Object Class Recognition using Images of Abstract Regions

Yi Li, Jeff A. Bilmes, and Linda G. Shapiro
Department of Computer Science and Engineering
Department of Electrical Engineering
University of Washington

Problem Statement

Given: Some images and their corresponding descriptions

{trees, grass, cherry trees}

{cheetah, trunk}

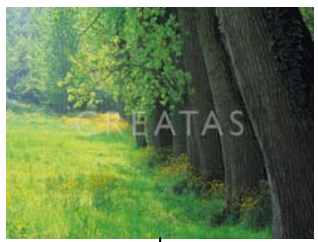
{mountains, sky}

{beach, sky, trees, water}

●●●

To solve: What object classes are present in new images

?



?

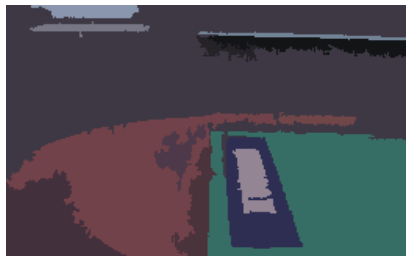
?

?

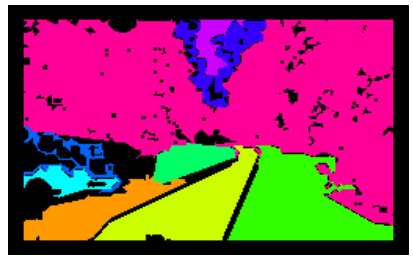
●●●

Image Features for Object Recognition

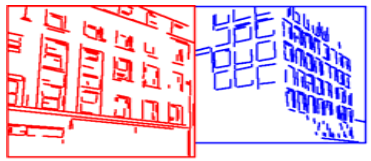
- Color



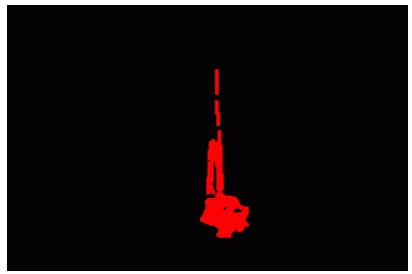
- Texture

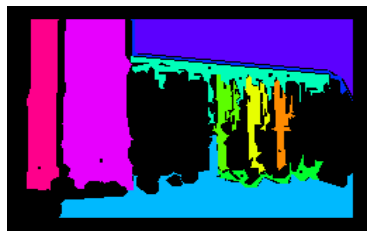
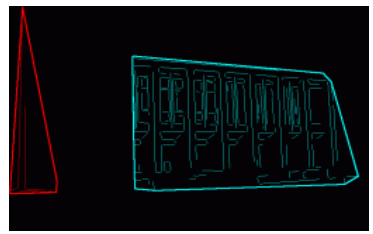
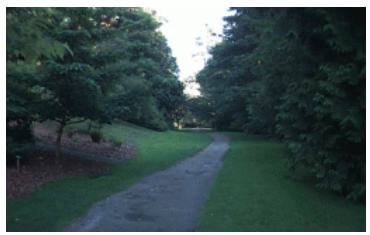
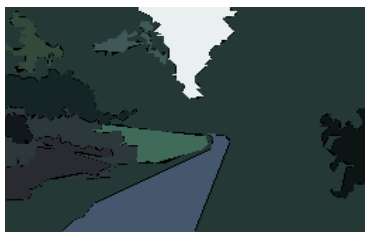
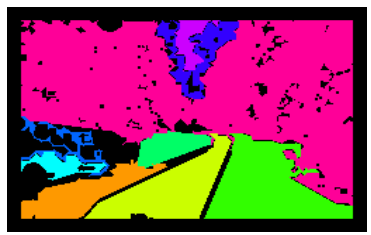
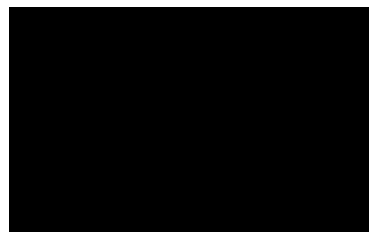
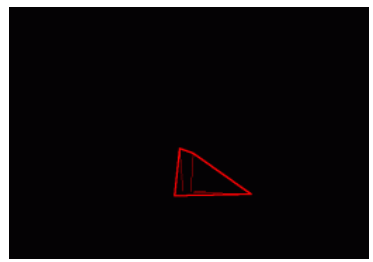
- Structure

