The Aggregation of Pixel Data into Mapped Area Features

A Case Study based on the Wrangell-St Elias Mapping Project

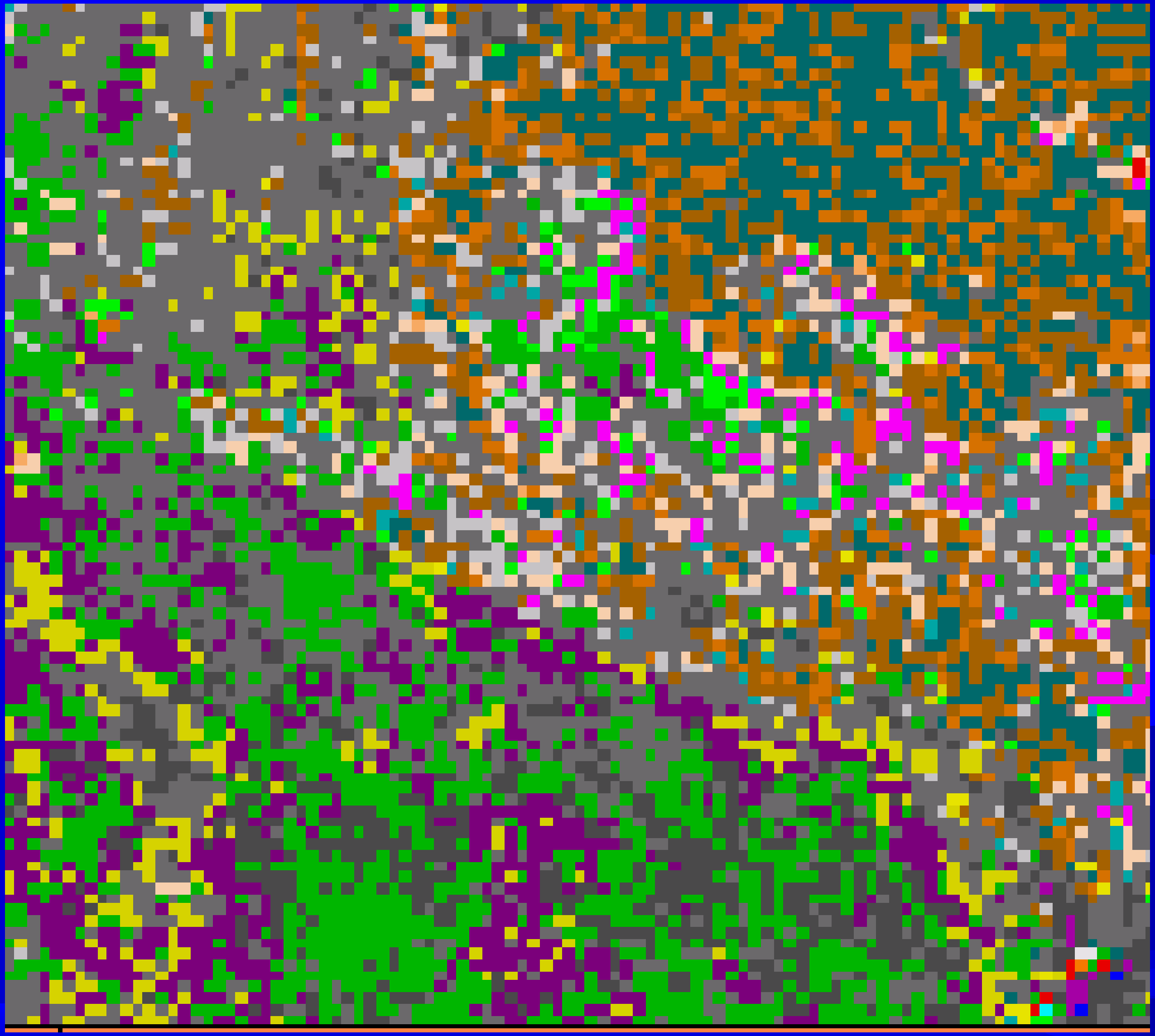
Ken Stumpf

Geographic Resource Solutions

Project Classification Results

Vectorize This

Or How About This ???



Unfortunately, most land cover is not distributed homogeneously in large areas !

Conversion to vectors is difficult.

Why don't we just leave the data set in pixel format ??

? Image Processing Myth ?

- Pixels are just fine.
- Polygons are unnecessary - we can live without them!

Reality
Pixel Heterogeneity Confounds the User

- Excessive information and detail - resolution is often too small for most projects
- Represents data at a level that is not manageable
  - Difficult to process and query
  - Difficult to summarize and evaluate
  - Difficult to evaluate for accuracy

Why We Like Polygons ?

- Enable area queries
- Easier to understand and map
- Can describe spatial relationships of types - corridors, buffers, adjacencies, and edges
- Results in ‘type’ level information that may not be present at the pixel level
- Easier to test accuracy - we can develop and use statistics at the polygon level

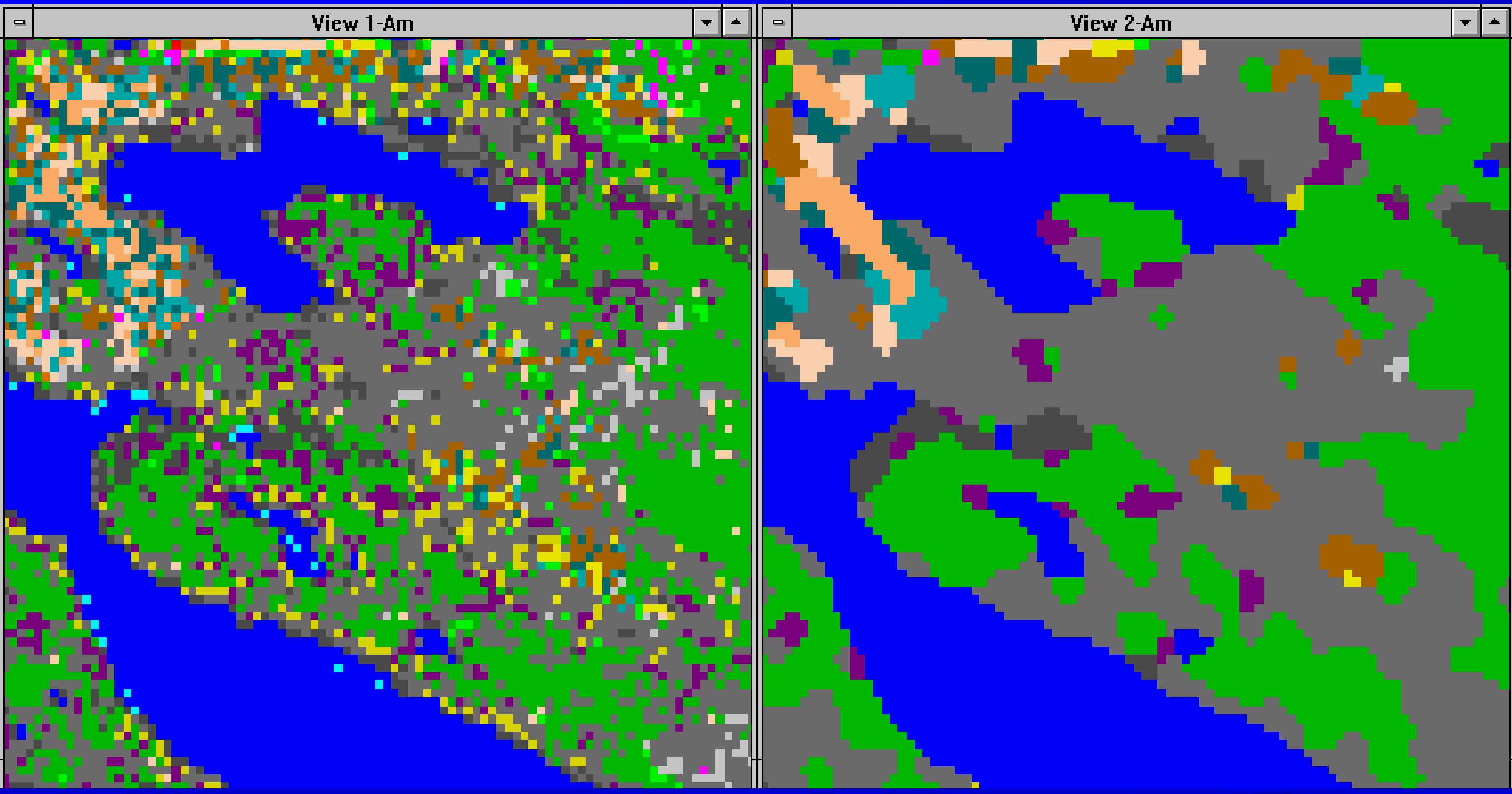
Traditional Approach to Pixel Cleaning or Polygon Formation

Filtering, Scanning, Smoothing, and Merging
or
“How to distort data”

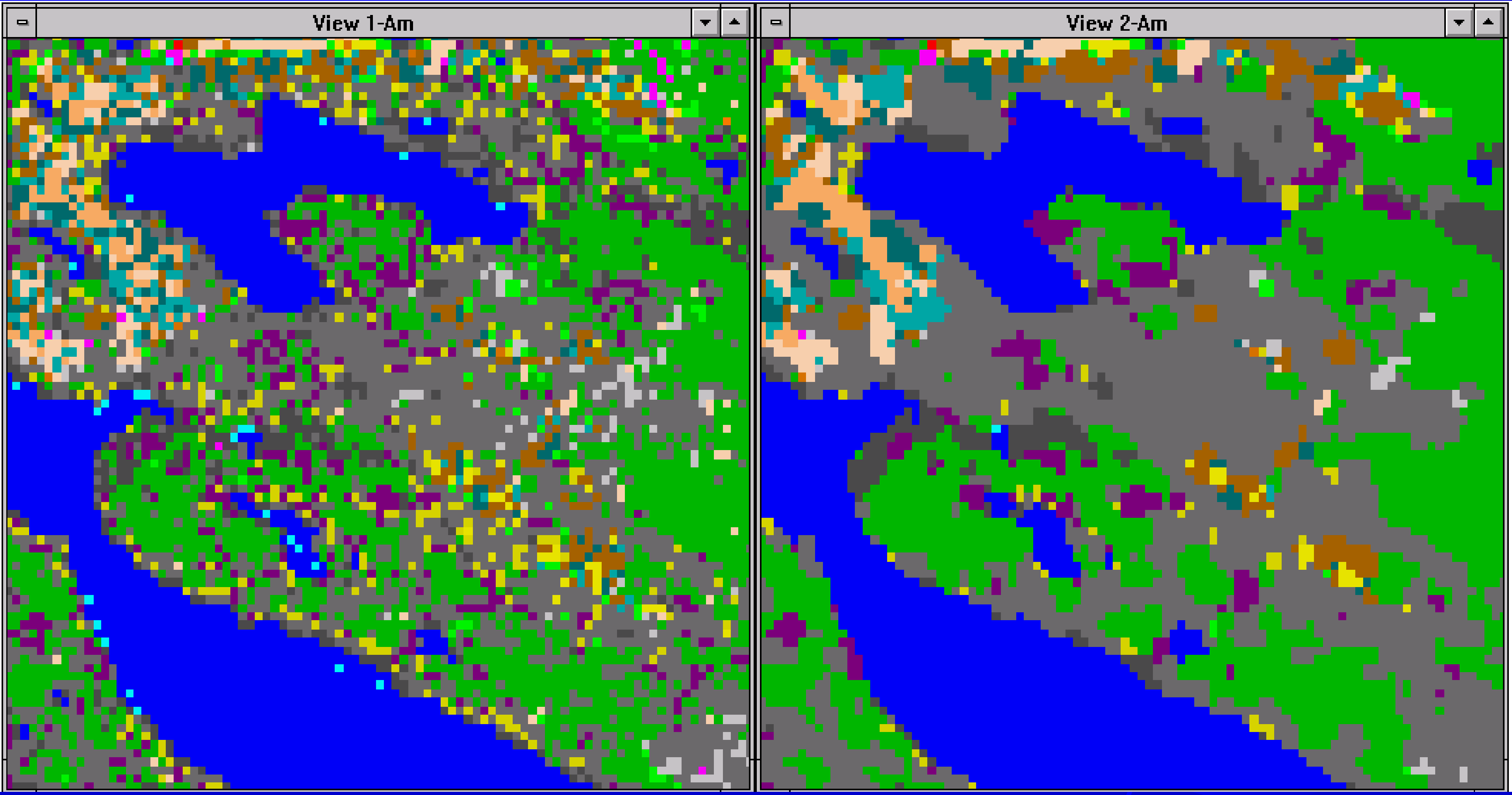
? More Myths ?

