Quantitative accuracy assessment

Developed by remote sensing specialists at the USFS Geospatial Technology and Applications Center (GTAC), located in Salt Lake City, Utah

Accuracy Assessment

What is it and what is its role?

- Accuracy Assessment…
  - is the process by which the accuracy or correctness of an image classification is evaluated
  - involves the comparison of the image classification to reference data that are assumed to be true
  - supports the spatial information used in the decision making process

Importance of Accuracy Assessments

Measure uncertainty (truth vs. agreement with reference data)

- Accuracy assessments are required for decision-making
- They allow us to compare approaches

Accuracy Assessment Process

1. Sample design
    1. sampling methods
    2. sample size

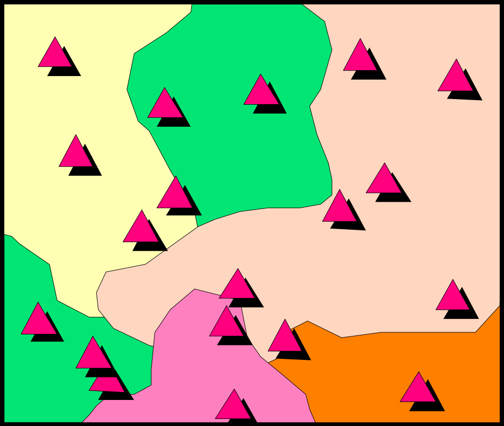
Sample Design

Objectives…

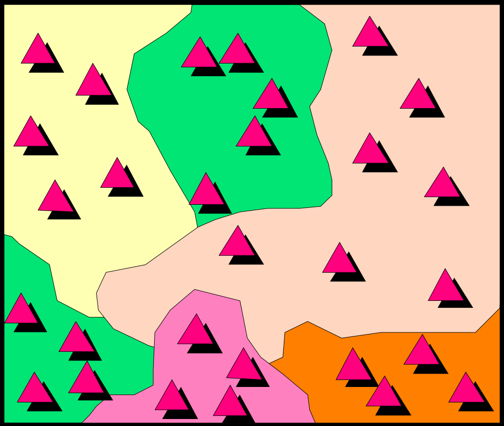
- The major objective of the sample is to provide information that is representative of the population
  - It's easy to calculate sample statistics
  - If the sample is representative of the population, then the sample statistics are good and valid estimates of the population parameters
- However, sampling is very expensive. We must balance sampling validity with cost when considering:
  - Sample size
  - Site locations (remote sites are very expensive)
  - Field measurement requirements etc...

Sampling Methods

Several Major Types (slide 1 of 3):



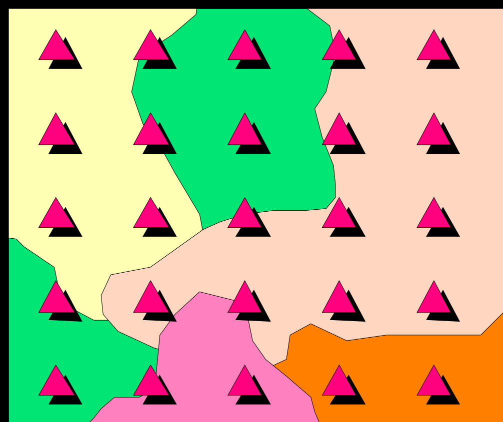
Simple Random Sampling:
observations are randomly placed.



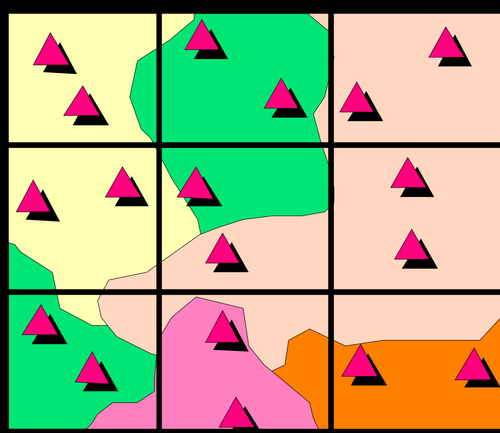
Stratified Random Sampling:
A minimum number of observations
Are randomly placed in each category.

Sampling Methods

Several Major Types (slide 2 of 3):



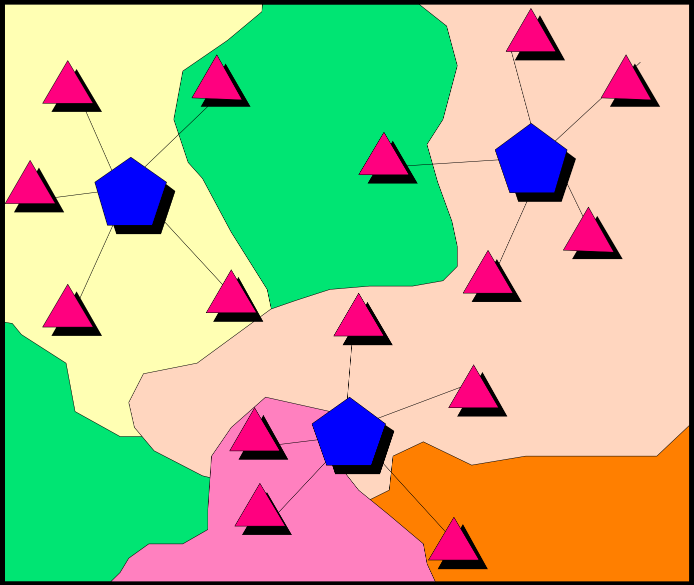
<u>Systematic Sampling</u>: observations are placed at equal intervals according to a strategy.



