Super-resolution mapping of rural land cover objects from fine spatial resolution satellite sensor imagery

M. W. THORNTON\textsuperscript{1}, P. M. ATKINSON\textsuperscript{1} and D. A. HOLLAND\textsuperscript{2}
\textsuperscript{1}School of Geography, University of Southampton, Highfield, Southampton SO17 1BJ, UK
\textsuperscript{2}Research and Innovation, C530, Ordnance Survey, Romsey Road, Maybush, Southampton S016 4GU, UK

Abstract

Mapping rural land cover features, such as trees and hedgerows for ecological applications is a desirable component of the creation of cartographic maps by the Ordnance Survey. Based on the notion of spatial dependence, super-resolution mapping can be used to increase the spatial resolution of satellite sensor imagery and provide increased mapping accuracy of such features. 2.6 m spatial resolution Quickbird imagery was first was soft classified using a supervised fuzzy $c$-means algorithm and then a simple super-resolution pixel-swapping algorithm was applied. Sub-pixels within pixels were iteratively swapped until the spatial correlation between sub-pixels for the entire image was maximised. Mathematical morphology was used to suppress error in the super-resolved output, increasing overall accuracy. Field data were used to assess the accuracy of the resultant image. Soft classification accuracy was between 70\% and 80\%, resulting in the super-resolution method producing reasonably accurate results of between 50\% and 75\%.

1. Introduction

1.1 Land cover mapping

Rural land cover objects such as hedgerows and trees are important biological and ecological components of the natural world and their accurate mapping is important for land management strategies. Ordnance Survey, the national mapping agency in the UK, needs to map cartographically rural land cover objects, at a variety of scales (local, regional and national).

A generic definition of a rural land cover object is extremely hard to achieve (Baudry and Bunce, 2001). In most cases, these features have indeterminate boundaries and complex intrinsic geometric properties (e.g. they vary greatly in width, height and, perhaps most importantly, length) (Boutin, \textit{et al.}, 2001). Classification of remotely sensed imagery is one possible method of mapping these features. Research into increasing classification accuracy has often focused on the problem of mixed pixels (Smith \textit{et al.} 1990; Fisher, 1997; Cracknell, 1998; Bardossy and Samaniego, 2002) where pixels within an image represent more than one land cover on the ground. In most cases, allocating a mixed pixel to a single land cover class by traditional hard classification is inappropriate and will not provide a realistic or accurate representation of land cover (Foody, 1996).
Research into mixed pixels led to the development of a new breed of classification technique, subsequently known as soft classification (often termed fuzzy classification) (Bastin, 1997; Foody, 2002b). Soft classification techniques predict the proportions of each individual class within a single pixel by ‘unmixing’ all available land class information in a pixel and assigning a pixel to multiple classes. This is useful in relation to the prediction of small rural features, which often occupy only a small proportion of a pixel. Soft classification techniques predict the proportions of each individual class within a single pixel. However, the location of classes within the pixel is unknown. Predicting the position of individual classes within pixels is the goal of super-resolution mapping.

1.2 Super-resolution mapping

Many super-resolution techniques are based on an expectation of spatial dependence. This is the phenomenon that observations close together are more alike than those further apart (Curran, et al., 1998; Chiles and Delfiner, 1999). Super-resolution techniques increase the spatial resolution of a classified image by decomposing the pixel into smaller units, known as sub-pixels.

Super-resolution techniques are a relatively new development within the remote sensing field, and accordingly, there are various operational difficulties implementing them (Tatem, et al., 2003). One such difficulty is the effect of error in the soft classification, which is inevitable when using real imagery.

The objective of this paper is to use real soft classified fine spatial resolution satellite sensor imagery to super-resolve fine rural land cover objects, such as hedgerows and trees. To deal with the deleterious effects of error in the soft classified imagery, a mathematical morphology technique will be introduced and the benefits of the approach discussed.

2. Methods

There are three stages to the methodology. The techniques involved at each stage are discussed in this section.

2.1 Soft classification

Two soft classification algorithms were chosen - the supervised fuzzy c–means classifier (FCM) (Bezdek et al., 1984) and the linear mixture model (MM) (Settle and Drake, 1993). They were chosen because of their computational simplicity and common use in research (Bastin, 1997; Atkinson, 2001).

In the supervised FCM, in which the mean values for each class obtained from training data are manually input to the classifier, the classifier assigns pixels to multiple classes by measuring iteratively the Euclidean distance of each pixel to the mean value of each class in feature space.

