GeoInformatica 8:2, 143–155, 2004

Predicting Forest Age Classes from High Spatial Resolution Remotely Sensed Imagery Using Voronoi Polygon Aggregation

TRISALYN NELSON,[1] BARRY BOOTS,[1] MIKE WULDER[2] AND ROB FEICK[3]
[1] *Wilfrid Laurier University, Department of Geography and Environmental Studies, Waterloo, Ontario, N2L 3C5, Canada*
E-mail: nels2122@machl.wlu.ca; bboots@wlu.ca
[2] *Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, 506 West Burnside Road, Victoria, BC, V8Z 1M5, Canada*
E-mail: mwulder@pfc.forestry.ca
[3] *University of Waterloo, Department of Geography, Waterloo, Ontario, N2L 3G1, Canada*
E-mail: rdfeick@fes.uwaterloo.ca

Received November 21, 2002; Revised June 30, 2003; Accepted July 11, 2003

Abstract

Efficient identification of forest age is useful for forest management and ecological applications. Here we propose a user-assisted method for determining forest age using high spatial resolution remotely sensed imagery. This method requires individual trees to be extracted from imagery and represented as points. We use a local maximum filter to generate points that are converted to Voronoi polygons. Properties of the Voronoi polygons are correlated with forest age and used to aggregate points (trees) into areas (stands) based on three forest age classes. Accuracy of the aggregation ranges from approximately 68% to 78% and identification of the mature class is more consistent and accurate than the younger classes.

Keywords: Voronoi polygon, forest age, spatial analysis, point patterns

1. Introduction

Automated techniques for extracting features from high spatial resolution remotely sensed imagery are increasingly common and allow detailed information to be collected over large areas. While many ecological and management applications benefit from detailed data, the size of these data sets can lead to difficulties with data storage and may have limitations when attempting to understand general landscape trends. Therefore, methods to efficiently aggregate features provide a mechanism for generalizing detailed data and add flexibility to data sets generated by automated feature extraction.

One approach for extracting individual tree locations from high spatial resolution imagery is local maximum (LM) filtering [13], which represents trees as points. Commonly, however, forest information is presented as polygons. A technique to convert individual trees (points) into homogenous forest stands (polygons), allows the detailed data to be stored in a more traditional and, in some cases, a more efficient format.

In this paper, we focus on a method for generating polygons of homogenous forest age. Forest age mapping is useful in applications such as ecology and forest management. For instance, forest age mapping is used to predict susceptibility of forest stands to pests, such as the mountain pine beetle, which generally attack older trees [6], [8].

The spatial distribution (spatial pattern) of tree stems is one aspect of forest structure that has been shown to have an association with forest age [2], [5], [11]. The relationship between the spatial distributions of tree stems and forest age is best documented in literature of forest canopy gaps. For example, it has been shown that in coastal Douglas-fir forests, mature stands have a significantly higher number of gaps than younger stands [7]. Forest canopies of mature coastal Douglas-fir are also more open and heterogeneous than in younger stands [2]. The correlation between canopy openness and forest maturity can be related to changes in the spatial pattern of stems that occur as forests age. As forests mature, individual trees become larger requiring more space and therefore are found at lower densities than when stands are young. Eventually old trees die, due to small disturbances or inter-species competition, and new canopy gaps provide sufficient light and resources for younger trees to emerge [5], [12]. The emergence of younger trees increases the stem density of the forest. Therefore, under natural conditions, as a forest matures, individual tree density will first decrease and then increase slightly.

Relationships between the spatial patterns of stems and forests age are useful for aggregating individual tree locations, extracted from remotely sensed imagery, into polygons of homogenous forest age. In a previous study [9], forest age polygons were generated from points representing tree locations using a method based on nearest neighbor distances. Although the nearest neighbor method proved useful, here we improve the approach by using a set of techniques involving Voronoi polygons (VPs). In particular, the VP approach permits the automated conversion of points to polygons and also simplifies the delineation of non-forested areas.

