A LAND-COVER POST-CLASSIFICATION FILTER BASED ON NEURAL NETWORK END-NODE ACTIVATIONS

T. Van de Voorde, S. Somers, F. Canters
Centre for Cartography and GIS, Department of Geography, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussel
tvdoord, ssomers, fcanters)vub.ac.be

KEY WORDS: Mapping, Urban, Land Cover, Classification, Neural, High Resolution

ABSTRACT:

Supervised per-pixel land-cover classification techniques, when applied to Very High Resolution (VHR) satellite images, often fail to produce a useful land-cover product due to the noise (“salt-and-pepper effect”) that stems from pixel misclassifications. Improving classification algorithms might not solve this problem because it is mainly data-related, i.e. caused by spectral mixing, confusion or ambiguity in the image data. We propose a straightforward and easily implemented post-classification filtering technique to “brush-up” a per-pixel land-cover classification, making it more useful in combination with ancillary geographic datasets (e.g. census or cadastre maps). This filtering technique uses the activation levels of the end-nodes of the neural network that is used to produce the per-pixel land-cover classification. While relatively little improvement was noted when evaluating the results of the filtering on a pixel-by-pixel basis (Kappa Index of Agreement), the structural properties of the classification improved significantly, effectively resulting in a land-cover product that is more useful for further processing.

1. INTRODUCTION

Deriving land-cover maps from high-resolution satellite data is a common application for people using Earth observation data. Accurate maps that show what can be found on a particular part of the Earth's surface at a particular time are often very useful for professionals involved in the management of the human or natural environment. With the advent of Very High Resolution (VHR) satellite sensors, the production of even more detailed land-cover maps, suitable for planning purposes at the local or urban level, has become possible. However, the increase in spatial resolution has also lead to an information overload, making it more difficult to derive thematically correct and cartographically useful end-products.

A problem that often occurs with the results obtained from a per-pixel land-cover classification from VHR images, is the so-called “salt and pepper” effect: noisy or speckled results due to the misclassification of individual pixels or small pixel groups. This in turn occurs because of spectral contamination in the data registration process or spectral confusion between different classes, and cannot be easily overcome in the classification stage itself, because the problem is mainly data-related. Some classes, like roofing and asphalt, can simply not be spectrally distinguished with the limited number of available spectral channels of a VHR sensor. In highly fragmented landscapes (e.g. in urban areas) a large number of mixed pixels may occur, which cannot be unambiguously allocated to one land-cover class.

Some techniques have been proposed to overcome the problem of spatially noisy classifications. The image segmentation approach (Haralick and Shapiro, 1985), for instance, can be used to first cluster image pixels into similar regions, before the classification stage is entered. We propose a filtering technique that can easily be implemented after the classification stage, to make a land-cover map more useful for further processing.

2. BACKGROUND

Per-pixel, supervised image classification is about assigning individual pixels to one of a set of a priori defined information classes. This is usually accomplished with a classifier that is trained using a random set of training pixels per class. Each training vector contains the spectral radiance values of a small area on the Earth's surface (the image pixel), for a number of spectral channels (usually 4 for contemporary VHR sensors). The classifier applies a specific decision rule to assign each pixel in the image to one class, based on the pixel's position in the feature space, defined by the spectral variables.

Artificial Neural Networks (ANN) have been widely applied for satellite image classification since the early nineties, and their performance in comparison with more traditional, statistical approaches (e.g. maximum likelihood) has been extensively studied (Atkinson and Tatnall, 1997). A statistical classifier provides us with the probability that a pixel belongs to a certain class. Because each neural network’s output node corresponds to one pre-defined class, its activation level can also be considered as an indicator of the membership level of a pixel to that class (Foody et al., 1997).