- Modal or Majority (mathematical) filters are useful tools for forming polygons
- Cleanup and develop separate themes which are merged to form a final land cover map

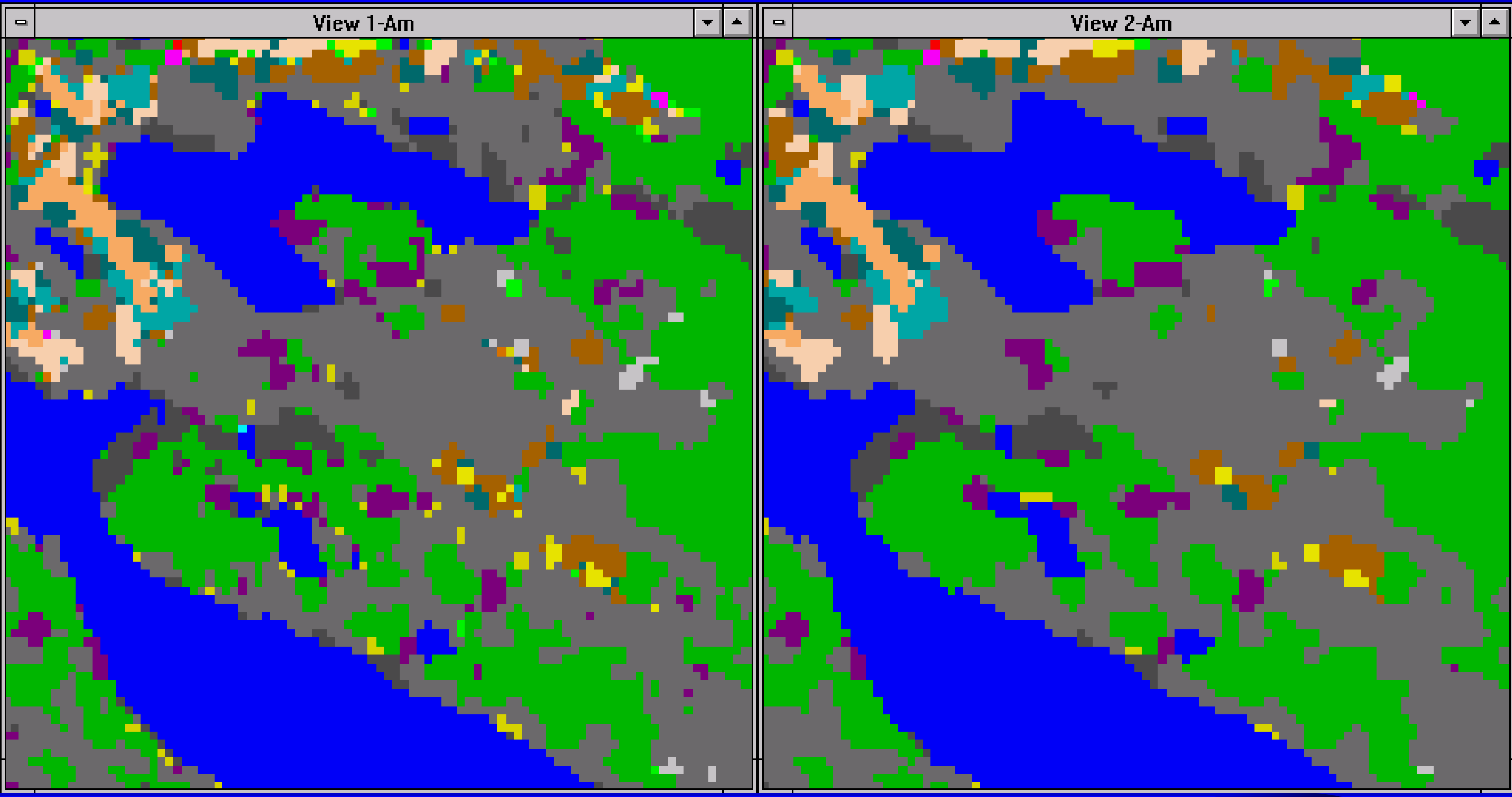
Modal Filtering - 4 Passes



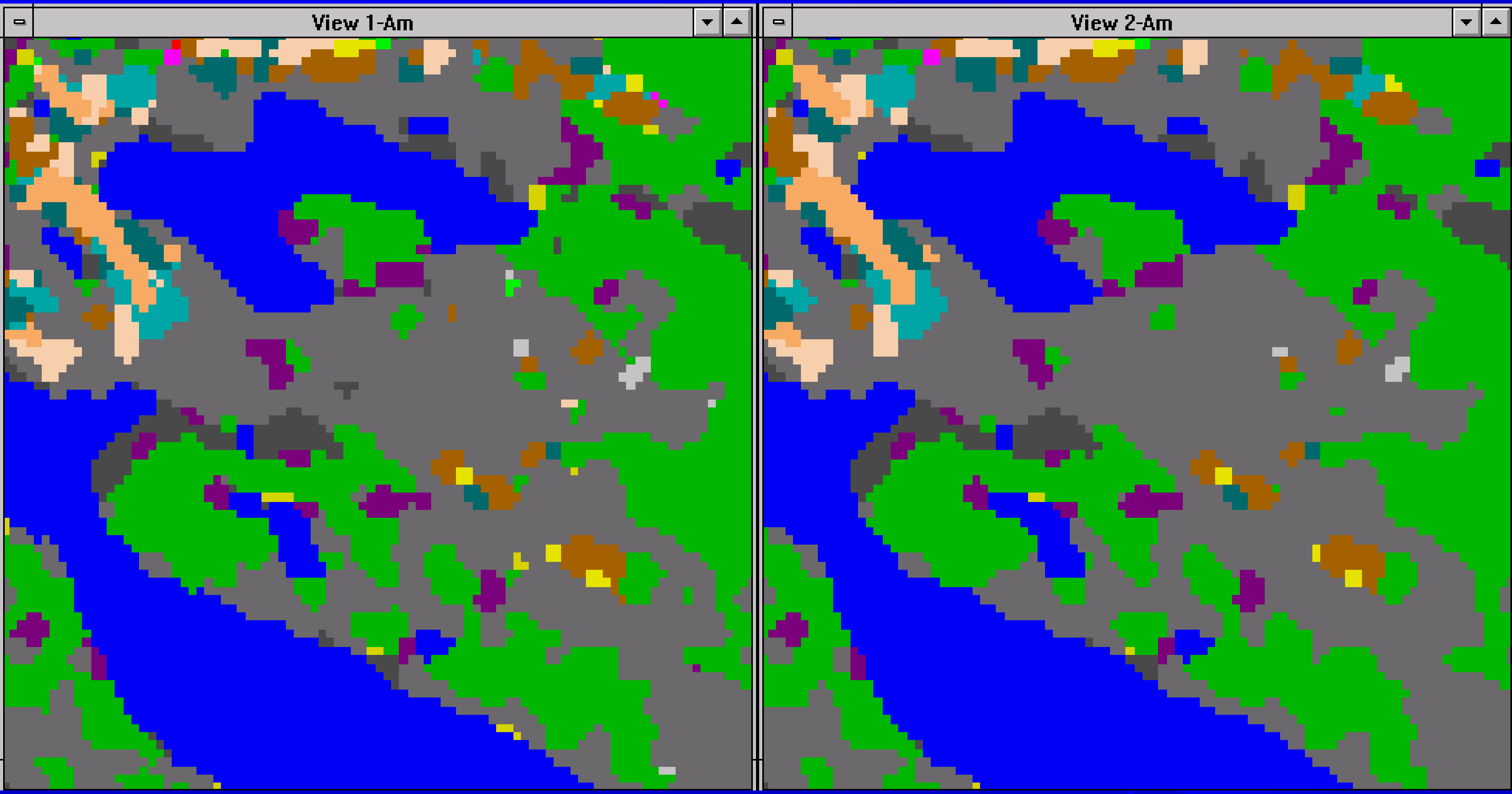
Modal Filtering - 1st Pass



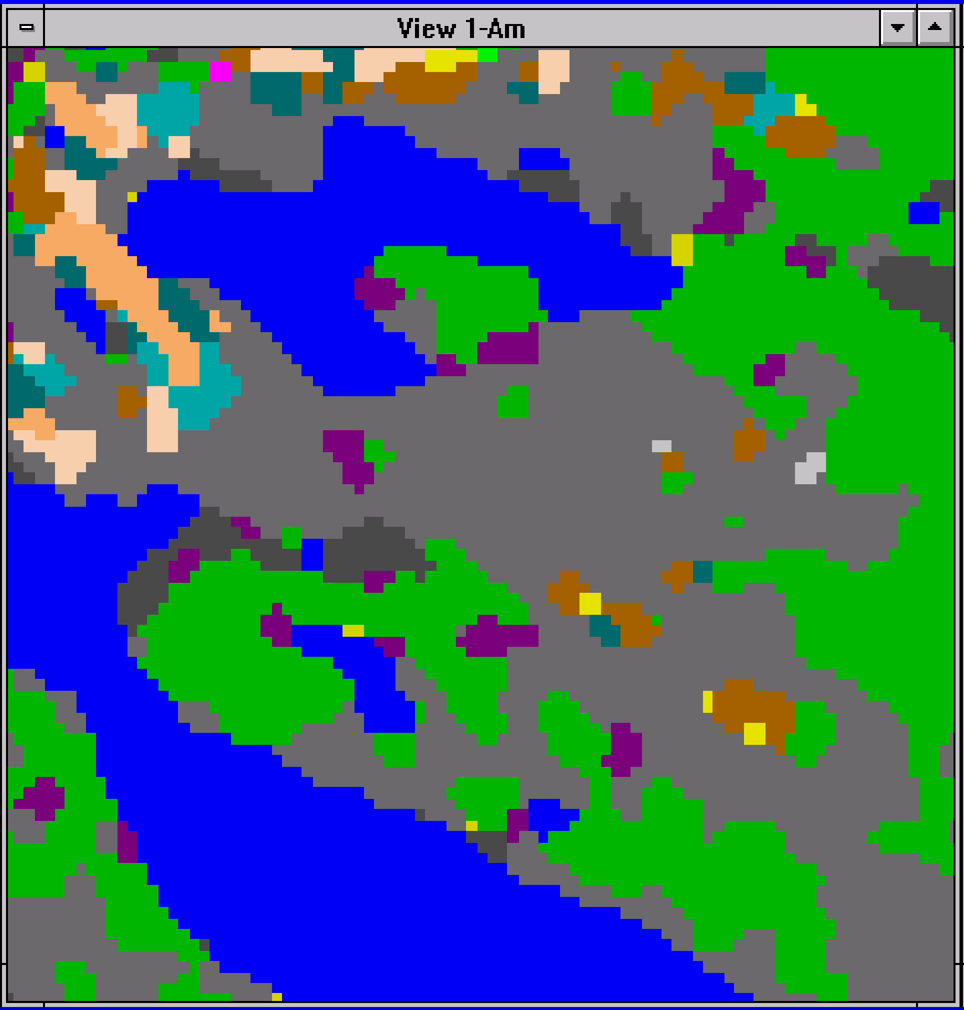
Modal Filtering - 2nd Pass

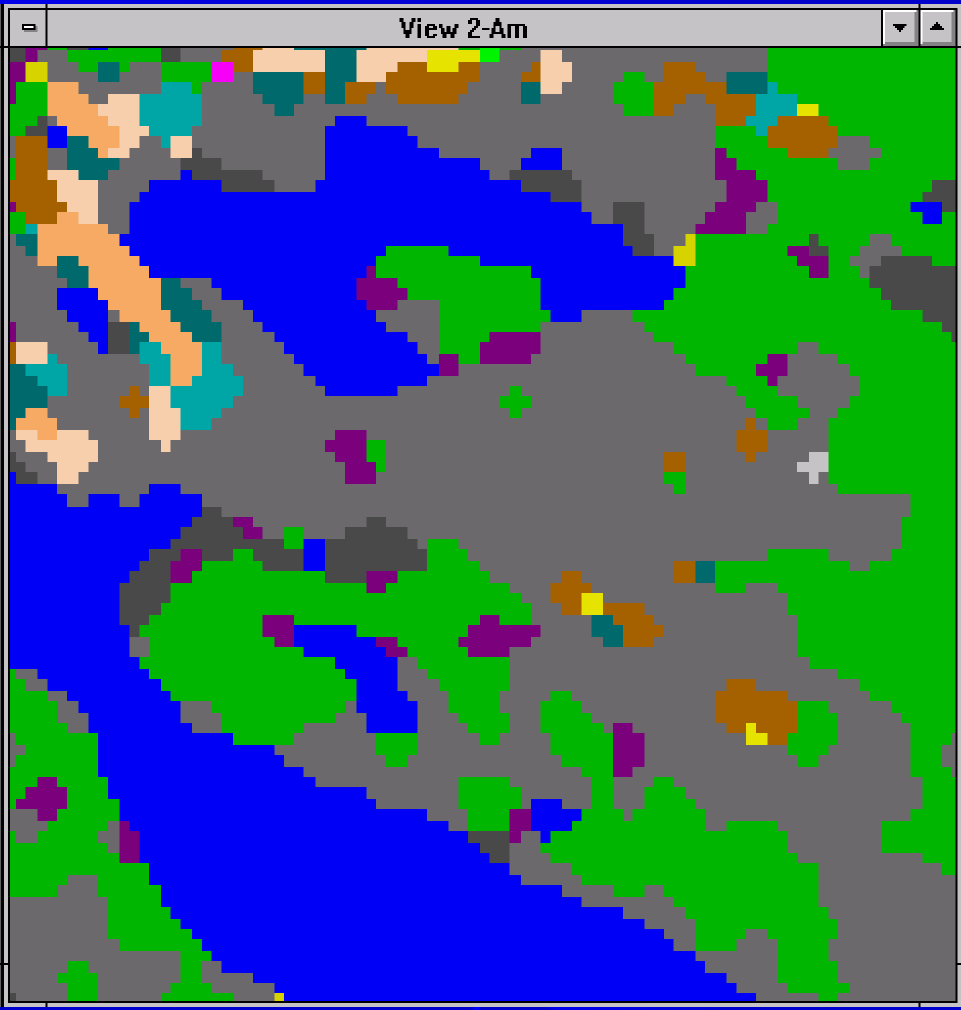