<u>Systematic Non-Aligned Sampling</u>: a grid provides even distribution of randomly placed observations.

Sampling Methods

Several Major Types (slide 3 of 3):



<u>Cluster Sampling</u>: Randomly placed "centroids" used as a base of several nearby observations. The nearby observations can be randomly selected, systematically selected, etc…

Accuracy Assessment Process

1. Sample design
2. Response design
    - Spatial assessment unit, sample frame
    - Sources of data (high res imagery, Landsat time series, field data)
    - Labeling protocol, schema
    - Defining agreement, map unit key
    - Reference classification uncertainty: geolocation and interpreter variability

Reference Data Collection

Reference data are assumed true!

- Sources of reference data may be:
  - An existing map
  - Existing inventory data
  - Photo-interpreted
  - Collected on the ground (usually preferred although most expensive)

Reference Data Collection

Reference data are assumed true!

- To avoid introducing errors because of poor reference data, the reference data must:
  - Be consistent (use standardized forms)
  - Use the same classification system as the map
  - Account for changes over time
  - Be precisely located
  - Have quality control

Assessing Accuracy

Budget Factors...

- By far, the most expensive component of an accuracy assessment is the reference data collection:
  - Number of sample sites
  - Type of sites (e.g., field vs photo)
  - Distribution of sites and ease of access to sites
  - Amount of data collected at each site
  - Level of effort required for measurement on the site
  - Reference data quality control.

Accuracy Assessment Process

1. Sample design
2. Response design
3. Analysis: error matrix

Reference Data

Map Data		CH	NCH	ROW TOTAL
	CH	27	6	33
	NCH	4	63	67
	COLUMN TOTAL	31	69	100

The Error Matrix

Comparing the reference data to the classified map...

- An error matrix is the standard way of presenting an accuracy assessment
  - The error matrix is a square array in which accuracy assessment sites are tallied by both their classified category in the image and their actual category according to the reference data
  - For example, you classified an image into 3 categories: Hardwoods, Conifers and Other. You field collected reference data from 100 randomly selected sites and built this 3x3 error matrix…

Error Matrix Components

1 of 3

Reference Data is referenced along the top.

Classification Data is referenced along the side.

		Reference			
		H	C	O	Σ Row
Classification	H	28	14	15	57
	C	1	15	5	21
	O	1	1	20	22
	Σ Col	30	30	40	100

Error Matrix Components

2 of 3

Diagonals: these represent the sites that were classified correctly according to the reference data.

Reference

Classification	H	C	O	Σ Row
H	28	14	15	57
C	1	15	5	21
O	1	1	20	22
Σ Col	30	30	40	100

Row Totals: used to give individual class accuracies.

Column Totals: used to give individual class accuracies.

Error Matrix Components

3 of 3

Off-diagonals: these represent the sites that were misclassified according to the reference data.

Reference

	H	C	O	Σ Row
H	28	14	15	57
C	1	15	5	21
O	1	1	20	22
Σ Col	30	30	40	100

Classification

Off-diagonals.

Measures of Accuracy

Overall Accuracy...

Reference

Classification	H	C	O	Σ Row
H	28	14	15	57
C	1	15	5	21
O	1	1	20	22
Σ Col	30	30	40	100

Overall accuracy is given by the sum of the diagonals divided by the grand total (expressed as a percent):

3 (diagonals = 28 + 15 + 20
=63

Overall accuracy = 63/100
=63%

Measures of Accuracy

Individual Class Accuracy...

		Reference			
		H	C	O	Σ Row
Classification	H	28	14	15	57
	C	1	15	5	21
	O	1	1	20	22
	Σ Col	30	30	40	100

Individual Class Accuracy is given by the diagonal value divided by the row or column total (expressed as a percent). For example, the conifer class:

Conifer = 15/21 = 71%

Or…

Conifer = 15/30 = 50%

But how can there be two values and what do they mean…?

Measures of Accuracy

Individual Class Accuracy...

How can there be two values and what do they mean...?

- There are two types of errors: Omission and Commission
- Any site that is omitted from the correct class is committed to an incorrect class
- These two types of errors correspond to two types of class accuracies: User's and Producer's...

Measures of Accuracy

Producer's Accuracy

- The percentage of time a class identified on the ground is classified into the same category on the map.
- (Forest) Omission – When forest is classified as other.