In this study individual trees were extracted from a 1 m spatial resolution IKONOS satellite image using a LM filter and represented as points; points were then converted to VPs. Next, forest inventory data were used to identify training areas for nine inventory age classes. Summaries of VP area and perimeter attributes were calculated for age stratified training areas to determine the correlation between VP attributes and forest age. Age summaries of VP attributes were used to determine possible forest age classes and VP attribute thresholds for aggregation. Subsequently, individual trees were aggregated into three classes: Young (1–20 years), Intermediate (21–120 years), and Mature (> 120 years). Finally, the accuracy of this aggregation was determined with reference to forest inventory data. While we demonstrate this method with trees in a forestry context, it can be applied to spatial distributions of other point objects.

2. Study area and data

The study area is located in the Sooke Watershed, 40 km northwest of Victoria, British Columbia, Canada (48° 34′ N, 123° 42′ W). The Sooke Watershed is dominated by coastal Douglas-fir (*Pseudotsuga menziesii*). Stand age ranges from 0 to > 250 years [4], although there are no stands aged 120–141 years and few stands exist between 60 and 100 years due to harvesting operations. The forest stands younger than 120 years are managed plantations, while the older stands are naturally occurring.

The forest information required to develop the aggregation technique and perform accuracy assessments was obtained from the Sooke Watershed forest inventory map (scale 1:15000) [4]. This map was generated via airphoto interpretation in 1991, at which time stand age attributes representing the age expected in 1998 were predicted. As with all of British Columbia Ministry of Forests' inventory data, age is represented with nine classes (table 1) [3].

The locations of individual trees were extracted from a panchromatic IKONOS image with 1 m spatial resolution. This imagery was collected on June 3, 2000 at 11:05 PST and covers an area of approximately 8 km^2. The upper left corner of the image is 441962E, 5386346N and the lower right corner is 451162E, 5377371N.

Tree locations were isolated from the IKONOS imagery using the well-documented LM filter [13]. Our goal is not to revisit the LM technique, but rather to develop an approach for generalizing a variety of objects identified with different feature extraction approaches. Therefore, our description of the LM filter is brief and we refer readers to the cited reference for further details. The LM filter locates individual trees based on the assumption that tree crowns are convex. If this assumption is correct, the approximate tree apex should occur where image radiance values are the highest and radiance values should decrease towards the edge of a tree crown [13]. Using a moving window, the LM filter locates a tree only if the radiance value of the central pixel is greater than all other pixels in the window. For this study a 3 by 3 window was used as it best represents the size of mature Douglas-fir tree crowns. To reduce image noise, which causes trees to be falsely identified, the IKONOS image was smoothed using a 3 by 3 averaging filter prior to feature extraction. A drawback to smoothing an image is the reduced likelihood of finding trees with small crowns.

Details of feature extraction results and accuracy can be found in Nelson et al. [9] where it was demonstrated that 70% of mature trees were identified correctly, while only 37% of young trees were accurately located. Less than 1% of trees located by the LM filter were false.

Table 1. British Columbia ministry of forests' inventory age classes.

Forest Inventory Age Class	*Age (years)*
1	1–20
2	21–40
3	41–60
4	61–80
5	81–100
6	101–120
7	121–140
8	141–250
9	> 250

3. Methods

An overview of VP methods for aggregating trees into forest polygons based on tree age can be seen in table 2. First, LM points were converted to VPs. Given a set of points in a continuous space, VPs partition the space so that all locations in the space are associated with the closest point [10]. VPs are advantageous for spatial analysis of point patterns since their construction automatically defines spatial relationships between the points even if the pattern is not homogeneous (e.g., neighboring points can always be defined as those points whose VPs share a common edge regardless of variations in the density of the points across space). As well, attributes of the VPs, such area and perimeter, can be assigned to the points. Area and perimeter attributes are particularly useful in this context as they are inversely related to the density of objects.

Successful aggregation of trees into forest age polygons requires analysis to determine which, if any, VP attributes are sensitive to changes in forest age. In order to determine

Table 2. Overview of methods for aggregating individual trees into forest age polygons using VP techniques.

