# Comparing Raster and Object Generalization

Nigel Daley [1,2], David G. Goodenough [1,2], A.S. (Pal) Bhogal [1], Quetzalcoatl Bradley [2], and Z. Yin [2]

[1]Pacific Forestry Centre, 506 West Burnside Rd, Victoria, BC, Canada, V8Z 1M5
[2]University of Victoria, Dept. of Computer Science, P.O. Box 1700, Victoria, BC, Canada, V8W 2Y2
Contact: ndaley@pfc.forestry.ca, 250-363-0153, fax: 250-363-0775

Abstract – Digital interpretation of imagery produces descriptions of the earth's surface, each description relying on the inherent resolution of the original image. Forest cover geographic information (GIS) files have been produced by interpretation of aerial photography. Common mapping scales in Canada for representing land information are 1:20,000 and 1:250,000. This paper discusses two methods to automatically generalize GIS from higher spatial resolution scales to lower scales. These two methods are a raster method (MapGen) for generalization developed by Pamap and the BC Ministry of Forests, and an object-oriented method (ObjectGen). The GIS data set consists of topographic data and forest cover files, both at 1:20,000 scale and placed on the same datum. In this presentation we compare the results for generalizing forest objects by these different methods. This work leads to segmentations of remote sensing images, at corresponding resolutions to the GIS files, being used to constrain the generalizations.

## INTRODUCTION

Remote sensing data is available from satellites and aircraft at multiple resolutions from 1 m to 1 km. For each application, there is a need to assess the utility of acquiring data at a wide variety of resolutions. In particular, can imagery at 1 m be used to derive digital interpretations appropriate for coarser resolutions? We chose to begin our investigation of how to generalize image objects by investigating the generalization of geographic information (GIS) files produced from interpretation of aerial photography. Our interest in this paper is in methods to automatically represent GIS information at other scales. We have implemented a raster method (MapGen) for generalization developed by Pamap and the BC Ministry of Forests (BCMOF) [1], and an object-oriented method described by Richardson [2]. The GIS data set consists of topographic data and forest cover files, both at 1:20,000 scale. In this presentation we compare the results for generalizing forest objects by these two different methods. The primary scales used for this test are 1:20,000 and 1:250,000. For remote sensing data, we will be segmenting images from the following sensors: MEIS (1 m), AVIRIS (20 m), TM (30 m), and AVHRR (1 km) for two test sites on Vancouver Island.

## THE DATA SETS

Forest cover and hydrology GIS data are available for British Columbia at 1:20,000 scale. We worked with 3 mapsheets, 082E062, 082E072, and 083E073, which cover the Okanagan Mountain Park area just south of Kelowna, BC. Forest cover data came from BCMOF in IGDS/Forest Inventory Planning Data Exchange Format (FIPDEF). These data files were digitized from 1994 airphotos on the NAD27 datum. The *Projected Type ID* attribute is used by BCMOF to generalize forest cover data. Values and descriptions for this attribute can be found in table 1.

Hydrology data (lakes and rivers) came from the BC Ministry of Environment, Lands, and Parks (BCMELP) Terrain Resource Information Management (TRIM) initiative. These data were digitized from airphotos using the NAD83 datum. They were converted to NAD27 to be compatible with the forest cover data.

## GENERALIZATION METHODS

Two methods of automated generalization systems were compared: raster generalization (MapGen) and object generalization (ObjectGen).

### Raster Generalization

MapGen, developed by Pamap (PCI Pacific) and BCMOF, is an automated raster generalization system. It is based on a set of polygon and vector generalization rules. Each polygon rule specifies how to combine neighbouring polygons. From

Table 1. Generalized Classes & MapGen Rules

| Class | Rule | Min Size | Merge List |
|---|---|---|---|
| 0 Water | | | |
| 1 Immature (stocking class 0) | PROJTYPEID = 1 | 15 | 3,2,4,9,5,6,8 |
| 2 Mature | PROJTYPEID = 2 | 15 | 1,3,4,9,5,6,8 |
| 3 Immature Residual | PROJTYPEID = 3 | 15 | 1,2,4,9,5,6,8 |
| 4 NSR | PROJTYPEID = 4 | 15 | 9,1,3,2,5,6,8 |
| 5 Non Commercial | PROJTYPEID = 5 | 15 | 6,4,9,1,3,2,8 |
| 6 Non Productive | PROJTYPEID = 6 | 15 | 5,4,9,1,3,2,8 |
| 8 No Typing | PROJTYPEID = 8 | 15 | 6,5,4,9,1,3,2 |
| 9 Silviculture NSR | PROJTYPEID = 9 | 15 | 4,1,3,2,5,6,8 |

0-7803-3836-7/97/$10.00 

table 1, a class 2 polygon (Project Type ID = 2) smaller than the minimum size will be merged with an adjacent class 1 polygon if it exists. If not, it will be merged with a class 3 polygon, and so on down the list. Vector rules specify the vector features to be displayed on the generalized map and a weeding tolerance for vector simplification. Feature attributes are stored in Oracle for fast sorting and selection.