User's Accuracy

- The percentage of time a class identified on the map is classified into the same category on the ground.
- (Forest) Commission – When other is classified as forest.

Measures of Accuracy

Individual Class Accuracy—User's...

		Reference			
		H	C	O	Σ Row
Classification	H	28	14	15	57
	C	1	15	5	21
	O	1	1	20	22
	Σ Col	30	30	40	100

<u>User's Accuracy</u> = the diagonal value divided by the row total. For example, the conifer class:

Conifer = 15/21 = 71%

This is a measure of commission error, and indicates the probability that a pixel classified into a given category actually represents that category on the ground.

Note: the formulas given only apply if Reference labels are along the top and Classification Labels are along the side (as shown)...

Measures of Accuracy

Individual Class Accuracy—Producer's...

		Reference			
		H	C	O	Σ Row
Classification	H	28	14	15	57
	C	1	15	5	21
	O	1	1	20	22
	Σ Col	30	30	40	100

Producer's Accuracy = the diagonal value divided by the column total. For example, the conifer class:

Conifer = 15/30 = 50%

This is a measure of omission error, and indicates how well training set pixels of a given cover type are classified.

Note: the formulas given only apply if Reference labels are along the top and Classification Labels are along the side (as shown)...

Measures of Accuracy

Individual Class Accuracy...

Reference

Classification	H	C	O	Σ Row
H	28	14	15	57
C	1	15	5	21
O	1	1	20	22
Σ Col	30	30	40	100

User's Accuracies:

Hardwood = 28/57 = 49%

Conifer = 15/21 = 71%

Other = 20/22 = 91%

Producer's Accuracies:

Hardwood = 28/30 = 93%

Conifer = 15/30 = 50%

Other = 20/40 = 50%

Error vs Difference Matrix

The error matrix really measures <u>differences</u> between reference and classification labels and may not be classification <u>error</u>...

- Some reasons for differences:
  - Classification errors
  - Registration differences between the reference and map data
    - Projections, datums, poor quality control
    - Lost--on the ground, image, or map
  - Changes in land cover between map and reference data (i.e., time)
  - Variation in interpretation of reference data
  - Mistakes in labeling the reference data
  - Data entry errors
  - Digitizing errors

Error vs Difference Matrix

Minimize reference label error by...

- Imposing rigorous quality control
- Locating accuracy assessment sites close to polygon centers:
  - Avoids sampling mixed classes near polygon boundaries and
  - Reduces errors due to mis-registration
- Measure variance within sites

Error Matrix

Some Limits of Analysis...

- Assumes each site can only be assigned to one map category
- Doesn't consider the accuracy of polygon boundaries
- Makes no distinction for magnitude of error:
  - Labeling a site as 30-40% crown closure (when its actually 45%) counts as a mistake just as much as labeling a Lake as a Mixed Deciduous Forest

Accuracy Assessment Conclusions

Accuracy Assessment:

- Is expensive
- Won’t improve the map accuracy
- Must be well planned
- Should be statistically valid
- Should be reported as an error matrix
- Is critical to any mapping project that may be used for decision making

Site Specific Accuracy - Summary

- It's more complicated than simply reporting overall accuracy…
  - Consider the relative costs of omission vs commission errors
  - These considerations should influence your choice of analysis methods, thresholds, post processing steps, etc.

If you choose to collect reference data via image interpretation…

- Here are some slides to consider when labeling plots

Image interpretation

- What will we cover?
  1. Basics of image interpretation
  2. Seven image characteristics
  3. Scale and the varying importance of the seven
  4. Land cover types (IPCC, NAMRIA’s, etc.) and their associated spectral signatures

Basic image interpretation

- Image interpretation is the "*determination* of the nature of objects in a photograph or image and the *judgment* of their significance."

"I love cutting trees in perfect circles because it drives aerial photography interpreters crazy."

*This skill is developed through hours of practice with photos, coupled with ground visits to check accuracy of interpretations

Basic image interpretation

- Fundamental to any remote sensing project
- Can reduce costs by up to 35% for mapping, inventorying, and planning for management of forests and rangelands
  - Estimated by the staff of DNR, State of Washington – Edwards, 1972
- Can be used in a variety of applications:

Digitizing

Sampling

Training Data

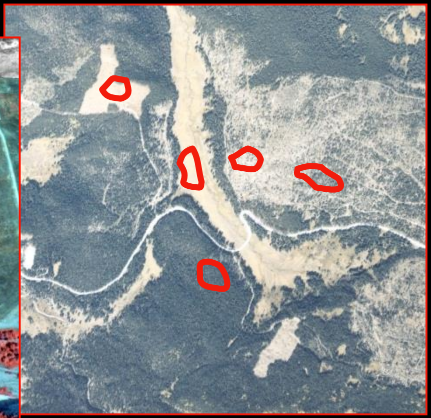


Image Interpretation v. Semi-Automated Classification

- The human brain-eye system is unsurpassed at interpreting high-resolution imagery
- Visual interpretation is the standard for evaluating any classification
- If you don't have the time or resources for a semi-automated method then image interpretation is a fit alternative.