Task	*Task Description*	*Task Goal*
LM filtering	A 3×3 LM filter is run on an IKONOS image smoothed with a 3×3 averaging filter.	Find all individual trees and represent locations with points.
Generating VP	VPs are generated for all trees identified via LM filtering.	Use VP attributes to aggregate individual trees into forest age polygons.
Select training areas	A forest inventory map is used to identify (50×50 m) training areas for each of the nine forest inventory age classes. These areas are selected throughout the entire IKONOS image.	Training areas stratified by forest age are required to test the relationship between VP attributes to forest age.
Assessing the relationship between VP attributes and forest age	VP attribute values, based on area and perimeter, are summarized for the age stratified training areas.	Identify which VP attributes are most capable of producing forest age polygons and determine age classes and attribute thresholds for aggregation.
Select sub-areas for aggregation	Three sub-areas (average area $= 633$ sq m) are selected from the IKONOS image and used for aggregation.	Due to the IKONOS image size, attributes are generated for three selected sub-areas.
Generating attributes	VP attributes shown to be the most sensitive to forest age are generated for all trees in the sub-areas.	VP attributes sensitive to forest age are required for aggregation.
Aggregation	Using age classes and thresholds suggested by analysis of the relationship between VP attributes and forest age, trees are aggregated into age polygons.	Generate forest age polygons.
Aggregation accuracy assessment	By comparing the aggregated sub-areas with forest inventory polygons two accuracy assessments are performed. The first assessment is at the VP scale and the second at the forest inventory polygon scale.	Determine the accuracy of the aggregation technique.

how well VP attributes reflect forest age 50 by 50 m training areas were selected for each of the nine inventory age classes based on the forest inventory map. For each age class, three to six training areas were used; more training areas were selected when age classes were well represented throughout the study area. We tested how well VP attributes correlated to forest age using both smoothed and unsmoothed attributes. To smooth VP attributes, all values within a neighborhood defined for VP i were averaged and the new value assigned to VP i. Two neighborhoods were defined for smoothing. Lag 1 neighborhoods include VP i and all adjacent VPs, whereas lag 2 neighborhoods also include all VPs adjacent to the lag 1 neighborhood. When attributes were smoothed, additional polygons outside the study area were used in neighborhood definition when necessary.

Smoothing the attributes compensated for variability in the training data, which was expected for two reasons. First, even within stands of a single age, variability in the distances between trees will occur due to factors such as soil moisture, soil nutrients, and competition. Second, training locations were selected from forest inventory polygons, which when created using traditional techniques such as airphoto interpretation, are considered homogeneous based on a set of characteristics such as age, species, height, and crown closure [3]. Consequently, single attributes, such as forest age, may vary slightly throughout the polygon.

The smoothed and unsmoothed area and perimeter attribute values were summarized for each of the forest inventory age classes (see tables 4 and 5 in Sections 4 and 5). By interpreting the summarized values, we determined which attributes were the most sensitive to forest age. Using natural breaks in the age stratified summaries of lag 2 VP attributes, we identified the number of age classes, range of each age class, and attribute thresholds to use for aggregations (see table 3). In this way, three age classes were created: Young (1–20 years), Intermediate (21–120 years), and Mature (> 120). Non-forest areas were excluded from the aggregation by disregarding all VPs that intersected roads, water, and other non-forested areas by at least 50%.

To assess the accuracy of the aggregation procedure three sub areas, having an average area of 633 m^2, were selected as representative of forest ages within the Sooke Watershed. In each sub area a digital version of the forest inventory polygons was overlain with the aggregation results and the accuracy assessment was conducted at two spatial scales. The first assessment was based on VPs where the age class assignment of each VP was compared with the forest inventory data. The second assessment was based on forest inventory polygons. For each forest inventory polygon we determined the assigned age class of all intersecting VPs. The modal age class of the intersecting VPs was then

Table 3. VP attribute thresholds used for aggregation. Units are in meters.

Age Class	*Lag 2 Area*	*Lag 2 Perimeter*
Young	< 50	< 28
Intermediate	50–72	28–34
Mature	> 72	> 34

compared to that of the forest inventory polygon and the percentage of instances where the age classes were the same was calculated.