Modal Filtering - 3rd Pass



Modal Filtering - 4th Pass



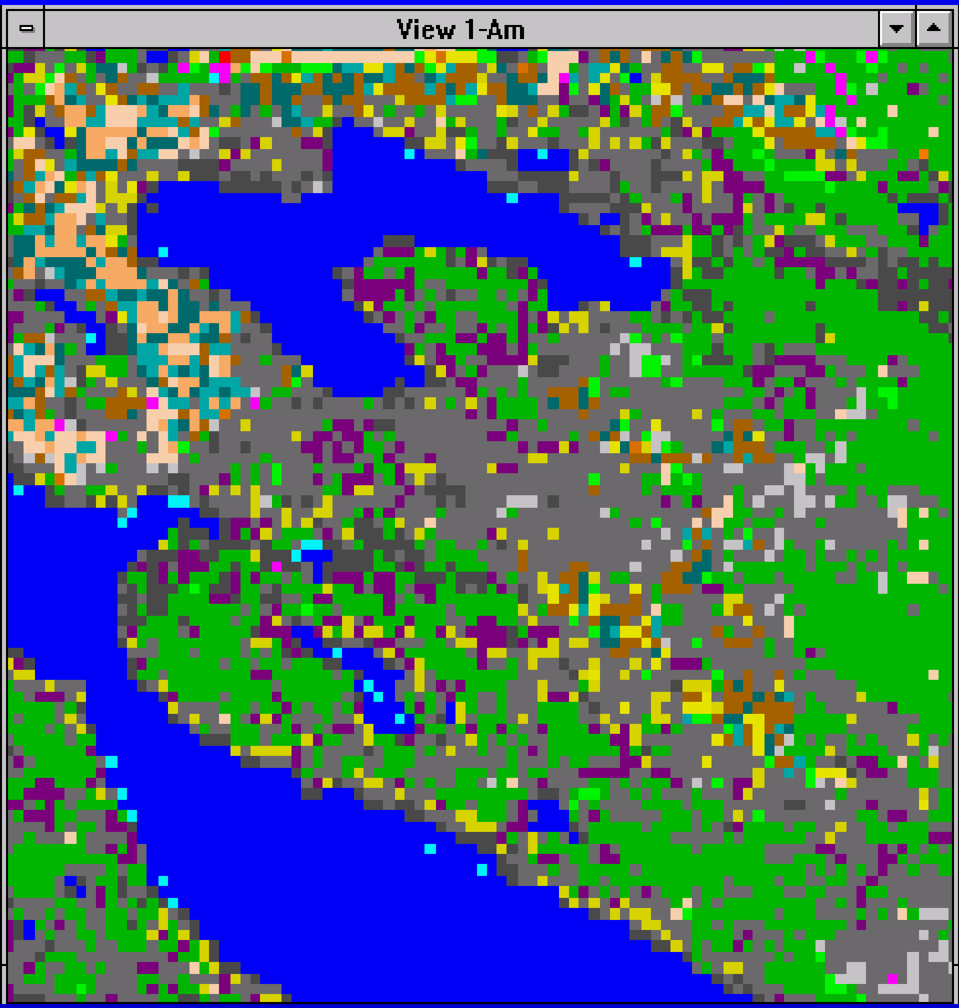


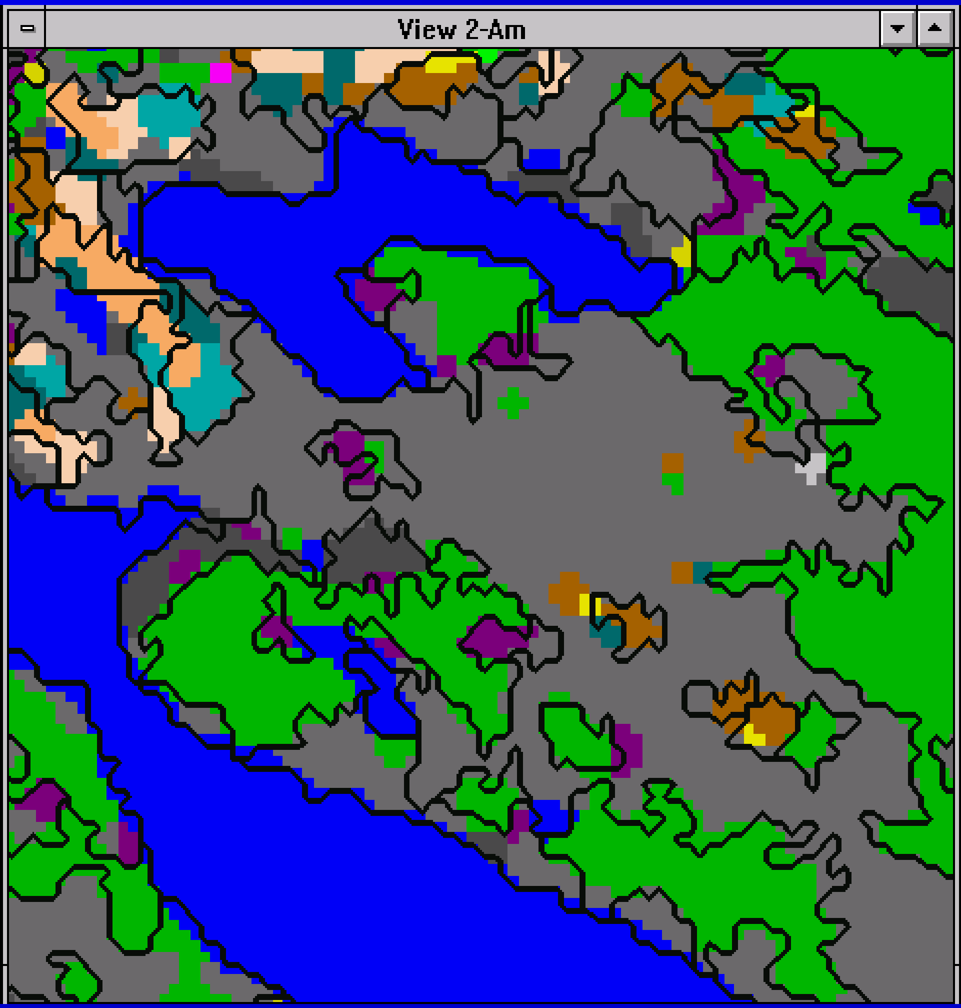
Reality

Mathematical Filters Do Not Approximate Ecological Relationships and Morphological Differences

- Feast or famine solution - e.g. shrub or tree when mixes should be developed
- Linear feature removal
- Edge degradation
- Minimum size problems - “When do you know you can stop filtering?”

