A CONTEXT BASED TECHNIQUE FOR SMOOTHING OF DIGITAL THEMATIC MAPS

Ben Yee
Dean Turpin
MacDonald, Dettwiler and Associates Ltd.
Richmond, B.C.

Evert Kenk
Mark Sondheim
B.C. Ministry Of Environment
Victoria, B.C.

ABSTRACT

This paper describes a context based technique for smoothing digital thematic maps produced by multispectral classification of Landsat Thematic Mapper data. The output of this technique is a "maplike" product which can be directly used as input to a geographic information system.

Keywords - classification, remote sensing, geographic information systems, image analysis.

INTRODUCTION

The work described here is the result of ongoing research at MacDonald, Dettwiler and Associates Ltd. into techniques for automated mapping utilizing remotely sensed data.

Multispectral classification techniques have been used on Landsat data to produce landcover maps. Traditional pixel-by-pixel multispectral classification techniques generally result in noisy or "speckly" images with a large number of small polygons which complicate the image and make the thematic image difficult to interpret. When classified imagery is used as an input to geographic information systems, the complexity of the classified image does not facilitate ease of map update or production. Thus there is a need for an effective technique to convert the classified image to a more cartographically acceptable product.

The role of a map is to effectively present information to users for their specific application at a given level of detail. In traditional mapping, a minimum mapping unit criteria is frequently used to simplify the map. In developing the final map, a cartographer takes into consideration contextual and esthetic factors.

Context plays an important role in cartography. It has been found that a large degree of the "errors" in automated computer classification is one of contextual interpretation. Even if an image could be classified with 100% accuracy, if it does not correspond to the interpretation desired by the mapper it would have "error". Interpretation frequently depends on the size, shape, and context. A small clearing inside a forest may still be interpreted as "forest" whereas a similarly sized bog inside a forest may be interpreted as "bog".

In this paper we demonstrate a technique which incorporates context in smoothing the digital thematic map to produce a more "maplike" product.

REVIEW OF EXISTING TECHNIQUES FOR SMOOTHING DIGITAL THEMATIC MAPS

Several different techniques for smoothing classified images have been documented including majority and minimum area filtering techniques [DAV76,SCH83].

In the majority filtering approach, the center pixel of an N-by-N neighborhood is replaced by the majority class of the neighborhood so small isolated polygons will be eliminated. Although this technique significantly reduces the number of polygons, small polygons still remain as there is no explicit control over the minimum polygon size. With larger window sizes there is a tendency for the majority filter to erode smaller features and affect the integrity of the polygon boundary location. Smaller window sizes however, do not produce adequate smoothing.

In the minimum area filtering approach, the class of an undersized polygon is converted to the the class of the polygon with which it shares the largest common boundary. Since there is no explicit control of the class conversion of undersize polygons, there may be undesired class conversions when an undersized polygon is converted to a very dissimilar class rather than to a more similar neighbor.

INCORPORATING CONTEXT IN SMOOTHING OF DIGITAL THEMATIC MAPS

Recognizing the importance of context in mapping we conclude that an effective technique for smoothing thematic maps should attempt to incorporate context.

As an example of the importance of context, we cite criteria used by the B.C. Forest Service in their production of forest cover maps [FOR82]:

1. Minimum Area Criteria:

   - "Mininum area may be fixed or variable, with clearly defined, important stands being mapped down to smaller areas than those which are less well defined and less important."
   - "Minimum type size of approximately 1.5 cm and 1.0 cm for forest and nonforest land respectively are recommended regardless of photographic scale."
   - Exceptions are made to the minimum area criteria in certain contexts (see 3).

2. Avoidance of Complicated Shapes:

   - "Connect small types with similar structure whenever possible."
   - "Avoid complicated, irregular type lines that hinder plotting, reading of the map, and history updating."

3. Context:

   - "Type out small nonforest patches isolated within high value types and, conversely, high value patches of timber within low value types."

We base our technique for smoothing of digital thematic maps on the concept of minimal mapping areas and use context to guide conversion of undersize regions. Regions are defined as a contiguous group of pixels of the same type. A region is undersized if it occupies an area less than that specified for its class. By allowing individual minimum sizes for each c'ass and user specifiable preferences for class conversion we can incorporate spatial context, the significance of the land cover type, and the context of the end application into the smoothing process.