Mixture modelling assumes that within a pixel, the spectral responses of each class are mixed linearly in proportion to the area that each class represents. Therefore, the spectral response of each pixel is a linear combination of the endmember spectra of each class and contains information on the proportion of each class found within
that pixel. Spectral endmembers for each class can then be used to unmix the pixels and predict the proportions of each class within each pixel.

Each of the soft classification methods above requires training (e.g. the selection of spectral endmembers representing pure pixels). For mixture modelling, for every class in the image a representative sample of pure pixels is required (e.g. 50-100 pixels per class).

Accuracy assessments on each of the soft classification techniques were performed. In each case, the root mean square error (RMSE) and Pearson’s product moment correlation coefficient ($r$) were calculated. The RMSE estimates the overall accuracy of the soft classification. It is calculated using (Tso and Mather, 2001):

$$\text{RMSE} = \sqrt{\sum_{k}^{K} \sum_{i}^{n} (z_k(x_i) - z_k^*(x_i))^2}$$

where $K$ is the number of classes, $z_k(x_i)$ is the proportion of class $k$ at the $i^{th}$ pixel location $x_i$ of the reference image and $z_k^*(x_i)$ is the value of the same pixel in the soft classified image.

The correlation coefficient ($r$) is a measure of the amount of association between the target and the predicted proportions and informs on the precision of the prediction. It is calculated per class by:

$$r_k = r_{uv} = \frac{C(uv)}{\sqrt{s^2(u) \times s^2(v)}} = \frac{C(uv)}{s(u) \times s(v)}$$

where $u = z_k(x)$ is the target image for class $k$, $v = z_k^*(x)$ is the predicted image for class $k$, $C(uv)$ is the covariance between $u$ and $v$, $s^2$ is the variance and $s$ is the standard deviation.

2.2 Pixel-swapping

The super-resolution pixel-swapping algorithm was first presented by Atkinson (1997) where it was used on simulated imagery. Land cover class proportions for each pixel are input to the pixel-swapping algorithm. In each pixel, a fixed number of sub-pixels are created, based on a suitable zoom factor. For example, if a zoom factor of five is used, then 25 sub-pixels in every pixel are created. The number of sub-pixels in each pixel remains fixed throughout the procedure. During this process, each sub-pixel is allocated to a single hard land cover class, such that the original class proportions in each pixel, output from the soft classification, are maintained. Furthermore, these proportions are maintained throughout the procedure.

In every iteration, the attractiveness ($A_i$) of each sub-pixel for a particular class $k$ is predicted as a distance-weighted function of its neighbours:
\[ A_i = A_k(x_i) = \sum_{j=1}^{n} \lambda_{ij} z_k(x_j) \]  

(4)

where \( n \) is the number of neighbours, \( z(x_j) \) is the value of the class, \( k \) (now constrained to be either 0 or 1), at the \( j \)th pixel location, \( x_j \), and \( \lambda_{ij} \) is a distance-dependent weight predicted as:

\[ \lambda_{ij} = \exp\left(\frac{-h_{ij}}{a}\right) \]  

(5)

where \( h_{ij} \) is the distance between pixel \( x_i \) (for which the attractiveness is desired) and the neighbour \( x_j \), and \( a \) is the non-linear parameter of the exponential distance-decay model. The algorithm ranks the attractiveness scores on a pixel-by-pixel basis. In each pixel, the locations of the least attractive (a 1 surrounded mainly by 0s) and the most attractive (a 0 surrounded mainly by 1s) sub-pixels are stored. If the attractiveness of the least attractive location is less than that of the most attractive location the sub-pixels are swapped - if it is more attractive no swap is made. One sub-pixel per pixel is swapped per iteration and the algorithm runs for either a specified number of iterations or until no swap is made.

2.3 Mathematical morphology and inverse-distance weighting

As previously discussed, the accuracy of the current super-resolution technique is affected adversely by error in the soft classification used as input. Therefore a two-step mathematical morphology (Heijmans, 1995; Serra, 1982) approach was used to remove small areas of error from the super-resolved output. Mathematical morphology techniques were developed to handle objects with a characteristic spatial structure, comprised of a specific arrangement of individual elements. Mathematical morphology techniques adjust and rearrange these individual elements to influence the state and appearance of the structure of objects. In this research, “closing and opening” operations were used. Initially, on a class-by-class basis, objects smaller than a specified size were eroded, by removing pixels around the objects’ edges. Then, the images were dilated, where sub-pixels were added back in. This removed small objects (misclassification) and left larger, accurate objects relatively untouched. These operations were achieved by passing a disc-shaped structuring element over the image, which is essentially a filter. The structuring element can vary in size and shape on a class-by-class basis.