Filtering Problems





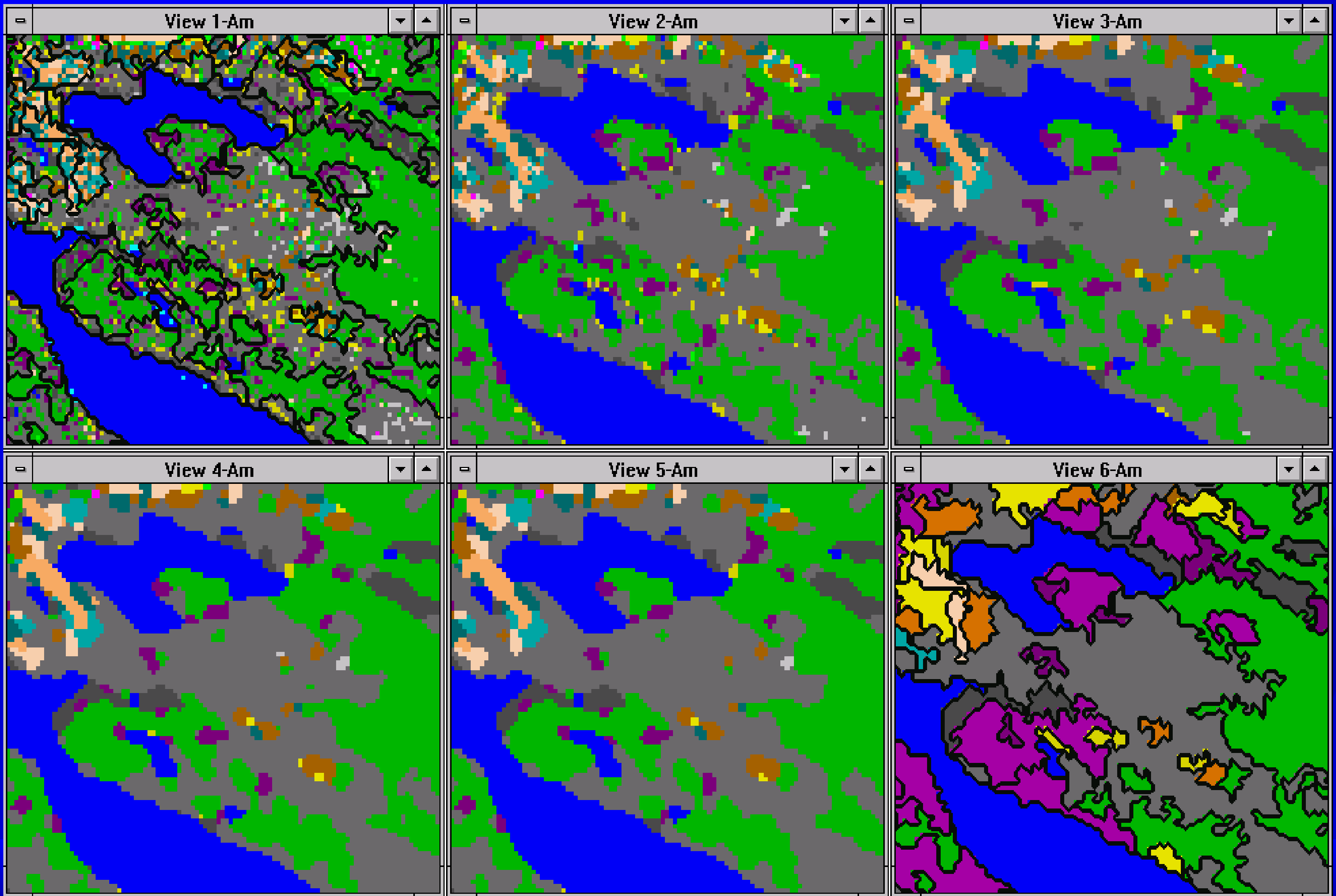
Reality - Vegetation/Land Cover Characteristics are Interrelated

- Cannot build separate themes and merge
  - make wrong decision about type boundaries
  - massive sliver problems
- Polygon attributes must be computed as weighted averages of values represented by pixels
- Polygons may yield new types not present in the classification

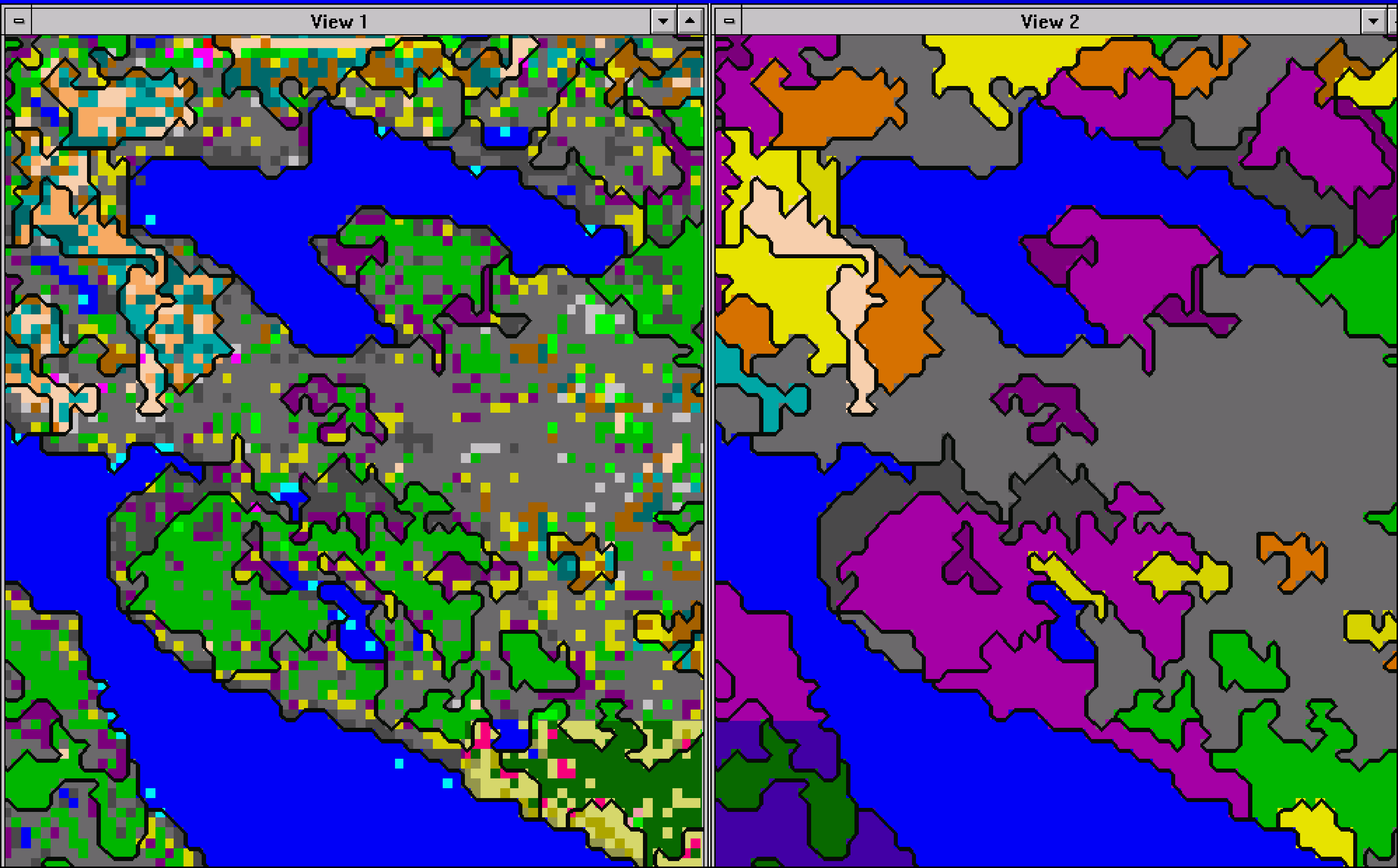
Solution: Ecological Rule-based Pixel Aggregation

- The classified pixel is a stratum in a stratification
- Each stratum represents a distinct set of cover, size, and species descriptions that are based on ground data collection efforts or other data descriptions
- Polygons are formed by grouping areas that have the most similar or related vegetation characteristics
- Process data until all polygons meet minimum mapping unit size limits

An Alternative Solution ...



… From Pixels to Polygons



Ecological Rule-Based Pixel Aggregation

- Based on:
  - Data
  - Rules and Relationships

Step 1: Develop Data Sets Representative of Pixel Data

- Ground Truth
- Classification Information and Associated Data

Data must represent all components of all types

Class: 2 **Name: White Spruce Open**

	% Cover	%Con/Hwd	% Species Cover
Trees: 42.50% cover comprised of:			
W Spruce	87.50%	87.50%	32.50%
B Spruce	87.50%	12.50%	4.50%
Total Conifer			**37.00%**
Hardwood	12.50%	100.00%	**5.50%**
Total Tree			**42.50%**

Shrubs:

Tall Shrub	45.00%
Low Shrub	5.00%
Dwarf Shrub	0.00%
Total Shrub	**50.00%**

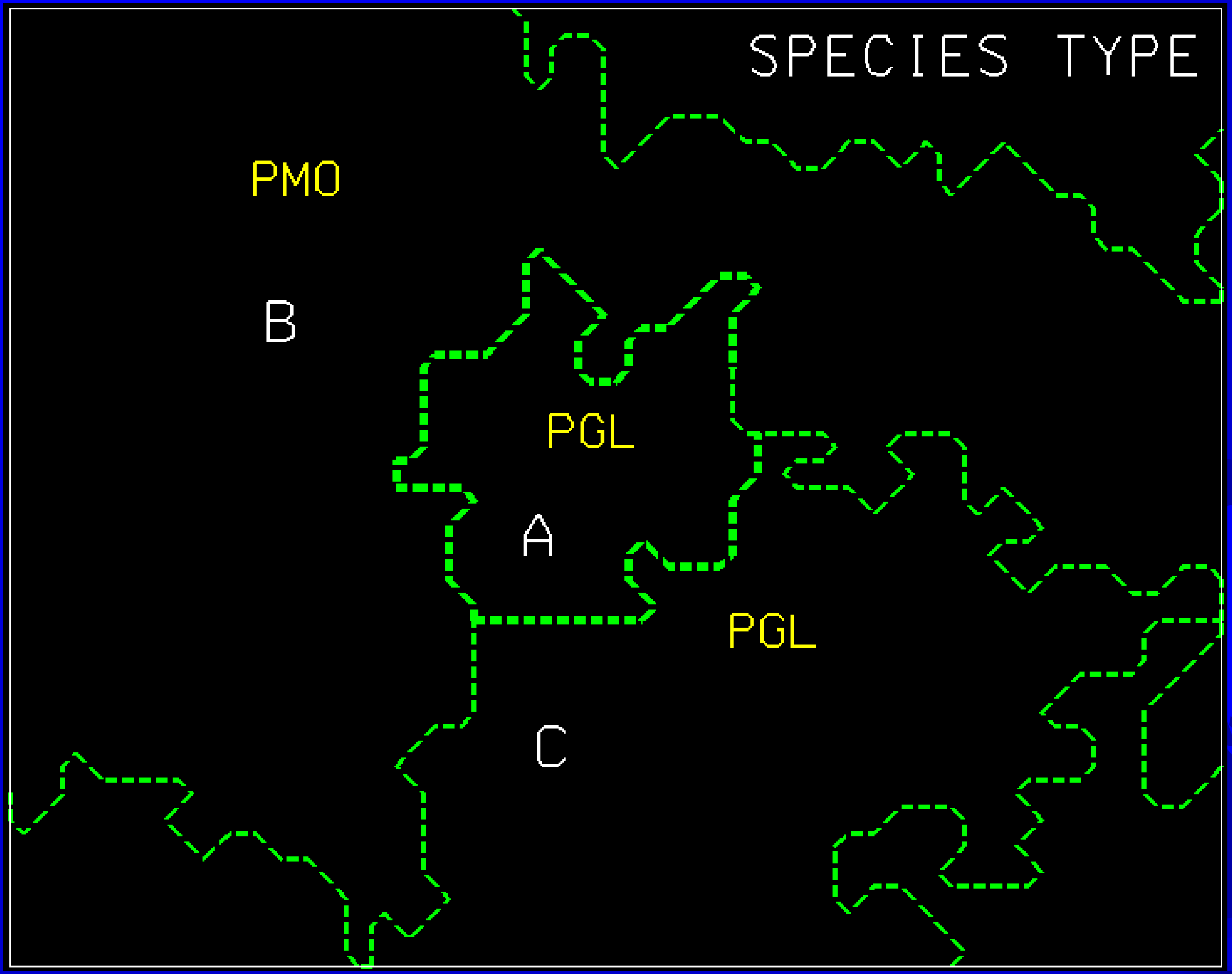
Forb:

Graminoid	0.00%
Forb	0.00%
Dry	7.50%
Wet	0.00%
Total Forb	**7.50%**

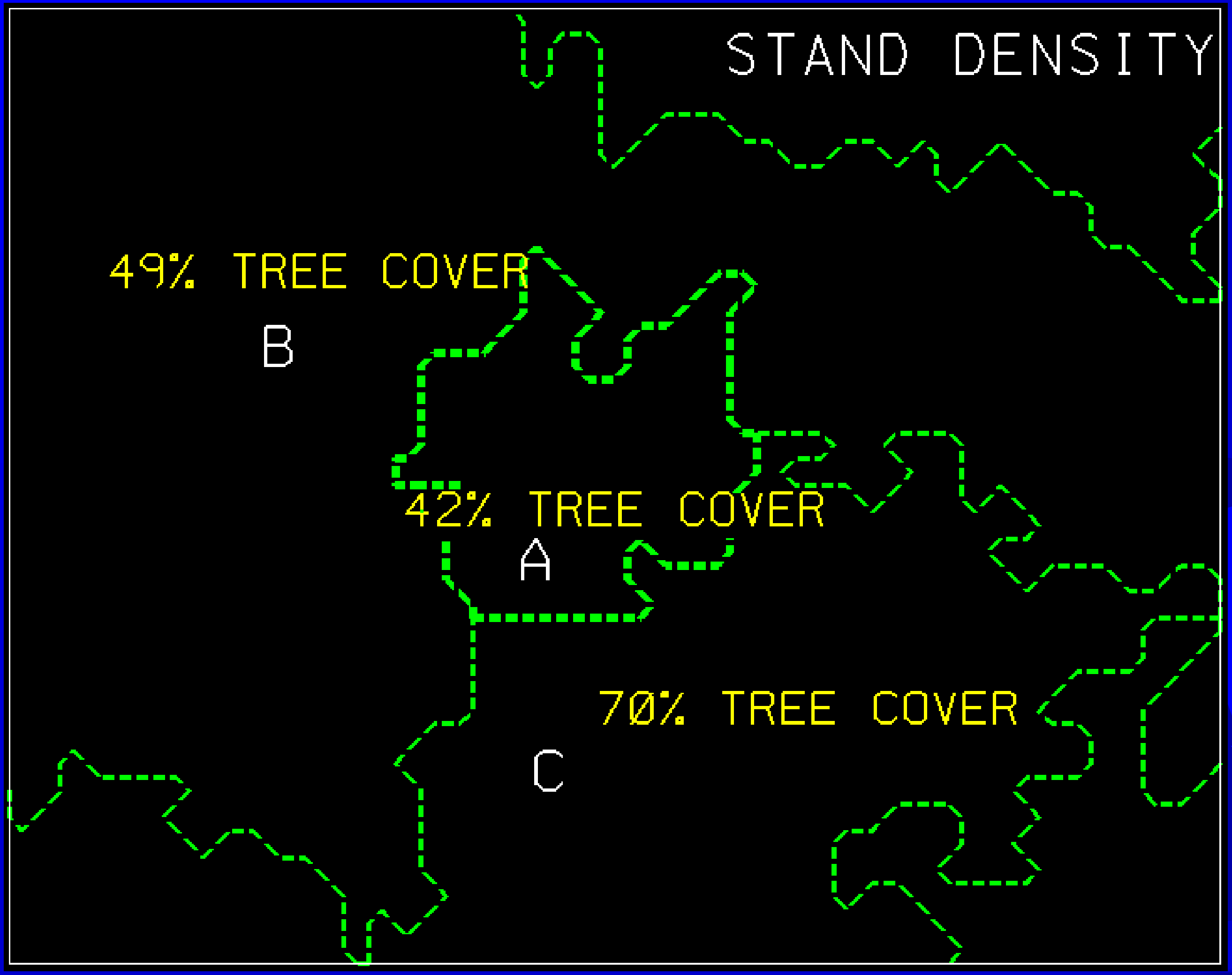
Step 2: Define Rules to Guide the Aggregation Process

- Vegetation classification definitions, relationships, and relative importance
- Minimum mapping unit size (by characteristic)

