue with random forest variable importance, however, is that it has a known bias towards correlated predictor variables (Genuer et al., 2010). Therefore when modelling textural characteristics derived with similar settings, the listing of variable importance makes no distinction of variables that are redundant due to high correlation with variables of greater importance.

1.6 Paper Aims

The central theme of this paper is to investigate the potential of random forest machine learning for the selection of optimal overall textures. In addition to this exploration, this paper will present a framework that utilises random forest machine learning for the rapid identification of the most relevant overall specific texture measures for UAV generated, ultra-high resolution imagery. The paper focuses upon a common approach to deriving texture, the grey level co-occurrence matrix (GLCM). Data-reduction techniques are further applied to the variable importance results of the random forest model to filter redundant texture layers to obtain a minimal class-specific number.

2 METHODOLOGY

2.1 Study Site and Data Collection

The Ralphs Bay area in Tasmania, Australia is a stretch of sheltered coastline that provides suitable habitat for cold-temperate salt marsh vegetation (see Fig.1). The distribution of salt marsh vegetation typically exhibits a zonation effect along gradients of water logging and/or saline tolerance, typically dictated by local topography. The lowest areas may gather salt water, excluding vegetation entirely. Lower, more disturbed areas are dominated by herbaceous vegetation, while higher less disturbed areas may allow the establishment of small woody species.

Figure 1: Ralphs Bay in Tasmania. QuickBird imagery illustrating salt marsh study site

UAV salt marsh imagery was generated for this study using a MikroKopter Oktokopter1 frame mounted with a canon EOS 5D digital SLR camera. Fluorescent discs served scattered throughout the study site served as ground control points. These ground control points were measured with a Leica dual-frequency Base RTK GPS. Four dominant land cover classes were identified: three plant genera and one bareground groundcover class (see Fig. 2).

The identified vegetation classes were Sarcocornia sp., Tecticornia sp. and Gahnia sp.. Sarcocornia sp. and Tecticornia sp. are both halophytic succulent plants. Sarcocornia sp. is a herbaceous species that forms groundcover mats whilst Tecticornia sp. is a woody shrub that establishes in topographically higher areas. Gahnia sp. is a species of halophytic sedge whose relatively lower tolerance for salt-water inundation restricts its distribution furtherest away from the waters edge. A single bareground class was identified as water saturated mudflats.

2.2 UAV Orthomosaic Construction

The ultra high resolution UAV imagery was processed using PhotoScan2 to generate a georeferenced orthophoto of the study site (see Fig. 3).

1http://www.mikrokopter.de
2http://www.agisoft.ru
Gahn. sp.       Sarc. sp.         Tect. sp.         Mudflat

Figure 2: Terrestrial imagery of vegetation and non-vegetation salt marsh classes.

Figure 3: Generated UAV orthophoto. Zoomed subset illustrates the level of detail achieved with a spatial resolution of 1cm.

Figure 4: False colour composite of representative groundcover class tiles.

From this orthophoto, a mosaic of representative classes were extracted for the 4 groundcover classes. Each class is represented by 4 distinct sample sites (see Fig. 4).

2.3 Texture Measures

2.3.1 Grey Level Co-occurrence Measure The grey level co-occurrence measure (GLCM) is a calculation of image texture using a kernel that moves iteratively across an image. For each pixel location, a matrix is derived from the kernel based upon the frequency of pixel value pair cooccurrence at a predefined user offset (Haralick, 1973). From this frequency co-occurrence matrix, several different textural characteristics can then be derived from this co-occurrence matrix which include measures of variance, entropy and homogeneity. Modifying the size of the kernel sets the scale at which texture is derived. All GLCM texture measures were calculated using kernel sizes ranging from 3 to 31.

2.4 Random Forest and Texture Layer Filtering

Random Forest was implemented using the R statistical package and is controlled by the user by only 2 settings: the total number of decision trees in the ensemble, and the number of randomly selected variables used to determine the binary split at each individual node. The number of decision trees was set 4,000, while the number of variables to use at each node followed an established convention of using the square root of the total number of available variables.

A list of the overall variable importance in a descending order, was extracted from the random forest model. The correlation between texture layers derived from the class tile mosaic are iteratively compared with measures identified in the class lists. A correlation threshold of 0.8 is set, whereby a texture layer is removed from the overall texture list if it’s correlation with another texture layer of greater recorded importance exceeds this threshold.

Once the correlated, redundant texture layers are removed from the list, the final step is to reduce the overall list to a minimal number of textural measures. This was achieved by iteratively generating random forest models, adding successive texture layers based upon their recorded importance to the model, and recording the subsequent change in overall OOB error rate. Once an optimal overall OOB error rate is achieved, the remaining texture list is discarded.

3 RESULTS

3.1 Random Forest Model: Class Mosaic Digital Numbers

<table>
<thead>
<tr>
<th>Class</th>
<th>Gahnia</th>
<th>Mudflat</th>
<th>Sarcocornia</th>
<th>Tecticornia</th>
<th>ClassError</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gahnia</td>
<td>3191</td>
<td>590</td>
<td>192</td>
<td>1027</td>
<td>36.18</td>
</tr>
<tr>
<td>Mudflat</td>
<td>266</td>
<td>4582</td>
<td>21</td>
<td>131</td>
<td>8.36</td>
</tr>
<tr>
<td>Sarcocornia</td>
<td>210</td>
<td>124</td>
<td>4312</td>
<td>354</td>
<td>13.76</td>
</tr>
<tr>
<td>Tecticornia</td>
<td>1472</td>
<td>445</td>
<td>620</td>
<td>2463</td>
<td>50.74</td>
</tr>
<tr>
<td>Acc.:</td>
<td>72.74%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Random forest confusion matrix generated from class mosaic DN data only.