Mathematical morphology removed small areas of error. However, this left “holes” in the image where sub-pixels were not assigned to a class. Therefore, a simple inverse-distance weighting algorithm was applied, to replace these “holes” with the most likely class, based on its neighbours.

Confusion matrices were constructed for each site to assess the accuracy of the super-resolution before and after mathematical morphology had been applied. Pixels used in the accuracy assessment were chosen from the area immediately surrounding the feature only. That is, an invisible rectangular “box” describing the maximum
extent of the hedgerow was placed around the hedgerow and only pixels found in the box were used in the accuracy assessment.

3. Field sites

Three field sites (QB1-QB3; Figure 1) were chosen in the area around Burton, near Christchurch, Dorset, UK.

![Figure 1: Original subset image: (a) site QB1, (b) site QB2, (c) site QB3.](image)

The field sites were chosen through a combination of visual inspection of satellite sensor imagery and aerial photography followed by a field visit. After site selection, a subset image for each site was created from Quickbird™ satellite sensor imagery acquired on 2nd June 2001. The imagery has four wavebands at a spatial resolution of 2.6 m, co-registered to the British National Grid. The sites were purposefully kept small (< 400 pixels by < 400 pixels) to characterise the scene accurately yet limit computational overheads.

For each site, which contained a combination of trees and hedgerows, ground data were acquired. At 5 m intervals along hedgerows, the actual position, width, height
was measured using a GPS and any additional notes (position of trees, break in feature or other anomalies) were recorded. In addition to the measurements at 5 m intervals, measurements were also taken where the feature deviated significantly from the norm (loss of cover, change of geometry). In addition, the location and areal extent of land cover classes in the scene were recorded. These data were acquired during June 2003 to minimise phenological and other seasonal differences between the imagery and ground data.

4. Analysis

4.1 Soft Classification

The FCM and MM soft classifiers were applied to the imagery. In the FCM, a fuzzy exponent of 2 was used. Figures 2, 3 and 4 show the individual class predictions from the FCM for each of the field sites, QB1, QB2 and QB3 respectively, as a set of greyscale images.

Figure 2: Proportion images for site QB1: (a) “hedgerow”, (b) “cereal”, (c) “woodland”
Figure 3: Proportion images for site QB2: (a) “woodland”, (b) “hedgerow”, (c) “non-ripe cereal”, (d) “cereal”.
Figure 4: Proportion images for site QB3: (a) “hedgerow”, (b) “woodland” (i.e. individual trees), (c) “cereal a”, (d) “cereal b”, (e) “cereal c”.

In figures 2-4, white represents pixels completely covered by a particular class (100%); black represents zero cover of a particular class. Greys, therefore, indicate areas of mixing between classes. A visual inspection of these images indicates that the woodland class was separated from surrounding classes adequately. In other classes, however, the existence of greys suggests confusion between classes.
Table 1 shows the RMSE and correlation coefficients for each of the field sites for each of the soft classification techniques.

**Table 1. Accuracy assessment (a) site QB1, (b) site QB2, (c) site QB3.**

<table>
<thead>
<tr>
<th>(a) Site: QB1</th>
<th>FCM</th>
<th>MM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 (hedgerow)</td>
<td>11.55</td>
<td>0.97</td>
</tr>
<tr>
<td>Class 2 (woodland)</td>
<td>11.30</td>
<td>0.97</td>
</tr>
<tr>
<td>Class 3 (cereal)</td>
<td>12.36</td>
<td>0.97</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(b) Site: QB2</th>
<th>FCM</th>
<th>MM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 (woodland)</td>
<td>14.91</td>
<td>0.94</td>
</tr>
<tr>
<td>Class 2 (hedgerow)</td>
<td>25.95</td>
<td>0.81</td>
</tr>
<tr>
<td>Class 3 (non-ripe cereal)</td>
<td>6.56</td>
<td>0.99</td>
</tr>
<tr>
<td>Class 4 (ripe cereal)</td>
<td>24.64</td>
<td>0.83</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(c) Site: QB3</th>
<th>FCM</th>
<th>MM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1 (hedgerow)</td>
<td>18.15</td>
<td>0.90</td>
</tr>
<tr>
<td>Class 2 (woodland)</td>
<td>20.03</td>
<td>0.88</td>
</tr>
<tr>
<td>Class 3 (cereal a)</td>
<td>14.73</td>
<td>0.93</td>
</tr>
<tr>
<td>Class 4 (cereal b)</td>
<td>12.21</td>
<td>0.95</td>
</tr>
<tr>
<td>Class 5 (cereal c)</td>
<td>10.63</td>
<td>0.97</td>
</tr>
</tbody>
</table>