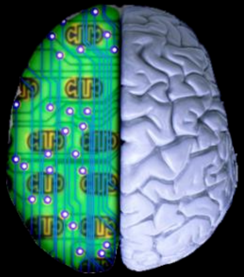
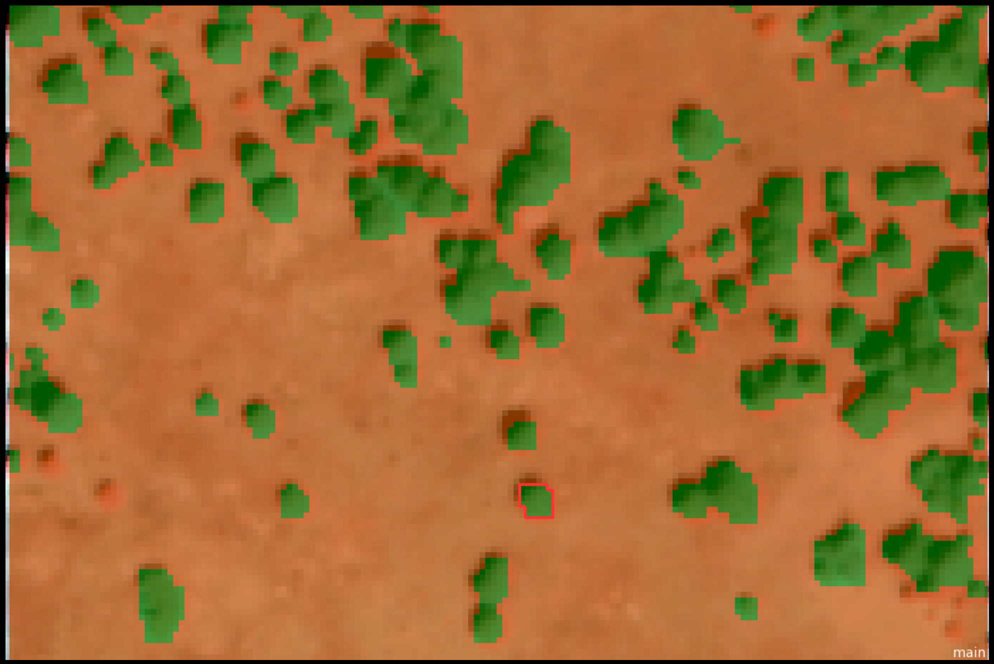
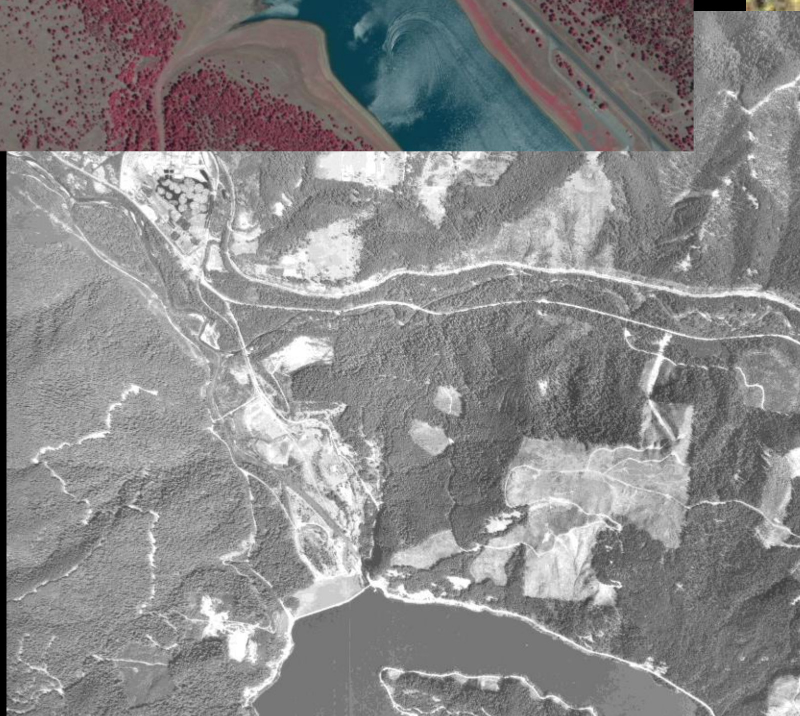


Image interpretation: 7 image cues

- Humans use any combination of seven image characteristics (tone/color, shape, size, association, shadow, pattern, and texture) for visual interpretation
  - Spectral classifiers use only tone/color values
  - Object-oriented classifiers use tone/color plus other cues, such as shape and association