4. Results and discussion

The sensitivity of VP attributes to forest age is summarized in tables 4 and 5. Both VP area and perimeter are correlated with forest age, with attributes smoothed with a lag 2 neighborhood being the most sensitive to changes in forest age. Although area and perimeter seem approximately equal in their sensitivity to forest age, perimeter attributes have less within class variation.

Generally, VP area and perimeter increase with forest age. The main exception to this trend is forest age class 9, which tends to have lower attribute values than age class 8. This trend reflects expected changes in the spatial pattern of trees as forests age. A natural stand is likely to self-thin once it reaches maturity (age class 8); eventually smaller trees will regenerate in the gaps causing the average spacing between trees to decrease (age class 9). While a similar trend is seen for class 6, it is likely an artefact of few training areas in this age class, which was not well represented in the study area.

Stratifying VP attributes by age allowed us to identify natural breaks in the relationship between VP attribute value and forest age, particularly when attributes are smoothed. For example, lag 2 area VP attributes show natural breaks between forest age classes 1 and 2, 3 and 4, and between 6 and 8. These natural breaks, as mentioned in the methods, became the aggregation thresholds; the threshold between forest age classes 3 and 4 was not applied due to a lack of forest representing age classes 4 to 5 in the study area. Similar age class thresholds were applied when aggregation was based on VP perimeter.

Aggregation results for the three sub areas are shown in figures 1–3. Aggregations were based on both lag 2 area and perimeter attributes, however only perimeter is shown as

Table 4. Comparison of mean values generated for each age class using the 50×50 m sub-areas. Area – Single Tree = average values calculated based on the area of the VP of each tree. Area – Lag 1 = the average area of the VP of each tree and the VPs of its lag 1 neighbors. Area – Lag 2 = the average area of the VP of each tree and the VPs of its lag 1 and 2 neighbors.

	Area – Single Tree		*Area – Lag 1*		*Area – Lag 2*	
FI-Age	*Mean*	*CV*	*Mean*	*CV*	*Mean*	CV
1	41.6	38.3	42.8	23.1	43.6	16.6
2	46.5	41.1	49.4	23.4	51.1	15.9
3	51.2	42.7	53.4	27.6	54.7	21.3
4	50.9	35.3	53.8	21.0	66.0	13.2
5	61.8	36.2	64.7	22.4	65.6	18.8
6	58.5	39.0	62.3	21.2	65.3	16.1
8	83.6	44.3	89.5	23.6	91.6	15.5
9	83.9	41.4	84.9	23.6	90.9	17.3

Notes. FI = forest inventory; CV = coefficient of variation.

Table 5. Comparison of mean values generated for each age class using the 50×50 m sub-areas. Perim – Single Tree = average values calculated based on the perimeter of each tree. Perim – Lag 1 = the average perimeter of each tree and its lag 1 neighbors. Perim – Lag 2 = the average perimeter of each tree and its lag 1 and 2 neighbors.

	Perim – Single Tree		*Perim – Lag 1*		*Perim – Lag 2*	
FI-Age	*Mean*	*CV*	*Mean*	*CV*	*Mean*	*CV*
1	25.8	17.8	26.1	11.7	26.3	10.1
2	26.9	19.2	27.6	11.9	28.2	7.7
3	28.1	20.5	28.7	14.3	29.0	11.1
4	28.5	16.1	29.2	10.2	29.8	7.6
5	31.8	16.5	32.3	11.2	32.6	9.5
6	30.8	17.8	31.6	10.7	32.3	8.1
8	36.2	18.8	37.3	11.1	37.8	7.8
9	35.5	22.0	36.2	13.2	36.6	10.6

Table 6. Percentage of VPs accurately classified. Each VP is assessed to determine if it was assigned to the correct age class.

	Young		*Intermediate*		*Mature*	
Age	*Area*	*Perimeter*	*Area*	*Perimeter*	*Area*	*Perimeter*
Young	42.5	42.3	48.8	50.1	8.7	7.6
Intermediate	20.6	20.5	57.2	58.4	32.1	31.1
Mature	0.4	0.3	11.7	12.5	88.0	87.2

Table 7. Percentage of forest inventory polygons correctly classified. Aggregation based on the lag 2 perimeter attribute.