Specifically the criteria that we use in our contextual smoothing technique are:

1. Regions must be larger than the specified minimum mapping size for its class or type (see Table 1).
2. If an undersized region is surrounded then it is merged with the surrounding region.
3. ...ions below the minimum size that a.e not entirely surrounded by a single class are merged with a neighbour whose class is most similar to the class of the undersize region. Table 2 shows a similarity matrix in which higher numbers represent an increasing degree of similarity.
4. If there is equal preference of a merge, merge with the region which shares the larger common boundary.

The minimum size of the regions and class similarities are specified by the user for the type of land cover being mapped, the mapping scale, and the end application of the map product.

Detailed explanation of the implementation of the technique is beyond the scope of this paper. In our work, we utilize a vector representation of the regions as the basis for our operations.

TEST RESULTS

Evaluation of the contextual smoothing technique was performed on thematic maps of two study areas in British Columbia, Canada: Adam River on Vancouver Island and Cranbrook in South Eastern British Columbia. The land cover maps produced from the classification of the Landsat scenes are for subsequent use in wildlife habitat mapping and thus the interpretation or smoothing of the land cover map was tailored for this purpose.

The Adam River and Cranbrook study areas were classified into 12 and 16 land cover classes respectively, using supervised maximum likelihood classification of Landsat 5 Thematic Mapper data. The results are shown in Figures 1 and 5.

Prior to contextual filtering, a 3-by-3 majority filter is applied to eliminate small isolated groups of pixels (see Figure 2). Minimum polygon sizes and similarity matrices suitable for the application are then specified. Tables 1 and 2 show the minimum polygon sizes and similarity matrix used in the Adam River study area. Figures 4 and 6 show the results of contextual filtering on the classified images shown in Figures 1 and 5.

Comparing the contextually filtered images to the raw classified images, we see a significant simplification of the images with a great reduction in the number of polygons. Table 3 shows a relative comparison of the number of polygons after the different smoothing operations for the Adam River study area. Comparing the results of the 5-by-5 majority filter (see Figure 3) to those of the contextual filter, we see that although the larger majority filter smooths the image significantly, it still leaves small insignificant polygons which the contextual filter has eliminated.

An important criteria in evaluating the smoothing technique is the accuracy of the resulting output. Evaluation of the accuracy of contextual smoothing (see Table 4) shows that the technique results in accuracies similar to those of large majority filters. The boundary locations of the polygons which result from this technique tend to be more accurate than those from majority filtering due to a tendency for the majority filter to erode smaller features. Furthermore, with a contextual basis, the conversion of the undersize classes will in general be logically more accurate than with techniques which do not take context into account.

Most importantly, the result is a product which is more "maplike", understandable, and compatible with geographic information systems.

Although not illustrated in this paper, an additional contextual criteria, which considers the significance of the polygon is considered, was tested and could be useful in some applications. The basic concept was that smaller polygons should be retained when it is of high significance than when it is of low significance. More simply stated, the minimum region sizes can change depending on the context; a smaller minimum size is warranted where the most similar adjacent polygon is significantly different (in our case we used a similarity score of less than 3) than in the case where the adjacent polygons are very similar. This criteria was utilized in further generalization of the image and it was found with that significant features were still retained.

CONCLUSION

In this paper we have shown an effective technique for converting classified image into a more cartographically acceptable product. The result is an output that can be vectorized and directly input to a geographic information system without further digitization or generalization.

We have shown the importance of context in mapping and demonstrated how context can be incorporated in a technique for smoothing of multispectral classification.

This technique is particularly effective for filtering of multispectral classified imagery, but can be equally effective when applied to other types of digital thematic maps.

ACKNOWLEDGEMENT

This work was performed under the financial support of an Industrial Research Assistance Program grant "Remote Sensing #2009".

REFERENCES

[DAV76] Davis, W.A. AND F. Peet., The Identification and Reclassification of Small Region On Digital Thematic Maps, Forest Management Institute Information Report FMR-X-90, 1976.

[FOR82] Forestry and Range Inventory Manual, B.C. Ministry of Forests Inventory Branch, 1982.