After selecting input mapsheets and rules, MapGen appends and converts the IGDS files into Pamap GIS files and the FIPDEF attribute files into Oracle tables. Forest cover polygons are then converted from a vector representation to a raster representation, merged according to the polygon rules, and then re-vectorized. Vectors layers, such as rivers, are simplified using the weeding tolerance with the Douglas-Poiker Algorithm [1]. The result is a Pamap GIS file with a layer for generalized forest cover polygons and layers for each vector generalization layer.

## Object Generalization

ObjectGen [3] is a modified implementation of the automated spatial and thematic generalization method outlined in [2] and [4]. ObjectGen uses an object class hierarchy to partition features into hierarchies of objects, classes and superclasses. In our implementation the forest cover superclass was divided into the 8 classes shown in table 1 and the hydrology superclass was divided into rivers and lakes. Objects of these classes were individual forest cover polygons, river segments, or lakes.

This data structure supports generalization at the superclass, class, or object levels by ordering the objects according to common superclass, class, or object attributes and applying a removal threshold. Richardson [2] calls this threshold the necessity factor and calculates it based on a matrix of map theme, target generalization scale, map object requirement (MOR) and map object functionality (MOF). MOF defines an object's usefulness in supporting and assisting in map reading and use. For example, map objects such as rivers, roads, and boundaries assist users in reading a map because they provide the reader with a sense of orientation, among other things. MOR defines the degree of need for an object class to appear on a map at a particular scale.

These two measures, ranging from 0 to 100, are quantification's of a map object's usefulness and necessity and are specific to each map theme. They are mathematically combined together to form the necessity factor (NF) which is applied to the data as a threshold, above which features are dynamically selected.

Our implementation [3] is a two phase process: data preparation and feature selection. Phase one has three steps: lakelines are automatically generated for lakes that have one inflowing and outflowing stream and manually digitized for the rest; Strahler stream orders [5] are calculated for the river network and lakelines; object attributes are transferred to Oracle. Phase two has four steps: necessity factors for each class are calculated based on MOF and MOR tables; forest cover is generalized at its superclass level; hydrology is generalized at its object level; the resulting generalization is viewed in ArcView. Phase two can be run repeatedly with attenuation factors applied to the necessity factor calculations. In forest cover superclass generalization the class partitions are ignored and objects are ordered by area, smallest to largest. The mean necessity factor, calculated as the mean of the 8 class necessity factors, is taken as the percentage of polygons to be generalized starting at the smallest.

In hydrology object generalization, each object is ordered in its class, and its necessity factor is applied to that class' objects only. Rivers are ordered first by stream order then by length. Lakes are ordered by area. If a lake is removed it is automatically replaced by its lakeline.

# RESULTS

The 8 classes are an aggregation of original polygons which contained descriptions of forest species, volume, site index, height, and so forth. These classes correspond to the standard followed by BCMOF for representation of forest cover at 1:250,000 scale. Our initial hypothesis was that object generalization would yield more accurate results than raster generalization. The results of the generalization are shown in table 2. The table lists the classes, the original area of the classes at 1:20,000 scale in hectares, the percentages of the original area by class, and the percentages of the original area by class for MapGen and ObjectGen after generalization to 1:250,000 scale. No significant difference in areas as a function of generalization method was seen. The 1:20,000 scale maps contained 520 polygons corresponding to the 8 classes. MapGen reduced this number to 160 polygons and ObjectGen reduced this number to 150 polygons. The loss of 370 polygons did not impair the visual result of the generalization as shown in Figure 1.

Table 2. Percentage Areas by Class and Generalization

| | **Class** | **Original (ha)** | **Map Gen** | **Original** | **Object Gen** |
|---|---|---|---|---|---|
| 0 | Water | 10,263 | 21.4% | 21.3% | 21.3% |
| 1 | Immature 0 | 11,333 | 24.8% | 23.5% | 24.1% |
| 2 | Mature | 13,366 | 27.6% | 27.8% | 27.8% |
| 3 | Immature R | 454 | 0.9% | 0.9% | 0.9% |
| 4 | NSR | 1,494 | 2.9% | 3.1% | 2.9% |
| 5 | Non Comm | 132 | 0.0% | 0.3% | 0.2% |
| 6 | Non Prod | 10,910 | 22.0% | 22.7% | 22.3% |
| 9 | Silvi NSR | 182 | 0.4% | 0.4% | 0.4% |
| **Subtotal** | | **48,132** | **100%** | **100%** | **100%** |
| **Number of polygons** | | | **160** | **520** | **150** |