- Context

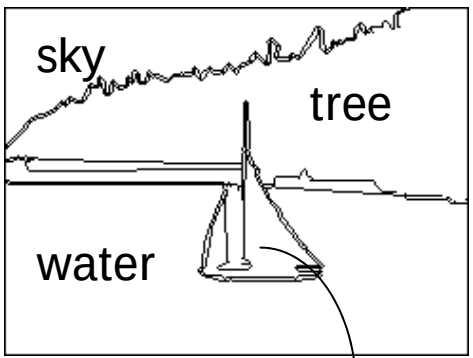



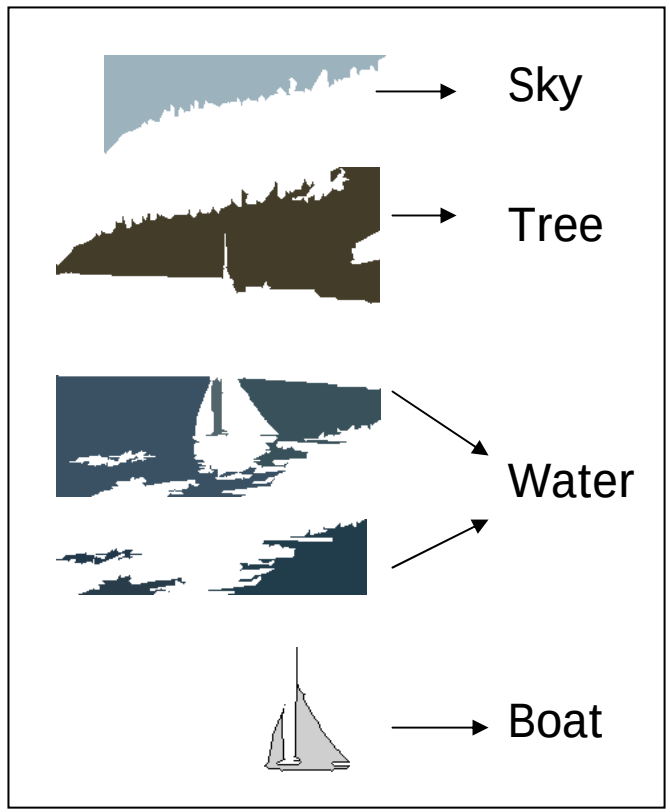
Abstract Regions

Original Images	Color Regions	Texture Regions	Line Clusters
			
			
			

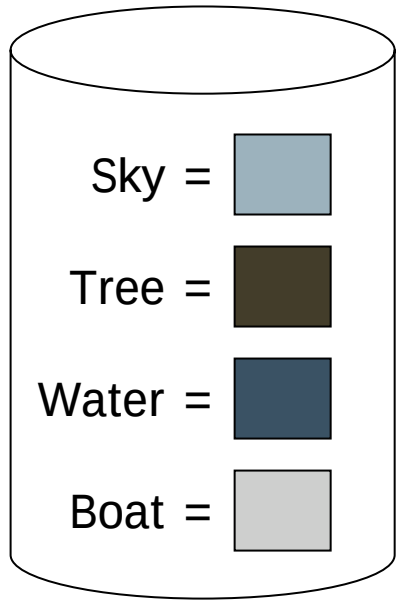
Object Model Learning (Ideal)

+





region attributes → object



Learned Models

Our Scenario: Abstract Regions

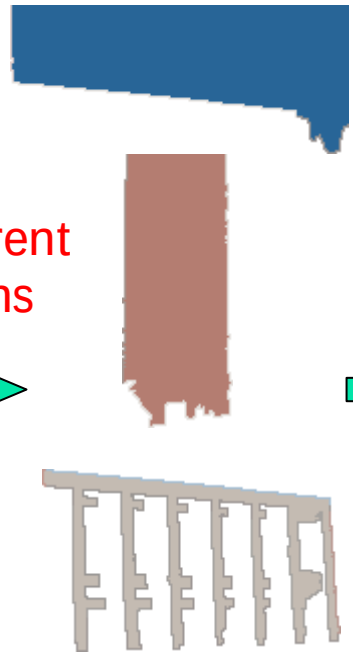
Multiple segmentations whose regions are not labeled; a list of labels is provided for each training image.