Table 1 illustrates the initial random forest model confusion matrix generated from class tile mosaic data digital numbers only. The classification performance is poor, with almost 50% of the data misclassified. The relatively spatially uncomplicated mudflats records the highest overall class accuracy, while the highest misclassification error rate occurs between the two most spatially more complicated classes: Gahnia sp., and Tecticornia sp..

3.2 GEOBIA : Preliminary DN Analysis

A preliminary GEOBIA was conducted upon the imagery using only the digital numbers (DN’s). The salt marsh image was first segmented into geospatial objects using the eCognition image analysis software package. Segmentation was conducted at object scales between 100 and 300 at increments of 50. The corresponding image objects were extracted and a random forest model generated from their spectral characteristics.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Overall Accuracy</th>
<th>Mudflat</th>
<th>Gahnia</th>
<th>Sarcocornia</th>
<th>Tecticornia</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>75.1527</td>
<td>95.76</td>
<td>15.23</td>
<td>96.07</td>
<td>61.59</td>
</tr>
<tr>
<td>150</td>
<td>75.4761</td>
<td>98.62</td>
<td>30.24</td>
<td>94.15</td>
<td>62.86</td>
</tr>
<tr>
<td>200</td>
<td>83.0892</td>
<td>98.05</td>
<td>35.45</td>
<td>92.8</td>
<td>83.12</td>
</tr>
<tr>
<td>250</td>
<td>81.8278</td>
<td>99.36</td>
<td>54.13</td>
<td>98.31</td>
<td>68.11</td>
</tr>
<tr>
<td>300</td>
<td>84.6395</td>
<td>98.68</td>
<td>51.76</td>
<td>98.31</td>
<td>76.88</td>
</tr>
</tbody>
</table>

Table 2: Overall and class-based accuracy of GEOBIA based on image DN’s

Table 2 illustrates the overall and class accuracies of the 5 segmentation results. Larger objects improved classification results for spatially complicated classes of Tecticornia sp. and Gahnia sp., improving the overall accuracy.

A random forest model was generated for all classes using all GLCM texture measures. Lists of texture measures in order of descending overall importance were derived from the random forest
model. These lists were further filtered to generate a list of optimal texture measures. An iterative analysis of additional texture layers to the model found that five additional texture layers effectively reduced the overall OOB error to under 1%.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Overall Accuracy</th>
<th>Mudflats</th>
<th>Gahnia</th>
<th>Sarco</th>
<th>Tect.</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>81.958</td>
<td>97.1</td>
<td>57.88</td>
<td>97.96</td>
<td>68.69</td>
</tr>
<tr>
<td>150</td>
<td>83.5164</td>
<td>97.31</td>
<td>55.24</td>
<td>96.61</td>
<td>74.81</td>
</tr>
<tr>
<td>200</td>
<td>85.1318</td>
<td>93.34</td>
<td>61.2</td>
<td>97.11</td>
<td>78.87</td>
</tr>
<tr>
<td>250</td>
<td>86.2142</td>
<td>99.36</td>
<td>71.94</td>
<td>99.15</td>
<td>73.25</td>
</tr>
<tr>
<td>300</td>
<td>87.5326</td>
<td>98.68</td>
<td>73.93</td>
<td>96.61</td>
<td>78.62</td>
</tr>
</tbody>
</table>

Table 3: Random forest confusion matrix generated from class mosaic spectral data only.

The identified texture layers were generated for the entire orthophoto, and textural information incorporated into the previously segmented objects for random forest object classification. Table 3 illustrates the results of GEOBIA with additional texture layers. While texture improved classification for all classes, Gahnia sp. was the most improved class, with per classification accuracy at the highest scale for DN only classifications comparable with the lowest scale segmentation with the addition of texture. The most optimal classification, GEOBIA with texture at a segmentation scale of 300, is presented (see Fig.5).

**Figure 5:** Final GEOBIA with texture classification of salt marsh vegetation

4 CONCLUSION

This study demonstrates that additional texture derivatives can help compensate for analyses of reduced spectral information. However, this remains true only for classes that are spatially complicated enough to benefit from the addition of texture. This is evident in the results, with GEOBIA alone capably handling the less spatially complicated mudflat and Sarcocornia sp. groundcover classes. Further analysis of spatial characteristics of these classes may provide insight for the potential of a spatial variability threshold from which to gauge the relevancy of additional texture. Furthermore, the results suggest that the selection of overall texture was most relevant for the lowest performing groundcover class Gahnia sp. In this regard, a more balanced approach may be to investigate the importance of texture on a per-class basis, to ensure all classes benefit adequately from the inclusion of texture.

In addition to improved overall accuracy, another important advantage that texture derivatives may provide is a reduction in primitive object size. In considering the accuracy of GEOBIA, shape is a fundamental factor. However, in the DN only classification, it was found that larger, more generalised objects were required to improve accuracy. The results suggest that the ability of texture to generalise spatial complexity may aid in driving down the minimum required object primitive size. Lowering the limit of object primitive size, and potentially avoiding overgeneralisation, may lead to a more accurate fitting of the boundaries between differing classes.

The implementation of texture often remains an arbitrary choice with little exploration of different settings. This is further complicated by the vast number of texture measures available, of which the GLCM is but one. This absence of selection methodology can be attributed to a number of factors including the sheer overwhelming number of methods and settings availability, the relative dependence of texture measures upon image characteristics (including image scale, the number of spectral bands, image context and class definition), the absence of an a priori approach for texture selection, as well as the high rate of classification improvement of texture inclusion even when arbitrarily applied. Until these issues are robustly addressed and a deeper link is established between texture and the biophysical characteristics being measured, texture selection methods may remain limited to cumbersome comparative approaches.

REFERENCES