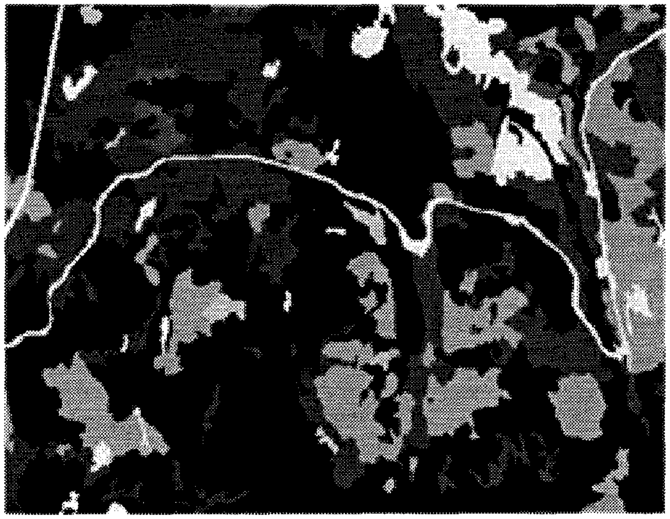

Figure 1a. Portion of 1:20,000 scale original forest cover data covering an area 8 km by 6 km.

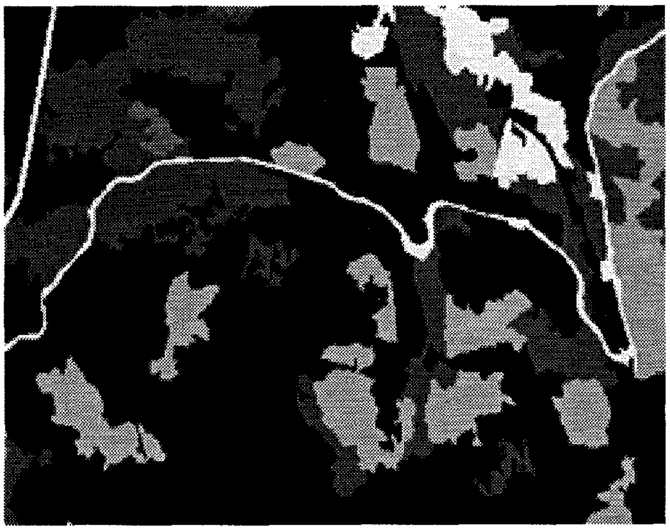

Figure 1b. 1:250,000 scale MapGen generalization of Fig 1a.

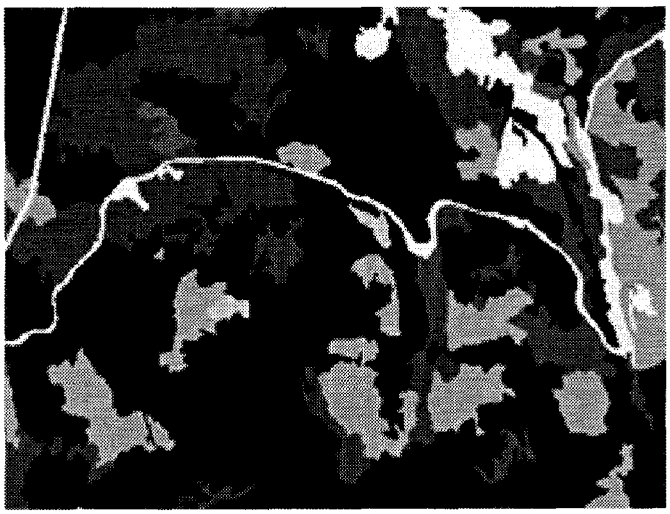

Figure 1c. 1:250,000 scale ObjectGen generalization of Fig 1a.

## CONCLUSIONS

Two methods of generalization were implemented, a raster generalization (MapGen) and an object generalization (ObjectGen). These methods were applied to three forest cover maps to create broad classes. There were no significant differences in class areas between the two generalizations and the original areas. Reductions in the original number of polygons of over 72% were achieved without significant errors in class areas. We are now investigating segmentation of remote sensing imagery, conversion of segments to labeled objects, and generalization of these objects.

## ACKNOWLEDGMENTS

This research was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada under grant OGP0184109, and by the Pacific Forestry Centre, Natural Resources Canada. We thank the Resources Inventory Branch of BCMOF and the Geographic Data BC Branch of BCMELP for their ongoing support.

## REFERENCES

[1] Pamap Technologies Corporation, "Map Generalization User Guide," PCI Pacific, Victoria, BC, Canada May 1994.

[2] D. E. Richardson, "Automated Spatial and Thematic Generalization Using a Context Transformation Model," Ph.D. thesis, University of Wageningen, Wageningen, The Netherlands, 1993.

[3] N. M. A. Daley and Q. Bradley, "Automated Spatial and Thematic Generalization: an implementation," University of Victoria, Victoria. BC, Canada http://www.aft.pfc.forestry.ca/ndaley/objgen/index.html, April 18 1997.

[4] D. E. Richardson, "Rule Based Generalization for Base Map Production," M.Sc. thesis, International Institute for Aerospace Survey and Earth Sciences, Enschede, The Netherlands, 1988.

[5] A. N. Strahler, "Quantitative Geomorphology of Drainage Basins and Channel Networks," in *Handbook of Applied Hydrology*, V. T. Chow, Ed., 1964.