Age	*Young*	*Intermediate*	*Mature*
Young	50.0	50.0	0.0
Intermediate	6.3	68.8	25.0
Mature	0.0	8.3	91.7

there are no visible differences. The VP level accuracy assessments are shown in table 6. Using the lag 2 area attribute, 67.9% of VPs were accurately classified. The corresponding value for the lag 2 perimeter was 68.3%. The Mature class was the most accurately classified and the Young class was the least accurately classified. Confusion in the classification occurred between Young and Intermediate classes, and Intermediate and Mature classes. This is not surprising as the VP attributes suggest the Intermediate class acts as a transition between the Young and Mature classes.

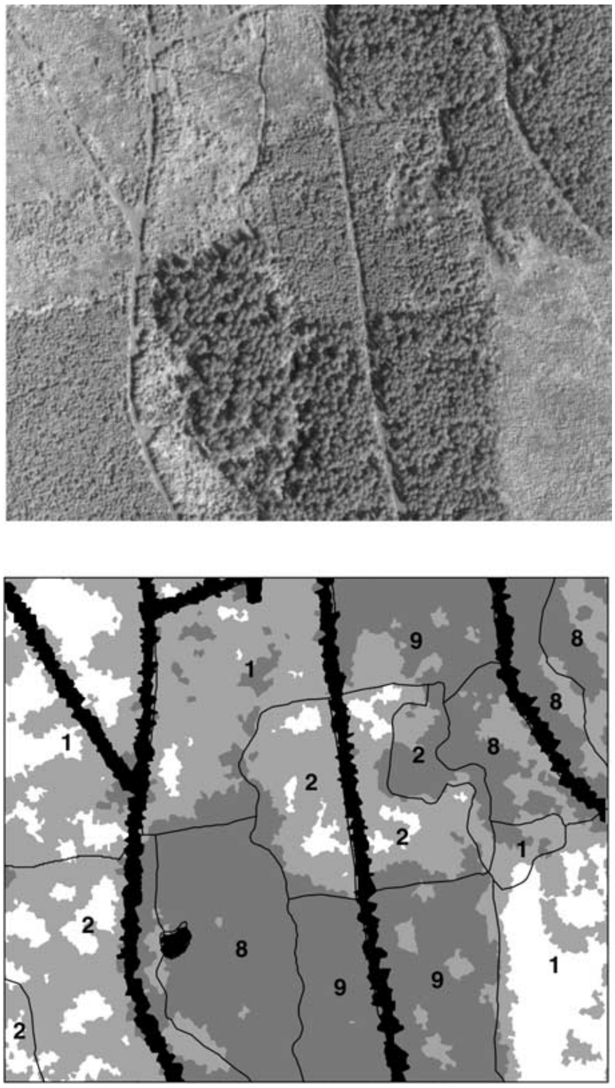

Figure 1. Site 1 IKONOS imagery is presented in the upper pane for comparison to the aggregation results, based on the average perimeter of polygons in lag 2 neighborhoods, in the lower pane. (White = young forest. Light gray = intermediate aged forest. Dark gray = mature class. Black = roads and non-forested areas. Numbers represent British Columbia Ministry of Forests' Inventory age classes.)

The forest inventory polygon based accuracy assessment showed the aggregation to be 78.3% correct (Table 7). The results of aggregations were so similar that the forest inventory based accuracy assessment was conducted for only the lag 2 perimeter method. Although the classification accuracy is higher when based on forest inventory polygons, trends are the same as for the individual VP accuracy assessment.