Traditionally, in remote sensing classification, a pixel is assigned to the most likely class (in the case of a statistical classifier) or to the class that corresponds to the highest activated output node (in the case of a neural network). The other activations or probabilities are mostly ignored. Nevertheless this information can be very useful to either describe the uncertainty that is present in the classification or to improve the classification result in a post-classification phase. As will be shown in this paper, end-node activation levels for different classes can be used in the process of removing structural clutter from a land-cover map, obtained through classification of VHR sensor data.
3. FILTER METHODOLOGY

One possible way to remove structural clutter from a land-cover classification involves the application of a simple majority filter (Gurney and Townshend, 1983) or a more sophisticated spatial reclassification technique (Barnsley and Barr, 1996; Gong and Howarth, 1992; Wharton, 1982) within a moving window of fixed size. The use of kernel-based approaches, however, has a number of disadvantages, including the difficulty of selecting an optimal kernel size, and the fact that the kernel is an artificial construct that does not refer to the spatial units that occur in the land-cover scene. To avoid the problems related to the use of kernel-based filtering methods, Barr and Barnsley (2000) propose a reflexive mapping procedure that operates on individual regions, i.e. groups of adjacent pixels that are assigned to the same land-cover class by the classifier. The procedure merges regions that fall below an a priori defined, class-specific area threshold with the smallest neighbouring region that exceeds the area threshold. After each re-assignment the region-based topological structure of the image is rebuilt.

A possible disadvantage of a region-based filtering approach, as described above, is that classification uncertainty at the pixel level, which is partly responsible for the presence of structural clutter, is not taken into account in the re-assignment process. It is therefore interesting to introduce classification uncertainty in the filtering process, and to find out if, by doing so, more accurate land-cover classifications can be produced. In this paper we present the first, experimental results obtained with a reflective approach to structural filtering that makes use of per-pixel classification uncertainty. The method resolves small regions that fall below an a priori defined area threshold, just like in the region-based approach described above. However, the re-assignment of the pixels that are part of a small region to one of the classes present in the neighbouring regions is performed on a per-pixel basis, using information about the identity of the second most likely class to which the pixel is assigned, as well as information about the uncertainty of the second most likely class assignment, produced by the classifier. In this study we used the end-node activation levels for the different classes, produced by a neural network classifier, as indicators of classification uncertainty.

3.1 Data Input

The filter we propose requires several data items, acquired or derived from the classification. The starting point is a neural network land-cover classification of which we not only use the output node with the highest activation for each pixel, but also retain the second most activated end-node, and the actual activation levels of those nodes.

From the map that shows the highest activated end-node for each pixel (the hard classification output), we derive unique spatial pixel groupings, so-called regions, i.e. pixels of the same class that are adjacent to each other. We also calculate the area (in number of pixels) for each of these regions.

Although we rely on the output of a neural network classification, similar information might, of course, also be obtained by applying other, probabilistic or fuzzy classification approaches.

3.2 Filter operation

The filter examines each pixel as follows.

If the area of the region to which the pixel belongs is larger than, or equal to a user-supplied area threshold, the pixel's land-cover class (node with the highest activation) is left unchanged. Otherwise, if there is a sufficiently large neighbouring region (according to the area threshold), for which the highest activated node equals the second highest activated node of the examined pixel, then the examined pixel's land-cover type is changed to the class corresponding to its second most activated node. If this rule cannot be successfully applied, the pixel is assigned to the class of the neighbouring region with which its own region shares the most border pixels (i.e. the highest connectivity). If two neighbouring regions that belong to different classes meet the highest connectivity criterion, then the pixel is assigned to the class to which the largest of the two regions belongs.

The rationale behind the procedure proposed is that the edge pixels that are part of a small region that is to be removed from the classified image are very often mixed pixels, which are partly covered by the land-cover class of the neighbouring region. If this is confirmed by a high end-node activation level for that class then the pixel is re-assigned to the class of the neighbouring region. This neighbouring region will, of course, not necessarily be the one that has the strongest spatial connectivity with the region to which the target pixel belongs. As a result, not all pixels of the region to be removed will be automatically assigned to one and the same land-cover class, as in a pure region-based approach. Pixels with a high end-node activation level for one of the land-cover classes of the neighbouring regions will be individually re-assigned to one of these classes.