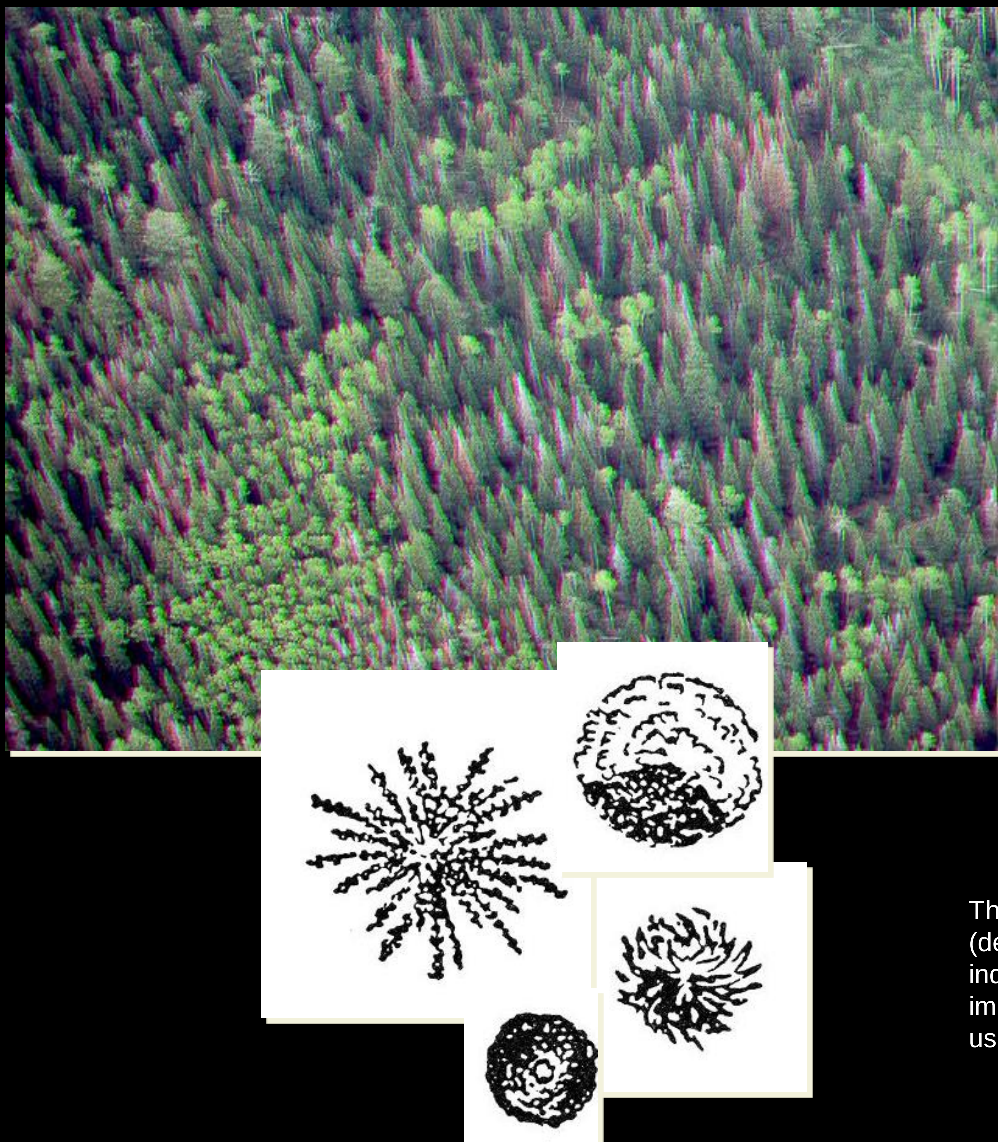


Seven characteristics: Tone/color



Tone (in the B&W image to the left) allows easy distinction between roads, forests, water and bare-ground. Color (in the upper image) allows easy distinction between coniferous trees, deciduous trees, senesced grasses and road surface types. Color (in the upper left image) in this Color Infrared image allows for added distinction between coniferous and deciduous forest, as well as between dead/dying and healthy forest. The brighter red vegetation is indicative of deciduous and the darker (finer textured too) red vegetation is indicative of conifers.

Seven characteristics: Shape



Shape is not necessarily diagnostic for interpretation: the same type of feature may have different shapes and different features may have the same shape.