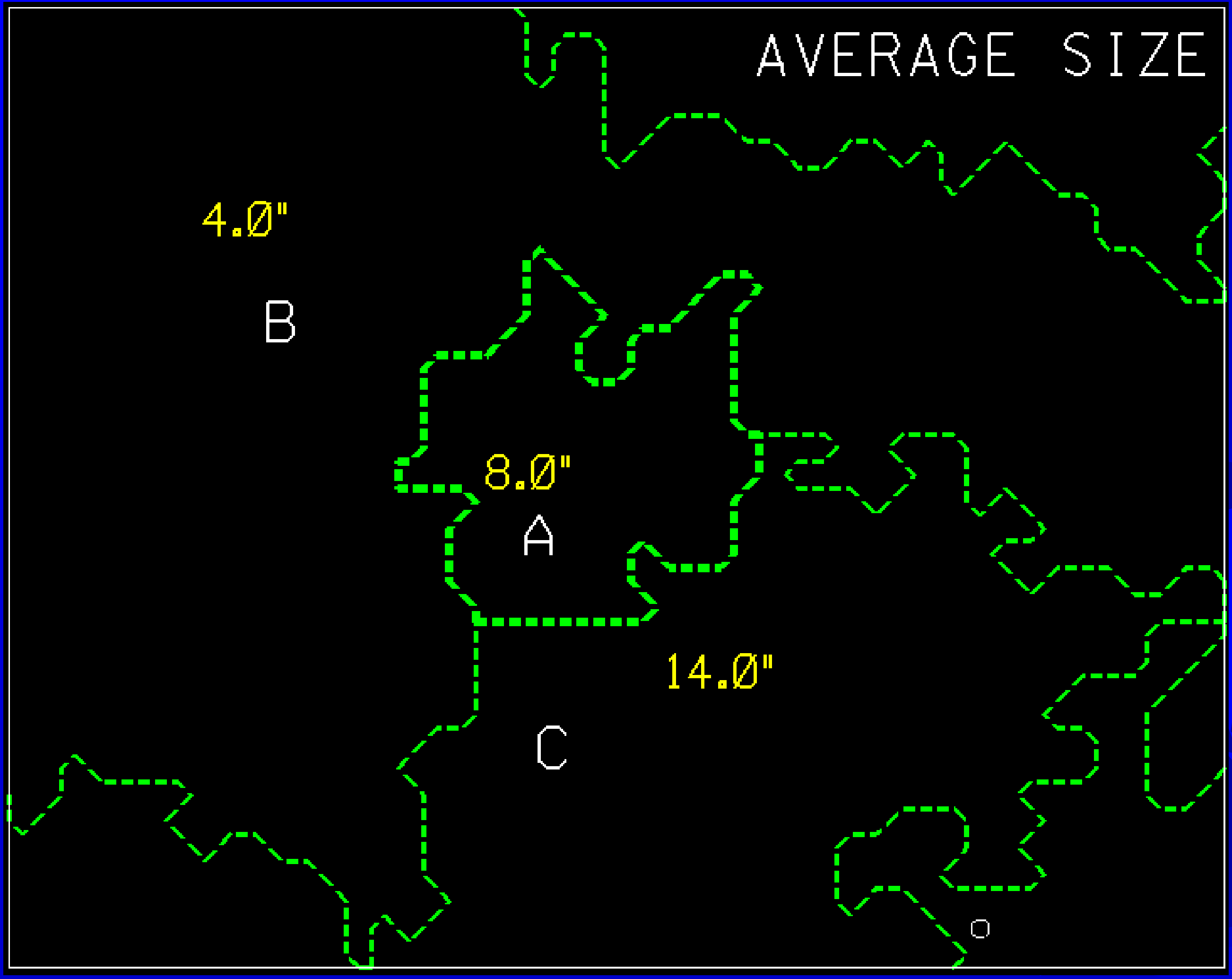
Similarity Of Features - Type



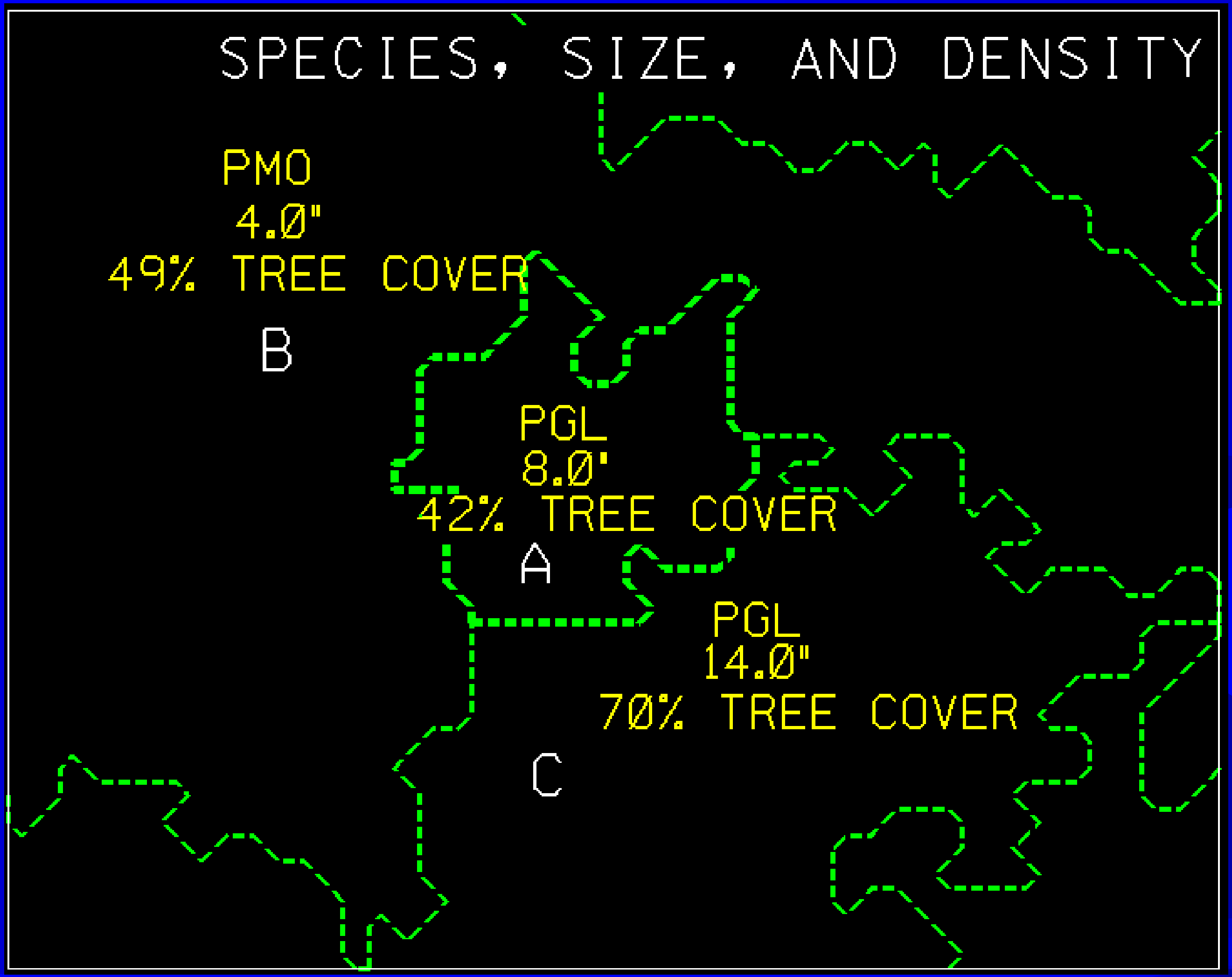
Similarity Of Features - Cover



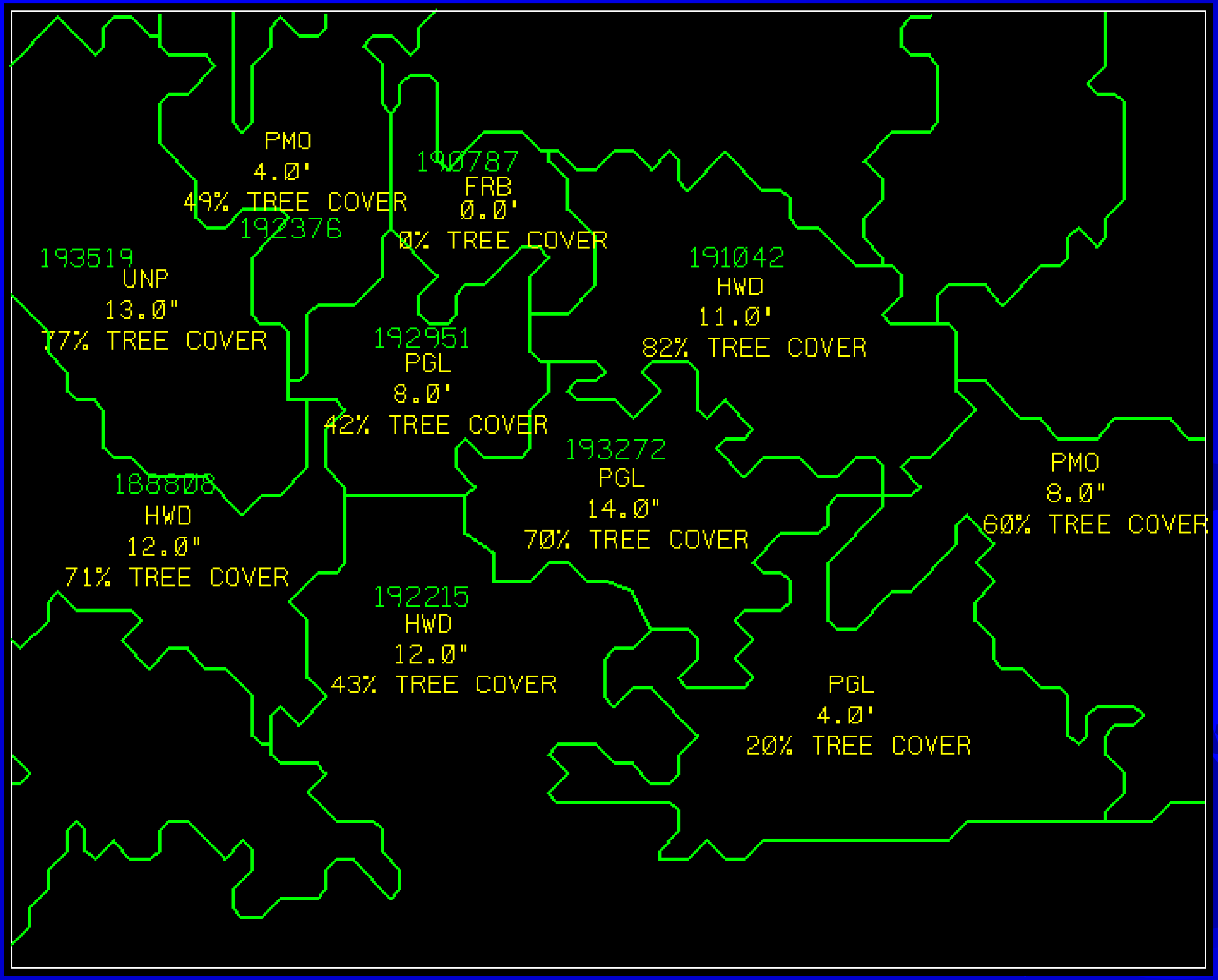
Similarity Of Features - Size



Similarity Of Multiple Features



Similarity Of Multiple Stands



Minimum Size Mapping Unit

- Degree of Similarity - Similar Vs Dissimilar
  - Desirable Limits
  - Critical Limits

Minimum Acreage Limits

Land Cover Type	Desirable Minimum Size	Critical Minimum Size
PGl	20.0 acres	3.0 acres
PMo	20.0 acres	3.0 acres
UnP	20.0 acres	3.0 acres
PHw	20.0 acres	3.0 acres
Hwd	20.0 acres	3.0 acres
TSh	20.0 acres	5.0 acres
LSh	20.0 acres	5.0 acres
DSh	20.0 acres	5.0 acres
MSh	20.0 acres	5.0 acres
Frb	20.0 acres	5.0 acres
Lch	5.0 acres	3.0 acres
H2O	5.0 acres	3.0 acres
…..		

Step 3: Evaluate Similarity and Merge With Most Similar

- Represent the rules as a function and attempt to quantify similarity
- For each subject area evaluate all adjacent areas and determine the most similar area
  - Merge the subject area into the most similar area
  - Recompute merged area attributes
- Stop when minimum mapping unit thresholds are met

Sample Similarity Estimates

Stand = 81373

----

stand#	iw	ip	cover	pctcon	shr	hrb	mtype	psp	ltype	pixels
81373	PGl	PGl	43	82	0.0	0.0	0	11	12	134
84939	PHw	PGl	35	57	0.0	0.0	0	11	20	55
* 10.5			2.0	2.5	0.0	0.0	0.0	0.0	6.0	

----

stand#	iw	ip	cover	pctcon	shr	hrb	mtype	psp	ltype	pixels
81373	**PGl**	**PGl**	**43**	**82**	**0.0**	**0.0**	**0**	**11**	**12**	**134**
82936	**PGl**	**PGl**	**55**	**85**	**0.0**	**0.0**	**0**	**11**	**12**	**55**
*** 3.3**			**3.0**	**0.3**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	

----

stand#	iw	ip	cover	pctcon	shr	hrb	mtype	psp	ltype	pixels
81373	PGl	PGl	43	82	0.0	0.0	0	11	12	134
85658	Hwd	Hwd	39	22	0.0	0.0	0	22	25	20
* 22.3			1.0	6.0	0.0	0.0	0.0	5.5	9.8	

Aggregate stand 81373 with stand 82936

Step 4: Report Polygon Attributes

- Summarize weighted averages of pixel characteristics within the polygon boundary
- Develop discrete estimates and variances from weighted averages as polygon attributes
  - variance of tree cover is related to spatial distribution of cover
  - variance of tree size is related to stand structure
- Develop single theme maps from polygon map through reclassification of database characteristics
- Develop categorical estimates from discrete estimates - no need to jaywalk anymore

Land Cover Density Summary:
Stand ID: 1789
Total Number of Pixels: 50
Contributing Pixels: 50

Size Class:	0-4"	5-8"	9-12"	13"+	Total	Other	Total
White Spruce	0.0%	0.0%	27.1%	0.0%	27.1%		27.1%
Black Spruce	0.0%	0.0%	10.2%	0.0%	10.2%		10.2%
Hardwood	0.0%	0.0%	6.4%	0.0%	6.4%		6.4%
Tall shrub						39.8%	39.8%
Low shrub						8.1%	8.1%
Wet moss						1.8%	1.8%
Lichen						5.3%	5.3%
Total Cover	0.0%	0.0%	43.7%	0.0%	43.7%	56.3%	100.0%
Total Tree Cover					**43.7%**		