The visual inspection is supported by the accuracy assessment, where the accuracy of the woodland class was higher than for other classes. In the less complex, 3-class QB1 site, there was little difference in accuracy between the FCM and MM. Both have accuracies of > 85% for each class. However, on the more complex sites, where there are both more classes and classes that are less spectrally separable from others, the accuracy of the MM is less than that of FCM. Accordingly, only the output of the FCM was used as input for the super-resolution mapping.

4.2 Pixel-swapping
The pixel-swapping algorithm was applied to the soft classified imagery using a zoom factor of 5 for each of the three field sites (figure 5).

Figure 5: Super-resolution output: (a) site QB1, (b) site QB2, (c) site QB3.

Table 2 shows the confusion matrix.
Table 2. Confusion matrices, super-resolved output: (a) site QB1, (b) site QB2, (c) site QB3.

(a) | Feature | Not Feature | Totals | PA (%) |
---|---------|------------|--------|--------|
Feature | 19      | 26         | 45     | 42.2   |
Not Feature | 5       | 40         | 45     | 88.8   |
Totals | 24      | 66         | 59     |        |
UA (%) | 79.2    | 60.6       |        |        |

(b) | Feature | Not Feature | Totals | PA (%) |
---|---------|------------|--------|--------|
Feature | 15      | 3          | 18     | 83.3   |
Not Feature | 1       | 17         | 18     | 94.4   |
Totals | 16      | 20         | 32     |        |
UA (%) | 93.7    | 85.0       |        |        |

(c) | Feature | Not Feature | Totals | PA (%) |
---|---------|------------|--------|--------|
Feature | 17      | 20         | 37     | 45.9   |
Not Feature | 4       | 33         | 37     | 89.1   |
Totals | 21      | 53         | 50     |        |
UA (%) | 80.9    | 62.2       |        |        |

Mathematical morphology and inverse-distance weighting were then applied to the super-resolved output (figure 6).
Figure 6: Super-resolved output with mathematical morphology applied: (a) site QB1, (b) site QB2, (c) site QB3.

In each case, the structuring element was disc-shaped and the size of the structuring element was varied on a class-by-class basis for each field site. The morphology was applied initially on all classes except the feature class (i.e. all non-feature error pixels were removed first) to maximise the availability of feature class pixels in the inverse-distance weighting step. For example, in field site QB1 a basic [0,1,1] structuring element was used (where no morphology was applied to the feature class, and the structuring element for the other 2 classes was of size 1). In the remaining sites, where there were more classes and more significant error, larger elements were used, for example, [0,2,2,2] on QB2 and [0,1,2,3,3] on QB3. The morphology was then re-applied, this time acting only on the feature class (to remove feature class error pixels e.g. [1,0,0] on QB1, [1,0,0,0] on QB2 and [1,0,0,0,0] on QB3).
A visual inspection of the images shows that trees within hedgerows resemble their initial shape and hedgerows reflect the approximate size and geometry of the features in the original subset images. The approximate widths of the hedgerows were: QB1, 3-5 m, QB2: 20-23 m, QB3, 1.8 - 3 m, and in most cases continuous features were predicted. In some cases, particularly in QB3, small gaps appeared within the hedgerows.

In areas of dense mixing on the ground, particularly found within some of the ‘cereal’ classes, error arises as a function of error in the soft classification. Table 3 shows the confusion matrices.

Table 3. Confusion matrices, super-resolved output with mathematical morphology applied (a) site QB1, (b) site QB2, (c) site QB3.