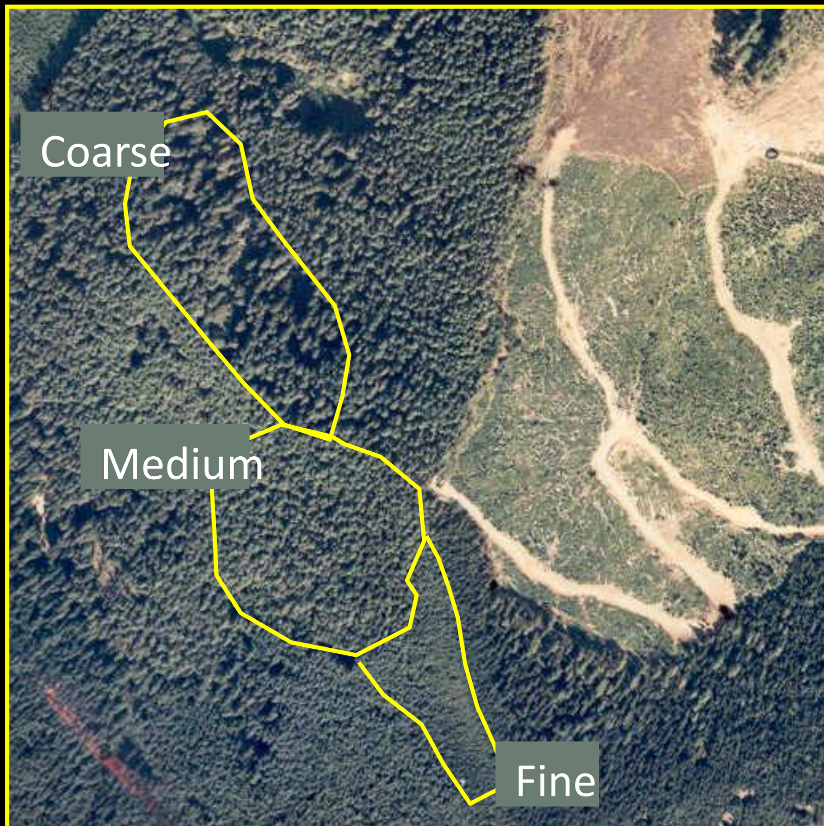
The shape of the tree crowns in the left image is indicative of the type (deciduous/conifer) or even species of tree. The image key graphics indicated the importance of crown shape on large scale photos. In the image above, the shape of the road and the circular tree crowns help us interpret those two features.

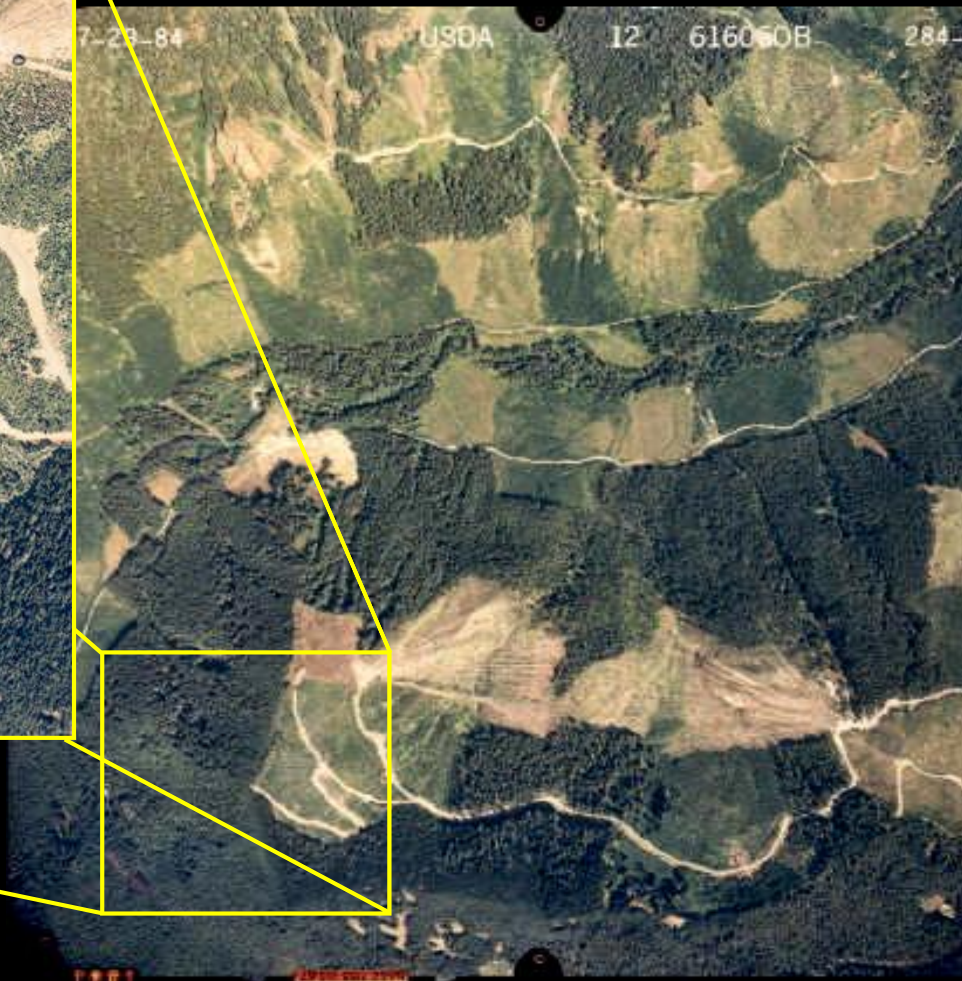
Seven characteristics: Size

Size is usually evaluated by looking at objects that the interpreter is familiar with and comparing the relative size to less familiar objects. At first glance the vegetation in the lower part of this image may look like trees, but in comparison to the size of the vehicles, the vegetation is likely shrubs (e.g. sagebrush).

Seven characteristics: Texture

Texture is the apparent roughness or smoothness of an image region created by the tonal changes of the image.