[SCH83] Schowengerdt, R.A., Techniques for Image Processing and Classification in Remote Sensing, Academic Press, New York, 1983.

[STR83] Strahler, A.H., Automated Forest Classification and Inventory in the Eldorado National Forest, U.S. Department of Agriculture Forest Service Report, 1983.

TABLE 1 ADAM RIVER - MINIMUM REGION SIZES

Class	Minimum Size (Pixels)	Color Code
Hemlock Clearcut	100	Dark Green
Hemlock Mature Seral	100	Olive Green
Forrested Rock	100	Dark Brown
Red Alder	50	Bright Green
Young Hemlock Clearcut	250	Bright Yellow
Huckleberry Clearcut	100	Bright Orange
Fireweed Clearcut	100	Medium Red
Recent Clearcut	100	Bright Red
Rock	50	Medium Grey
River Bar	25	Light Blue
Water	15	Dark Blue
Snow	100	White

TABLE 2 ADAM RIVER - SIMILARITY MATRIX

From \ To	Hemcc	Hemms	Forrock	Red Alder	Young Hemcc	Huckcc	Fweedcc	Reccut	Rock	River Bar	Water	Snow
Hemcc	5	4	4	3	2	1	1	1	1	0	0	0
Hemms	4	5	4	3	2	1	1	1	1	0	0	0
Forrock	4	4	5	3	2	1	1	1	1	0	0	0
Red Alder	4	4	4	5	2	3	3	2	1	0	0	0
Young Hemcc	2	3	2	1	5	4	4	3	1	0	0	0
Huckcc	2	2	2	2	4	5	4	3	1	0	0	0
Fweedcc	2	2	2	2	4	4	5	3	1	0	0	0
Reccut	2	2	2	1	4	4	4	5	2	0	0	0
Rock	1	1	4	1	3	3	3	4	5	0	0	2
River Bar	0	0	1	2	3	3	3	4	4	5	0	0
Water	0	0	0	0	1	1	1	2	3	0	5	3
Snow	0	0	1	0	0	0	0	1	4	0	0	5

TABLE 3 ADAM RIVER - NUMBER OF POLYGONS

Processing	Number of Polygons
Maximum Likelihood Classification (MLC)	39,316
3-by-3 Majority Filtered MLC	8,123
5-by-5 Majority Filtered MLC	3,798
7-by-7 Majority Filtered MLC	2,254
9-by-9 Majority Filtered MLC	1,433
Contextually Smoothed MLC	502

TABLE 4 ADAM RIVER - CLASSIFICATION ACCURACY

Processing	Accuracy (%)
Maximum Likelihood Classification (MLC)	74.9
3-by-3 Majority Filtered MLC	78.5
5-by-5 Majority Filtered MLC	80.6
7-by-7 Majority Filtered MLC	80.8
9-by-9 Majority Filtered MLC	79.1
Contextually Smoothed MLC	80.3