3.3 Case Study

We started developing the filter in the framework of a remote sensing research project funded by the Belgian Federal Office for Scientific, Technical and Cultural Affairs (OSTC). The main goal of this project is to improve spatial information extraction for local and regional authorities using very-high-resolution satellite data. Part of the research in this project involves deriving land-use maps at the urban level, suitable for urban planners. One of the possible strategies to derive urban land use from remotely sensed data is to first produce a land-cover map using a standard multispectral classification technique, and then try to infer land use from the morphological and topological properties of the obtained land-cover regions (Barnsley and Barr, 1997; Bauer et al., 1999). It is clear that the success of any approach of this type will depend on the quality of the initial land-cover classification. When we started testing several land-cover classification strategies we noticed that the results produced by per-pixel classification methods were stained with noise caused by misclassified individual pixels or small pixel groups. The region-based classifications we applied did not suffers from this effect, but also were not completely satisfactory. Therefore, after using standard majority filters unsuccessfully, the step towards more elaborate filtering was taken.

We did some initial testing of the filter described above on part of a neural network land-cover classification of a 4m resolution Ikonos Satellite image of a residential urban area. The classification consists of five classes: buildings (1), roads and parking (2), water (3), railways (4) and open spaces (5). The
The image on the right clearly shows that small artifacts can be removed with the rules described above, without using a moving window technique which suffers from edge-effects. Although the Kappa index (Rosenfield and Fitzpatrick-Lins, 1986) only increases from 0.56 to 0.58 after applying the filter, the filtered land-cover map has a much more consistent structure. This can be shown numerically by calculating the average fragmentation index (Monmonier, 1974) of the pixels of each class for the visually interpreted image (reference data), the per-pixel classification and the filtered image (table 1).

<table>
<thead>
<tr>
<th>Reference</th>
<th>Classification</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>379</td>
<td>5979</td>
</tr>
<tr>
<td>GYR. MN</td>
<td>6.96</td>
<td>1.40</td>
</tr>
<tr>
<td>CONTIG MN</td>
<td>0.74</td>
<td>0.24</td>
</tr>
<tr>
<td>PLADJ</td>
<td>93.31</td>
<td>77.33</td>
</tr>
</tbody>
</table>

Table 2: Selected metrics for class 1 (built-up)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Classification</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>34</td>
<td>2854</td>
</tr>
<tr>
<td>GYR. MN</td>
<td>68.14</td>
<td>1.41</td>
</tr>
<tr>
<td>CONTIG MN</td>
<td>0.94</td>
<td>0.18</td>
</tr>
<tr>
<td>PLADJ</td>
<td>98.07</td>
<td>93.57</td>
</tr>
</tbody>
</table>

Table 3: Selected metrics for class 5 (open areas)

4. CONCLUSIONS

Filtering a per-pixel land-cover classification using information that is implicitly present in the output of a neural network classification improves the spatial structure of the obtained land-cover map, both visually, and in terms of region-based class metrics. The technique may be very useful to improve the structural properties of a land-cover classification before attempting to infer land use from it.

So far only initial tests have been carried out with the filter. More work is needed to study its functioning in more detail and to improve its performance. One of the most important issues that need to be addressed is the optimal choice of threshold values. The user has to control the input thresholds (minimum area and activation level) carefully, to obtain equilibrium between noisy and over-generalized results. So far only general, class-independent threshold values were used. Better results might be obtained with class-specific threshold values.

The filter we propose here is part of an ongoing effort to improve urban land-use/land-cover extraction from very-high-resolution remotely sensed data. It represents one of the steps in the process of transforming per-pixel land-cover classification results into land-use related information that is meaningful for local and regional planning authorities.

References


Acknowledgements

The Belgian Federal Office for Scientific, Technical and Cultural Affairs is gratefully acknowledged for funding the work that is presented in this paper (project SR/00/02).