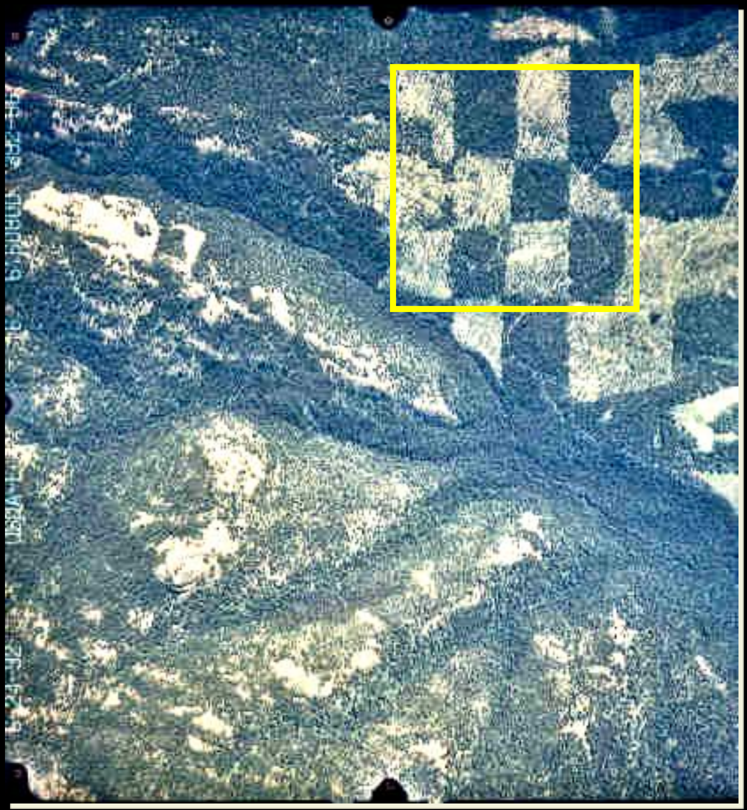
Texture can offer the interpreter clues about the density, age, and type of vegetation present. Note: texture becomes increasingly more important as the scale of the imagery becomes smaller.

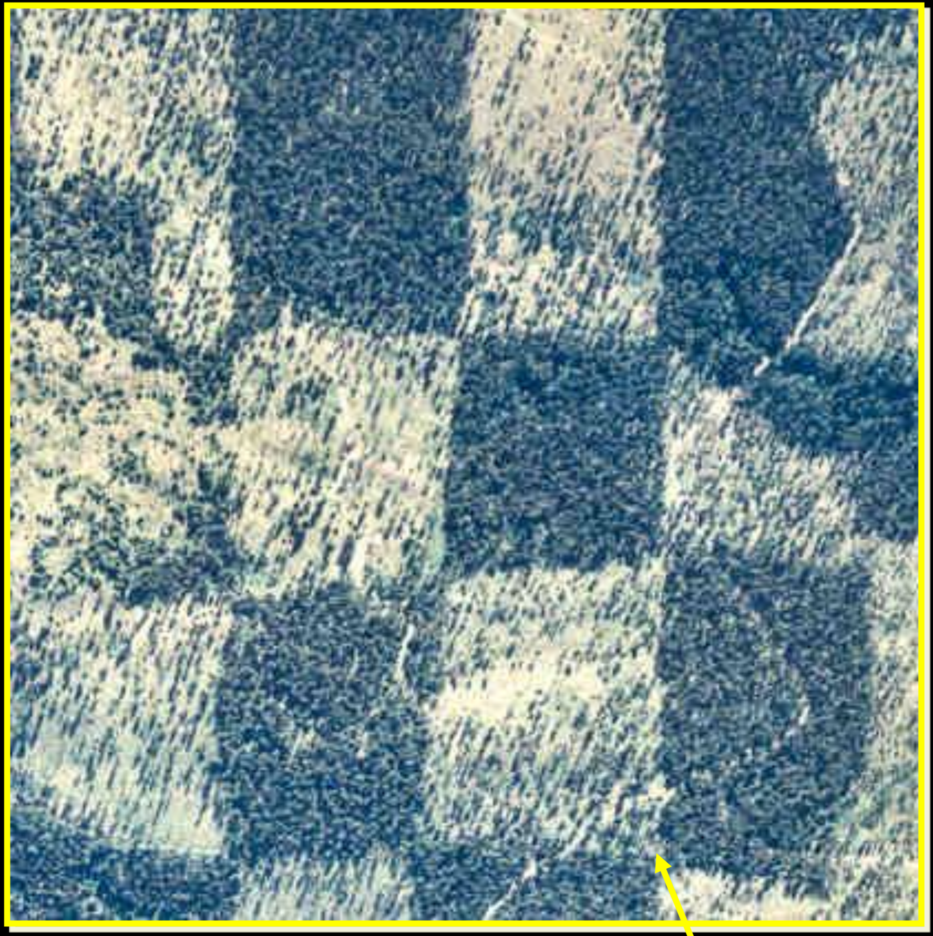
Seven Characteristics: Shadow

Shadows cast by objects in the image can give the interpreter information about the shape and size of certain features in the image.