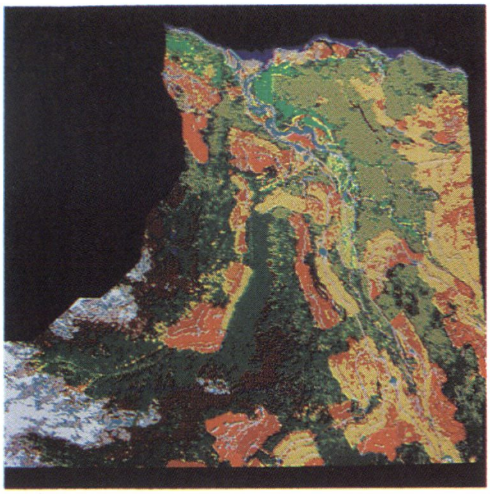

FIGURE 1 ADAM RIVER CLASSIFICATION

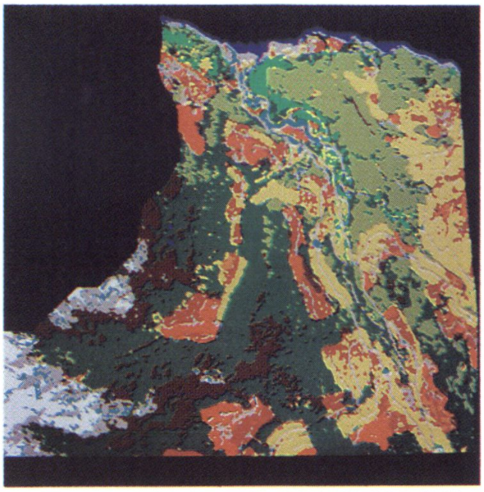

FIGURE 2 ADAM RIVER CLASSIFICATION
-3 BY 3 MAJORITY FILTERED

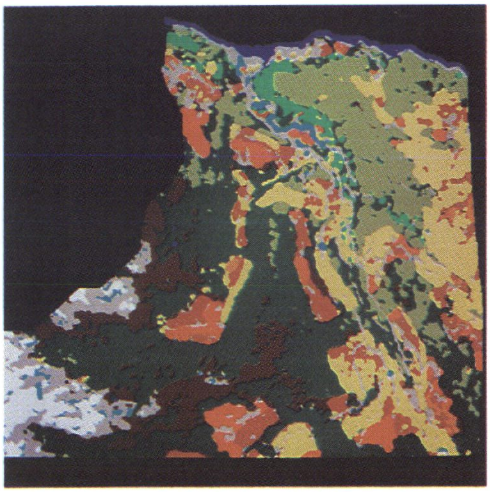

FIGURE 3 ADAM RIVER CLASSIFICATION
-5 BY 5 MAJORITY FILTERED

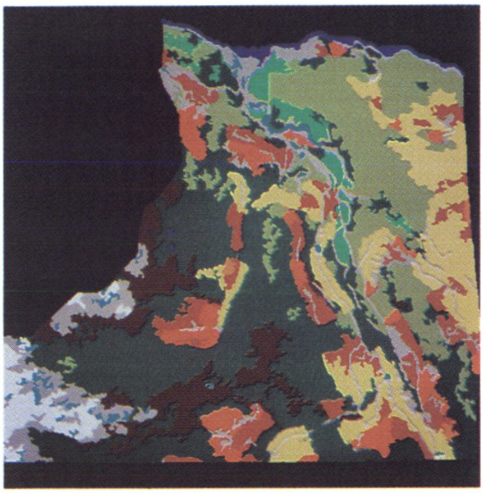

FIGURE 4 ADAM RIVER CLASSIFICATION
-CONTEXTUALLY SMOOTHED

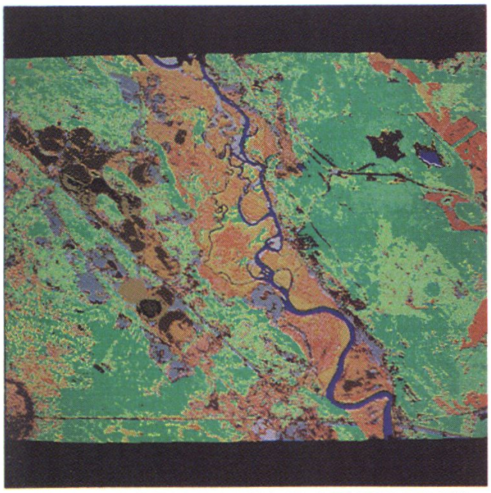

FIGURE 5 CRANBROOK CLASSIFICATION

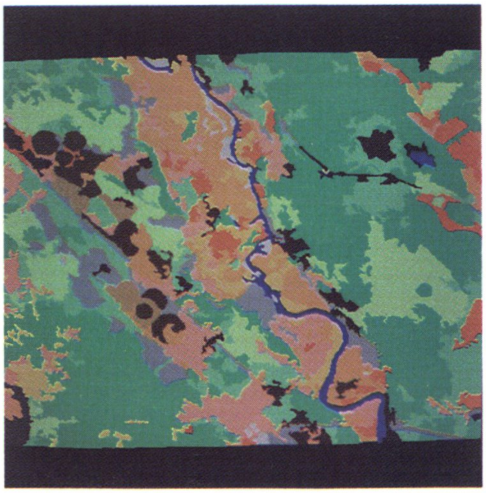

FIGURE 6 CRANBROOK CLASSIFICATION
-CONTEXTUALLY SMOOTHED