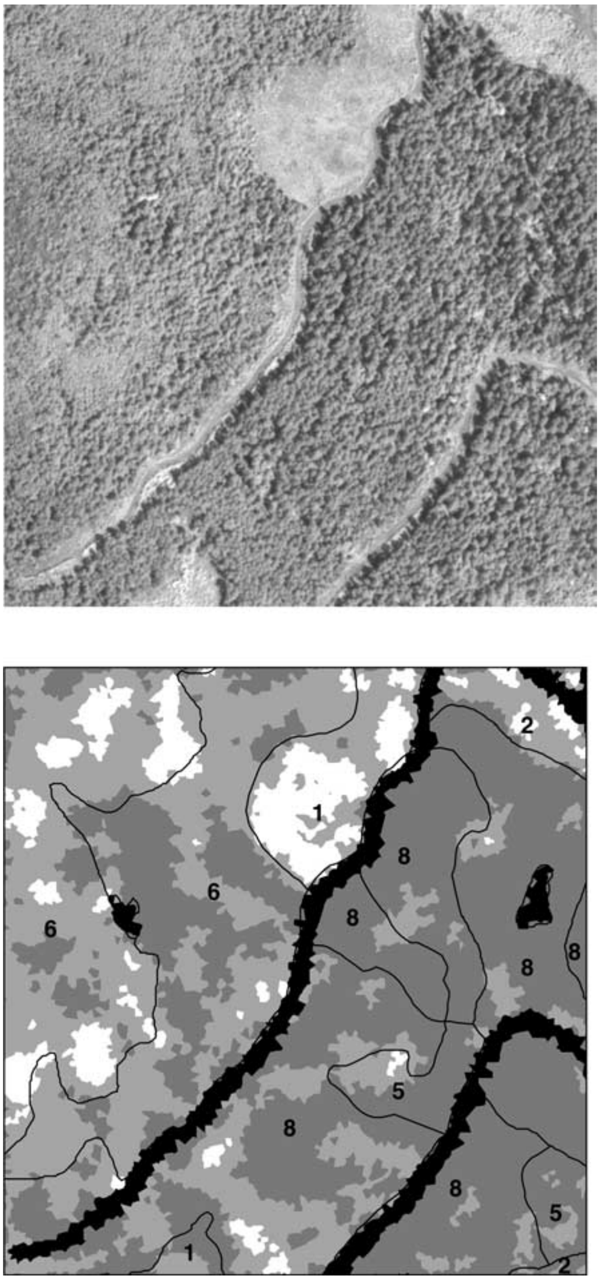

Figure 2. Site 2 IKONOS imagery is presented in the upper pane for comparison to the aggregation results, based on the average perimeter of polygons in lag 2 neighborhoods, in the lower pane. (White = young forest. Light gray = intermediate aged forest. Dark gray = mature class. Black = roads and non-forested areas. Numbers represent British Columbia Ministry of Forests' Inventory age classes.)