Seven Characteristics: Pattern

In general, man-made patterns can be easily distinguished. Natural patterns can also lead to information that will aid the interpreter's decision.

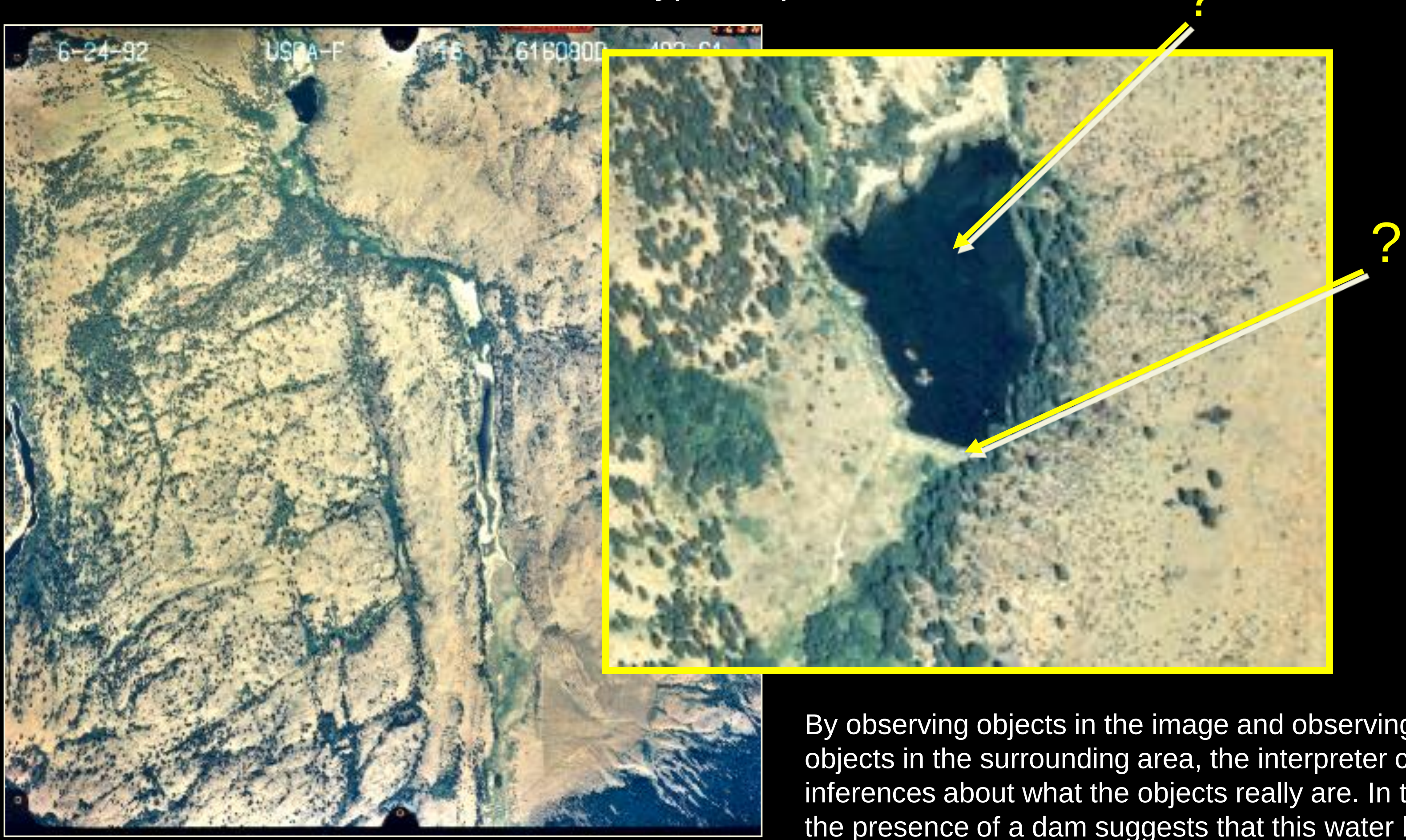




??

Seven Characteristics: Association

In forestry, biological association is particularly important in making vegetation-type maps.



By observing objects in the image and observing the objects in the surrounding area, the interpreter can make inferences about what the objects really are. In this case, the presence of a dam suggests that this water body is a reservoir. Without that association, the water body cannot be further classified as a reservoir.

Convergence of evidence

- Bringing together several types of information so that a conclusion may be drawn
- Inference using multiple (or all) image elements

An interpreter might look at the discrete vegetation dots in this part of the image and interpret them as trees.

United States Department of Agriculture

If you have questions please contact
Bill Silva,
at USFS Geospatial Technology and Applications Center (GTAC):
billsilva@fs.fed.us
801-975-3804

Please contact Sarah Marlay,
at USFS International Programs
to learn more about international training opportunities:
sarahemarlay@fs.fed.us

Forest Service