<table>
<thead>
<tr>
<th></th>
<th>Feature</th>
<th>Not Feature</th>
<th>Totals</th>
<th>PA (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>21</td>
<td>24</td>
<td>45</td>
<td>46.7</td>
</tr>
<tr>
<td>Not Feature</td>
<td>4</td>
<td>41</td>
<td>45</td>
<td>91.1</td>
</tr>
<tr>
<td>Totals</td>
<td>25</td>
<td>65</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>UA (%)</td>
<td>84.0</td>
<td>63.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>17</td>
<td>1</td>
<td>18</td>
<td>94.4</td>
</tr>
<tr>
<td>Not Feature</td>
<td>1</td>
<td>17</td>
<td>18</td>
<td>94.4</td>
</tr>
<tr>
<td>Totals</td>
<td>18</td>
<td>18</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>UA (%)</td>
<td>94.4</td>
<td>94.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>21</td>
<td>16</td>
<td>37</td>
<td>46.5</td>
</tr>
<tr>
<td>Not Feature</td>
<td>5</td>
<td>32</td>
<td>37</td>
<td>86.4</td>
</tr>
<tr>
<td>Totals</td>
<td>26</td>
<td>48</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>UA (%)</td>
<td>80.7</td>
<td>66.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5. Discussion

5.1 Soft Classification

Both classifiers displayed accuracies greater than 80%, yet both classifiers were less accurate where there was mixing within classes on the ground. This was most obvious in the more complex sites, such as QB2 and QB3, where there were both bare soil and cereals in the agricultural fields, even though only ‘cereal’ was used in training. Further, as Foody (2000) points out, the existence of untrained classes in the image may have led to reduced accuracy of the soft classification. Consequently, the accuracy of the pixel-swapping algorithm was affected, due to its reliance on accurate class proportion prediction.
5.2 Pixel-swapping

In each of the field sites several problems occurred. The primary problem was the appearance of error in the output. As the algorithm completed successive iterations, the algorithm erroneously swapped the error as if it represented actual land cover. After several iterations, the error began to take a structured form, and the accuracy of the pixel-swapping technique could not be increased.

There are several interleaved causes of error in the output. For example, in complex sites such as QB2 and QB3, there was class mixing on the ground. Such fuzziness and vagueness within the imagery caused problems during both soft classification and subsequently during pixel-swapping. The objects of primary interest, hedgerows and trees, were commonly spectrally similar to the background which made separating them a complex task.

Error in the ground data may have affected the assessment of the accuracy of the techniques. The phenology of the scene when the satellite sensor image was acquired is likely to differ slightly from the phenology of the scene when the ground data were acquired. Additionally, land cover features may have changed: the width of the hedge could have changed from the width of the hedge in the satellite sensor image, due, for example, to seasonal growth or hedgerow management practices. The GPS Position Dilution of Precision (PDOP) value, a measure of the current satellite geometry, affects the precision with which measurements are taken. PDOPs of between 2.2 and 4.0 were recorded which related to an average precision of 68%. Furthermore, georectification of the imagery is accurate to approximately 2 m. Given that the average width of some of the hedgerows was 3 m, this accuracy is critical to locating the hedgerow in the field. These problems had an impact on the accuracy assessment of the super-resolution output.

Despite the above problems, the result of super-resolution mapping is a map of small rural features at a spatial resolution of 0.52 m produced from space. The map reveals subtle details in these features (e.g., variation in width along their length) that are otherwise difficult to detect.

5.3 Mathematical morphology

A visual inspection of the super-resolved output, after mathematical morphology was applied, shows some improvements. In most cases, the mathematical morphology method removed large amounts of error from the super-resolved output. Table 3 indicates small increases in overall accuracy. In sites QB2 and QB3, where mixing between classes was extremely complex, particularly in the ‘cereal’ classes, some obvious error was still apparent even after application of the mathematical morphology. However, such error, which does not affect the prediction of the feature of interest, is of little concern.

5 Conclusion

The pixel-swapping algorithm is a simple and efficient super resolution technique. Initial testing on satellite sensor imagery yielded promising results, which, in addition to a novel mathematical morphology component to suppress error, provided visually
pleasing and moderately accurate results. This research has provided much information on the effectiveness of such a technique, and highlighted some of the problems associated with its operational implementation. This information will be used in future research in developing the algorithm for extraction of fine linear features.

Acknowledgements
The authors wish to thank the Ordnance Survey and the School of Geography, University of Southampton for funding this research.

6 References


ATKINSON, P.M., 2001, Super-resolution target mapping from soft classified remotely sensed imagery, Fifth International Conference on GeoComputation, University of Leeds, Leeds, CD-ROM.