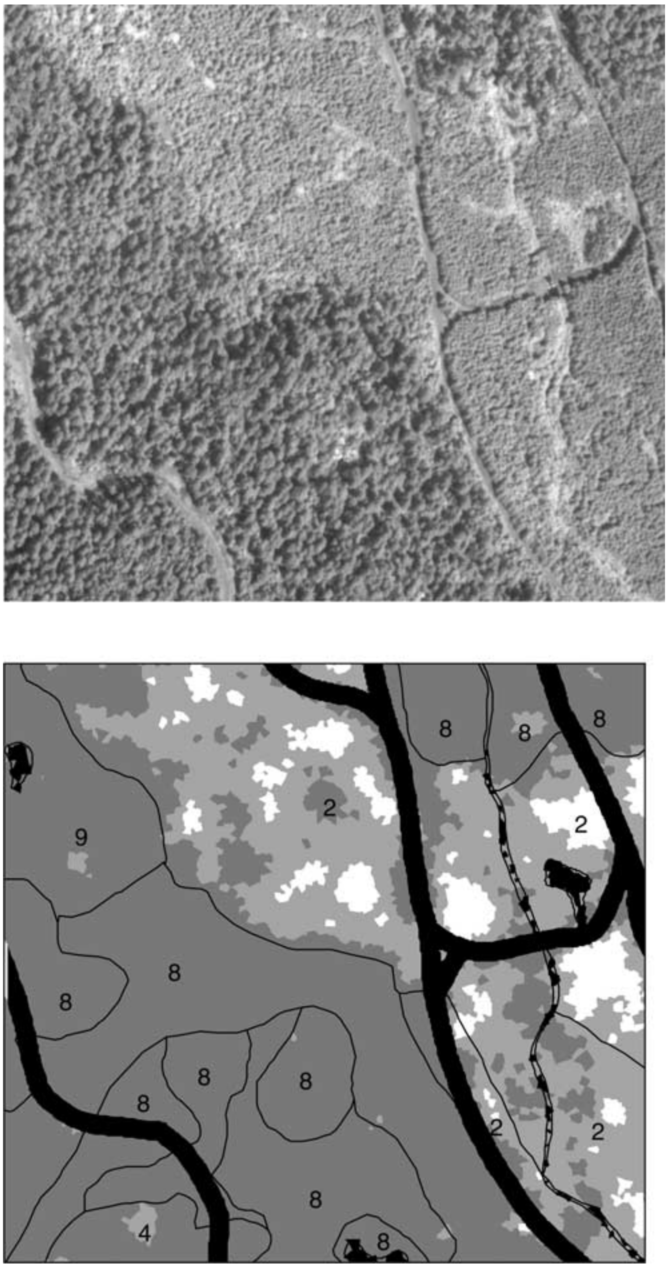

Figure 3. Site 3 IKONOS imagery is presented in the upper pane for comparison to the aggregation results, based on the average perimeter of polygons in lag 2 neighborhoods, in the lower pane. (White = young forest. Light gray = intermediate aged forest. Dark gray = mature class. Black = roads and non-forested areas. Numbers represent British Columbia Ministry of Forests' Inventory age classes.)

The accuracy assessment based on forest inventory polygons is subject to the same difficulties as discussed for the training data. It is possible that the accuracies are actually higher than reported, as the aggregation results, which are based only on age, are being compared with the forest inventory polygons that are relatively homogenous for a number of characteristics.

5. Conclusion

This paper demonstrates a method for extracting individual trees from remotely sensed imagery and aggregating tree points into forest age classes based on the properties of VPs. Although the accuracy of this method is similar to aggregations based on nearest neighbor distances, there are several advantages to using VPs. In comparison with the nearest neighbor approach [9], the use of VPs reduces the subjectivity of the aggregation method. For example, the spatial characteristics of VPs discussed above allow the spatial distribution of the trees to determine both the scale of analysis and neighbor relationships between trees, whereas aggregation using a nearest neighbor approach requires the scale of analysis and therefore the neighborhoods to be defined subjectively. Also, removal of non-forested areas is simplified using VPs. Finally, aggregating VP into classes automatically produces forest polygons while nearest neighbor methods require a post-processing technique to generate forest polygons from grouped points.

Our method is most accurate when trees are mature. This is most likely related to the accuracy of the LM filter, which in this study generally improves as tree size, and therefore age, increases. The relationship between LM filter success and tree size will change when using imagery with different spatial resolutions. When extracting individual trees in a coastal Douglas-fir forest from 1 m spatial resolution imagery, the high accuracy of mapping mature forests suggests that one of the best uses for these methods may be applications that require identification of broad mature age classes. For example, mountain pine beetles generally attack forests that are 80 years or older, therefore this method could be quickly implemented to identify the location of trees in age classes susceptible to mountain pine beetle attacks.

Our results also contribute to the general debate regarding the representation of forests using object and field models. Traditionally, forests are mapped as objects or discrete units, however boundaries in forests are not always crisp and forest polygons not always homogenous [1]. Using the VP method for generating forest polygons allows the variability within the crisp objects to be visualized, and may provide a mechanism for reporting uncertainty and fuzziness within forest boundaries. Further investigation is required to determine which variations are artefacts of the aggregation procedure and which represent true variation not reflected in the forest inventory data.

In this paper, we have demonstrated how aspects of the spatial distribution of trees can be used to generate age based forest polygons. Given that points can be used to represent many environmental and social objects, these methods may be suitable for other applications, where spatial distribution of objects is of interest.