Stand Tree Composition Summary:
Stand: 1789

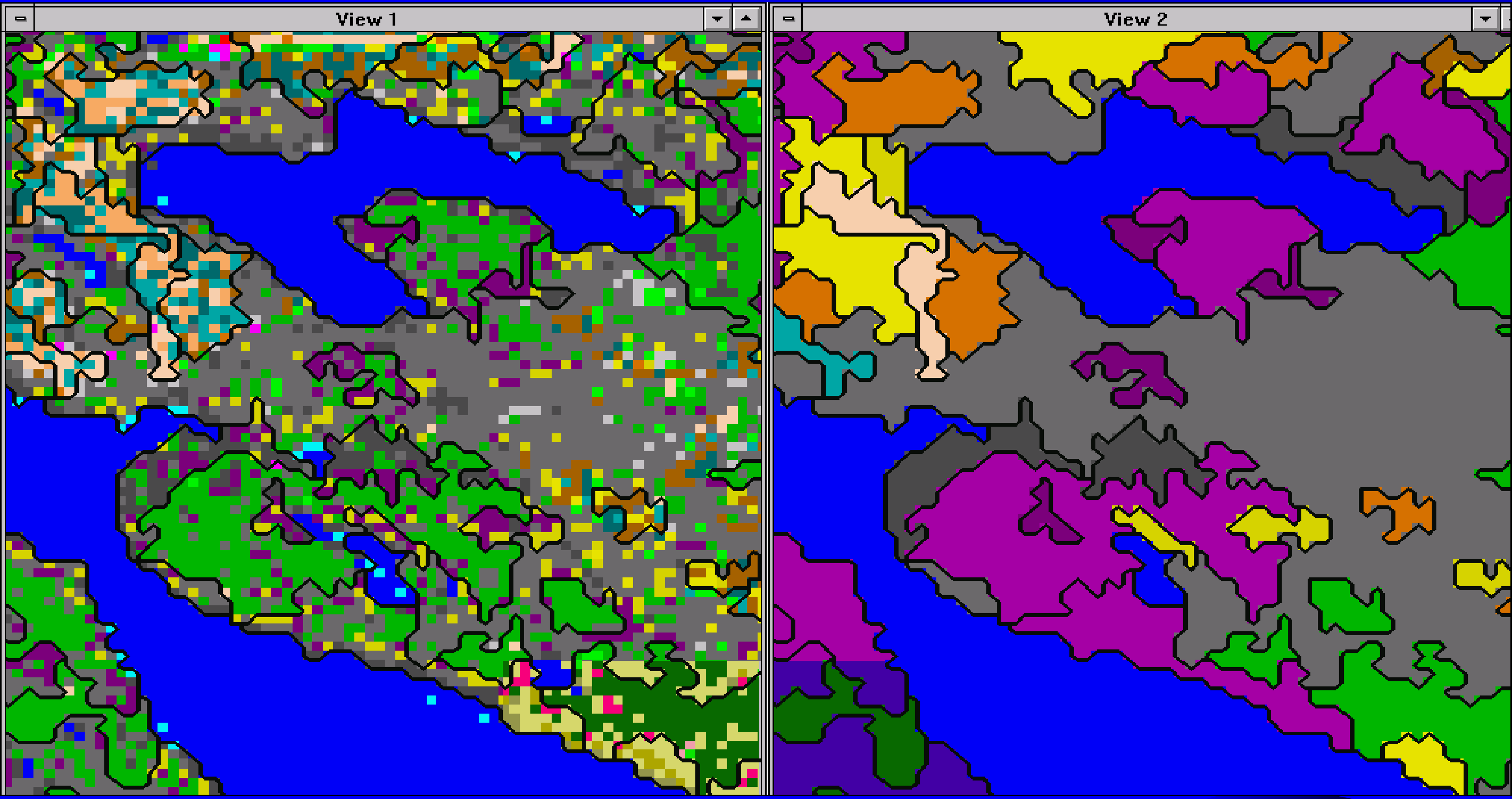
Size Class:	0-4"	5-8"	9-12"	13"+	Total
White Spruce	0.0%	0.0%	62.0%	0.0%	62.0%
Black Spruce	0.0%	0.0%	23.3%	0.0%	23.3%
Hardwood	0.0%	0.0%	14.7%	0.0%	14.7%
Total Tree Cover	0.0%	0.0%	100.0%	0.0%	100.0%

Table 7: Polygon Cover Description

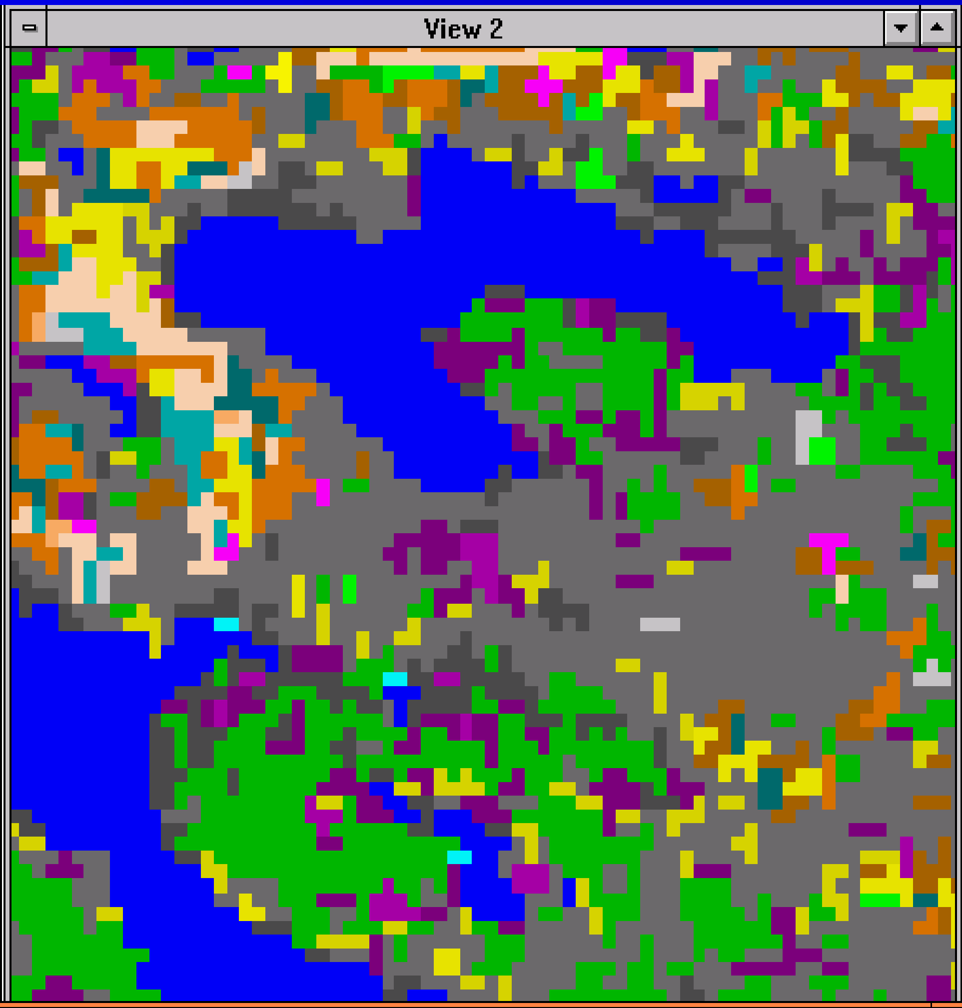
Table 9: Database Record Listing

wrangle_-id	[172598]
mapid	[100064]
lform	[s]
ltype	[UnP]
closure_class	[2]
density	[43.7]
pct_conifer	[85.3]
pct_hdwood	[14.7]
pr_species	[White Spruce]
pred_sp_pct	[62.0]
other_cover	[56.0]
cv_shr	[48.3]
cv_hrb	[7.6]
cv_bar	[0.3]
cv_oth	[0.0]
pix_ct	[50]
grid_val	[**1789**]
class_status	[8]
acreage	[11.1504]