Acknowledgments

We are grateful for the insightful comments of four anonymous reviewers which resulted in improvements to the final paper.

References

1. D. Brown. ''Classification and boundary vagueness in mapping presettlement forest types,'' *International Journal of Geographical Information Science*, Vol. 12:105–129, 1998.
2. G. Frazer, J. Troyfymow, and K. Lertzman. ''Canopy openness and leaf area in chronosequences of coastal temperate rainforests,'' *Canadian Journal of Forest Research*, Vol. 30:239–256, 2000.
3. M. Gillis and D. Leckie. ''Forest inventory mapping procedures across Canada,'' PNFI, Ontario, Canada. Information Report. PI-X-114. 79, 1993.
4. GVWD. ''Greater Victoria water district watershed management forest cover classification,'' Sooke Lake Watershed, Hugh Hamilton Ltd., Victoria. Scale 1:15000, 1991.
5. F. He and R. Duncan. ''Density-dependent effects of tree survival in an old-growth Douglas fir forests,'' *Journal of Ecology*, Vol. 88:676–688, 2000.
6. S. Larsson, R. Oren, R. Waring, and J. Barrett. ''Attacks of mountain pine beetle as related to tree vigor of ponderosa pine,'' *Forest Science*, Vol. 29:395–402, 1983.
7. K. Lertzman, G. Sutherland, A. Inselberg, and S. Saunders. ''Canopy gaps and the landscape mosaic in a costal temperate rainforest,'' *Ecology*, Vol. 77:1254–1270, 1996.
8. R. Mitchell and H. Preisler. ''Analysis of spatial patterns of lodegepole pine attacked by outbreak populations of the mountain pine beetle,'' *Forest Science*, Vol. 37:1390–1408, 1991.
9. T. Nelson, O. Niemann, and M. Wulder. ''Spatial statistical techniques for aggregating point objects extracted from high spatial resolution imagery,'' *Journal of Geographical Systems*, Vol. 4:423–433, 2002.
10. A. Okabe, B. Boots, K. Sugihara, and S. Chiu. Spatial Tessellations: Concepts and Applications of Voronoi Diagrams, Second Edition. John Wiley: Chichester, 2000.
11. C. Oliver and B. Larson. Forest Stand Dynamics. John Wiley & Sons: New York, 1996.
12. T. Spies and J. Franklin. ''Gap characteristics and vegetation response in coniferous forests of the Pacific Northwest,'' *Ecology*, Vol: 70:543–545, 1989.
13. M. Wulder, O. Niemann, and D. Goodenough. ''Local maximum filtering for the extraction of tree locations and basal area from high spatial resolution imagery,'' *Remote Sensing of Environment*, Vol. 73:103–114, 2000.

Trisalyn Nelson is a Ph.D. candidate in Geography and Environmental Studies at Wilfrid Laurier University. Her dissertation focuses on the spatial-temporal analysis of mountain pine beetle infestations. Her main interests are spatial pattern analysis, and detecting and representing change in space and time.

Barry Boots is a Professor of Geography and Environmental Studies at Wilfrid Laurier University. He received a Ph.D. degree form Rutgers University in 1972 and a D.Sc. from the University of Bristol in 1995. His main interests are measuring spatial structure and spatial relationships, spatial statistics, and Voronoi diagrams.

Michael A. Wulder is a Research Scientist with the Canadian Forest Service of Natural Resources Canada, located at the Pacific Forestry Centre in Victoria, British Columbia. He received his Ph.D. from the University of Waterloo in 1998. Dr. Wulder leads a research and applications program on ''Forest Geomatics'', that is focused on the use of spatial data to estimate forest inventory and structural attributes. Application research interests include, land-cover mapping of Canada with Landsat ETM+ data, high spatial resolution inventory attribute estimation, change detection for forest disturbance, spatial statistical description of forest structure, and lidar remote sensing of forest inventory attributes.

Rob Feick is an Assistant Professor in the Department of Geography and the School Planning at the University of Waterloo. His research interests include GIS, spatial decision support systems, multi-criteria analysis methods for group-based decision making, and spatial statistics.