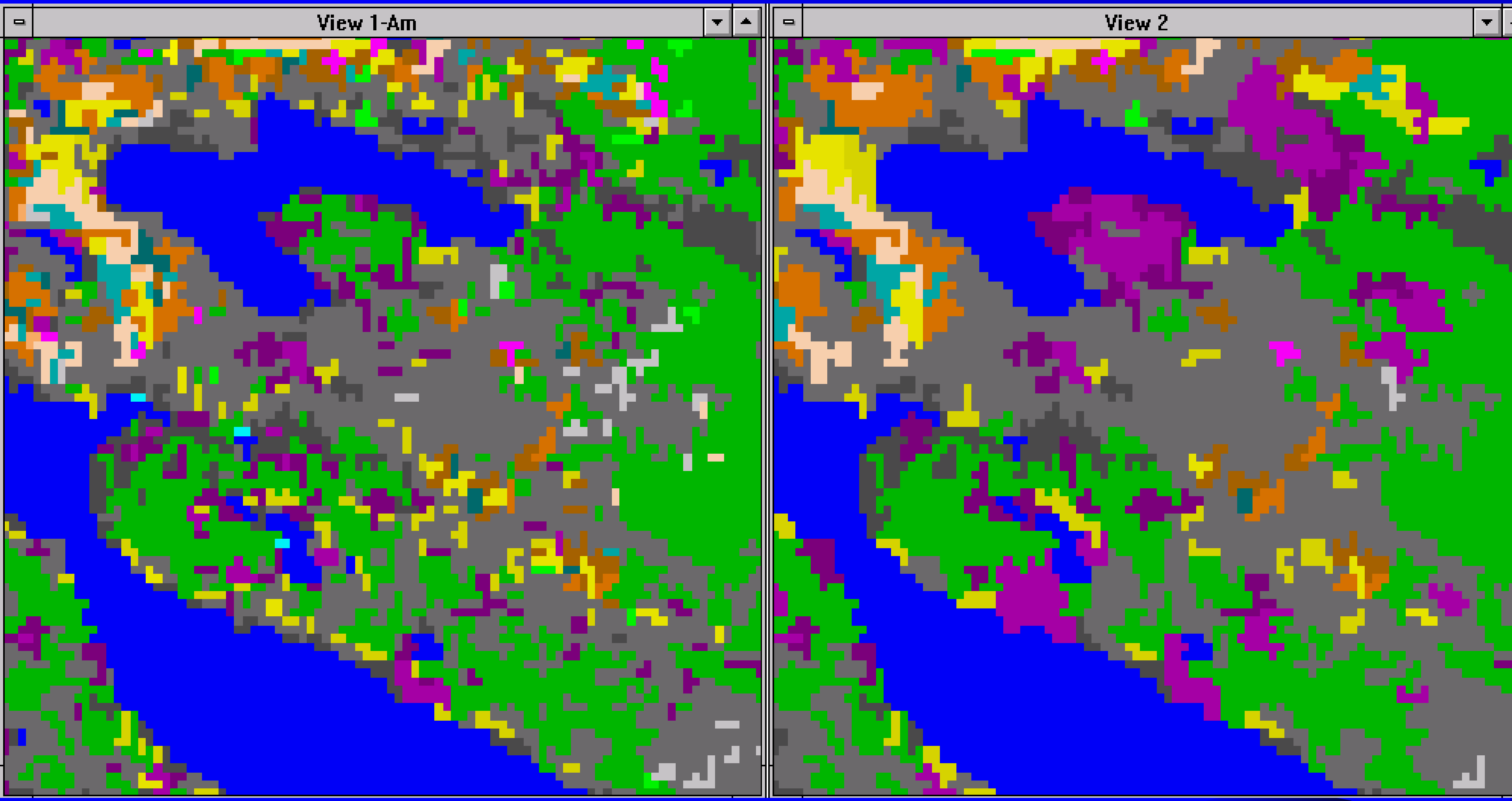
From Pixels to Polygons



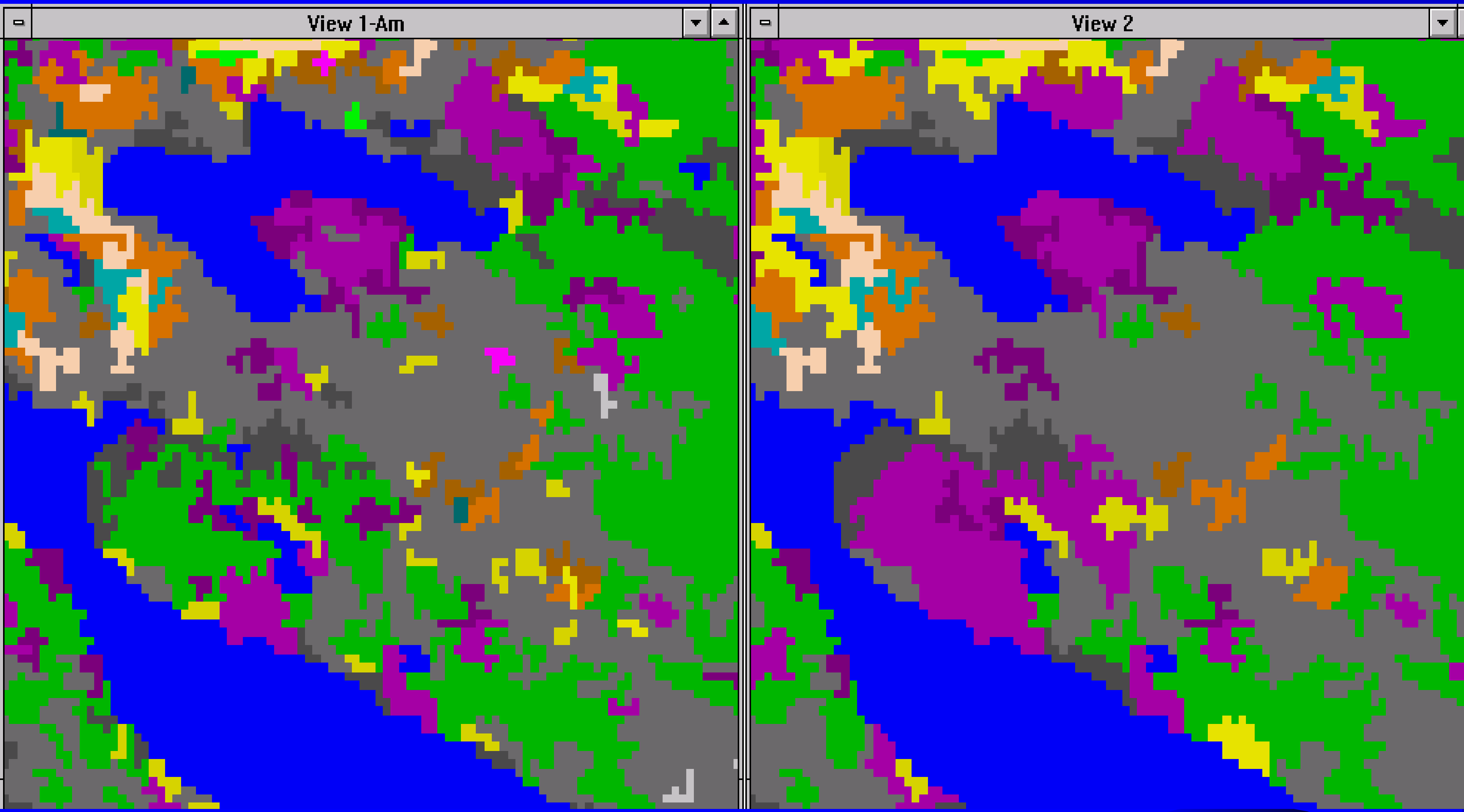
Iteration 1



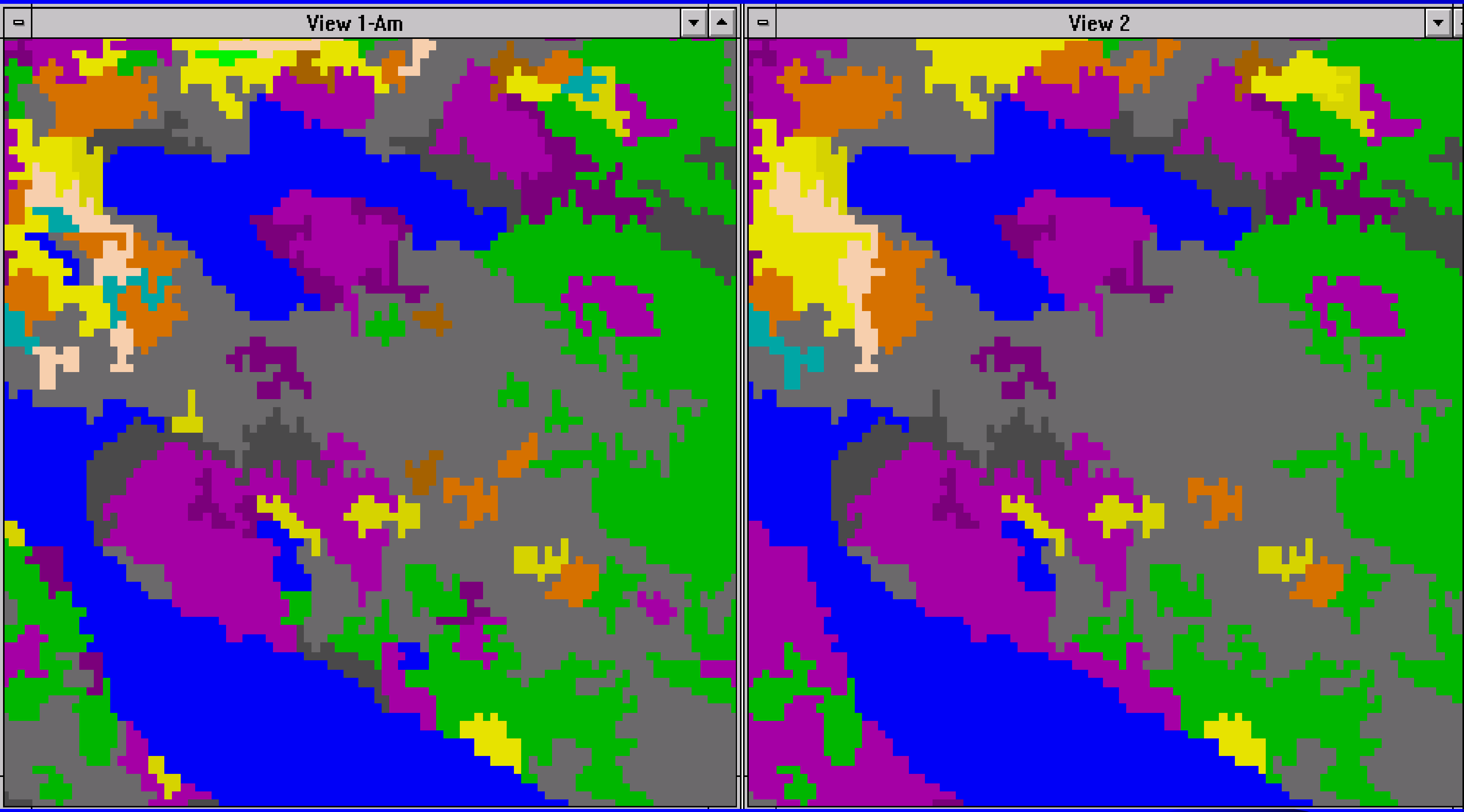
Iteration 2



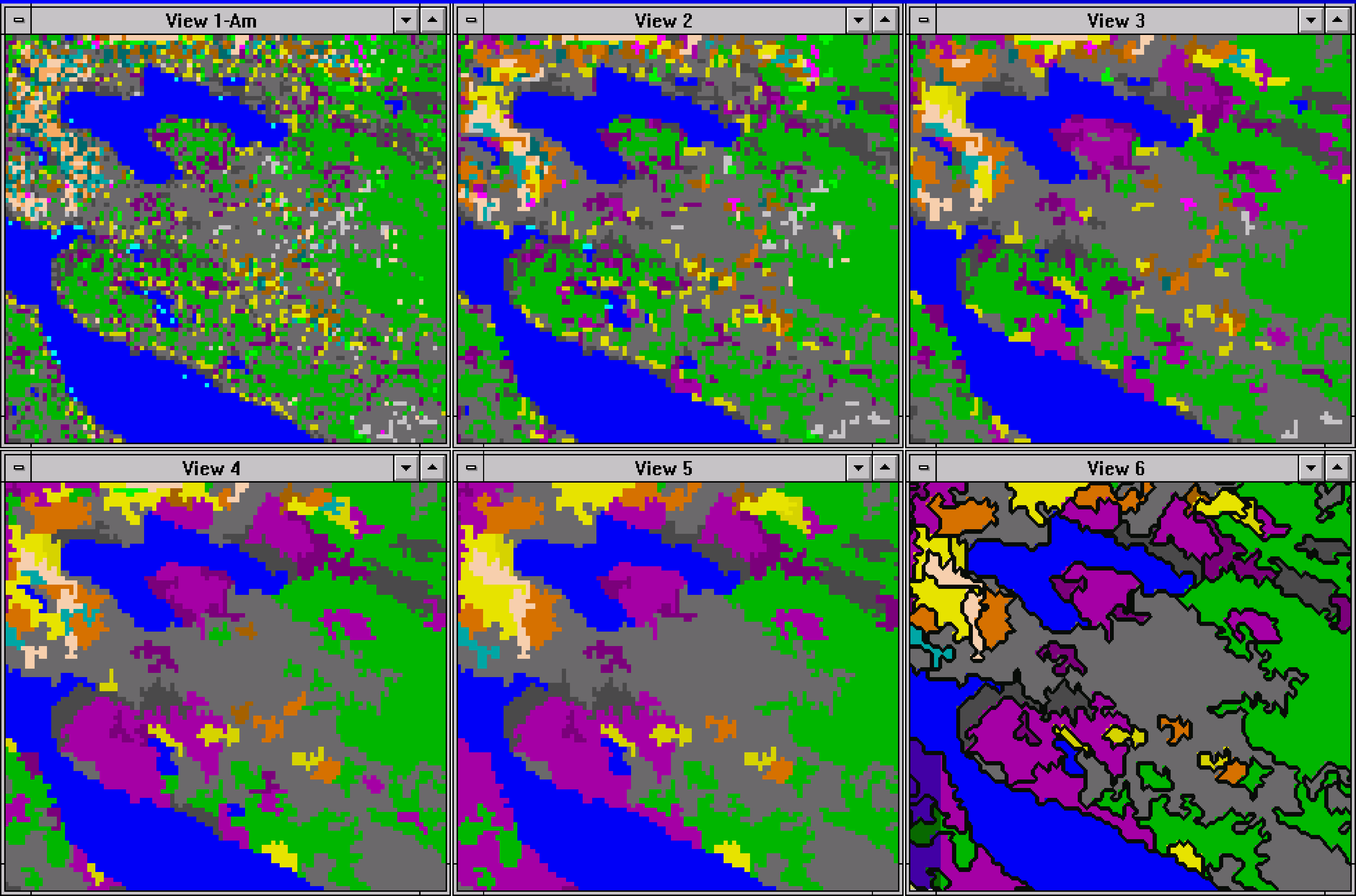
Iteration 3



Iteration 4



Start to Finish



Aggregation Results

- Overall balance of acreage by general type
- Movement towards mixed specie types
  - Unspecified Spruce
  - Spruce/Broadleaf
- Movement towards moderate density classes
- Development of new types
- Development of life form estimates based on attribute descriptions

Side by Side Results

Rule-based Aggregation Benefits

- Process millions of acres at one time
- Repeatable, consistent, and objective
- No human digitizing or editing of stand boundaries
- Can modify rules to change emphasis and produce different maps.
- Can aggregate using different vegetation classification schemes to develop different maps
- Similarity of values, <u>not classes</u>, yield polygons with lower within stand variation.
- Discrete estimates allow reclassification by user defined classes