image

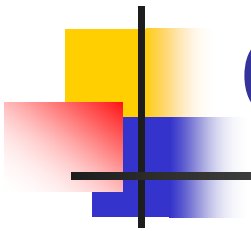
labels

{sky, building}

various different segmentations



region attributes from several different types of regions



Object Model Learning

Assumptions

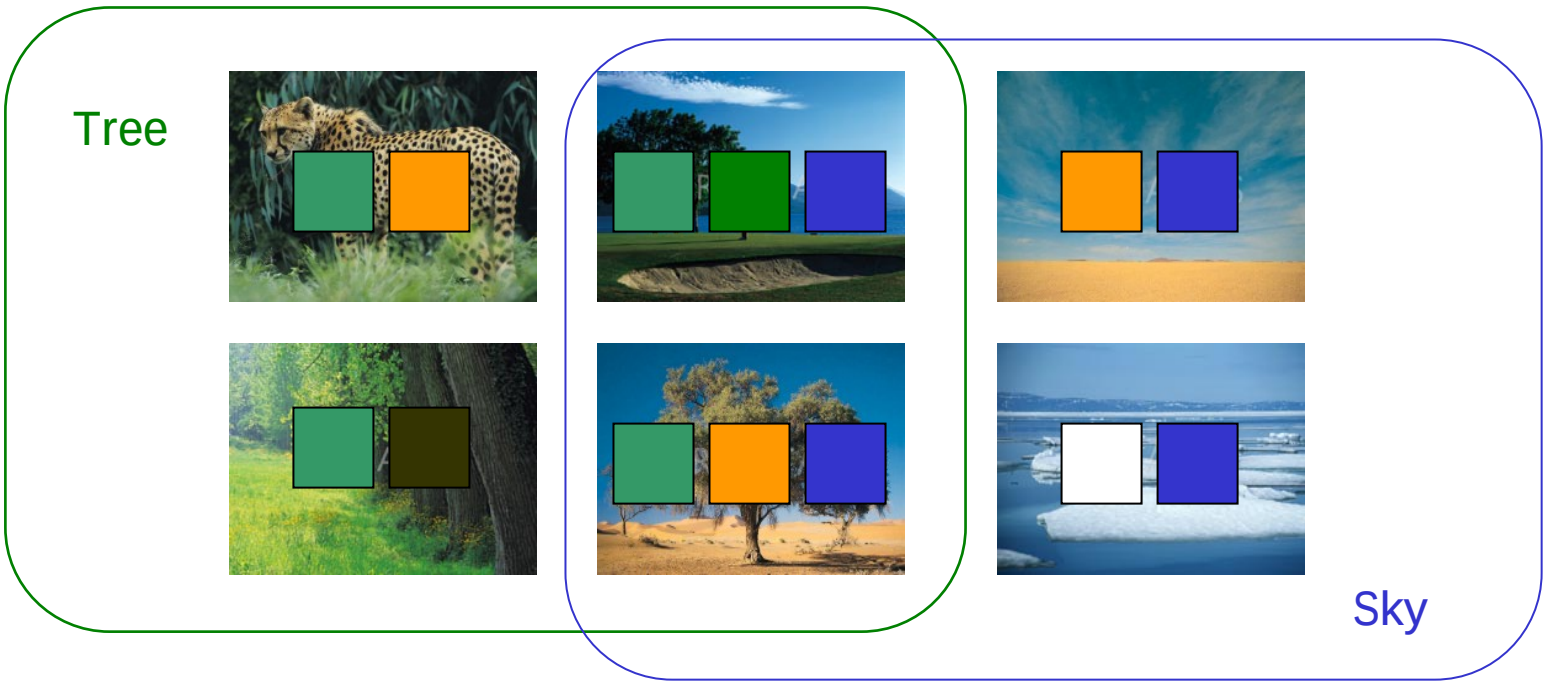
- The feature distribution of each object within a region is a Gaussian;

- Each image is a set of regions;

  each region can be modeled as a mixture of multivariate Gaussian distributions.

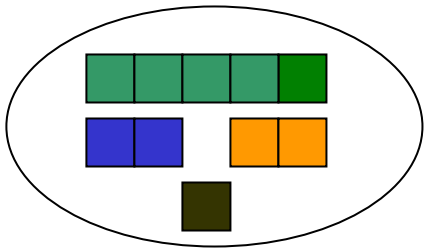
Model Initial Estimation

- Estimate the initial model of an object using all the region features from all images that contain the object

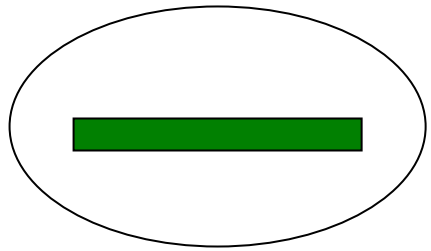


EM Variant

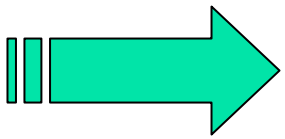
Initial Model for “trees”



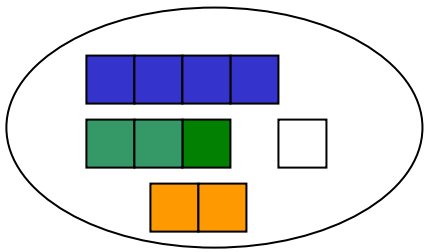
Final Model for “trees”



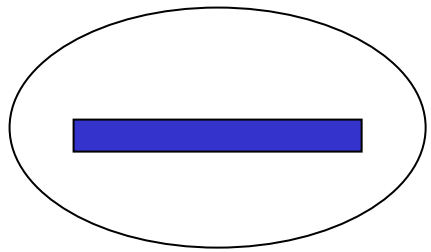
EM

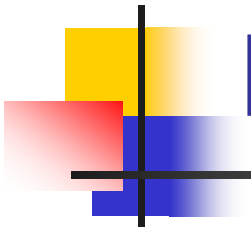


Initial Model for “sky”



Final Model for “sky”



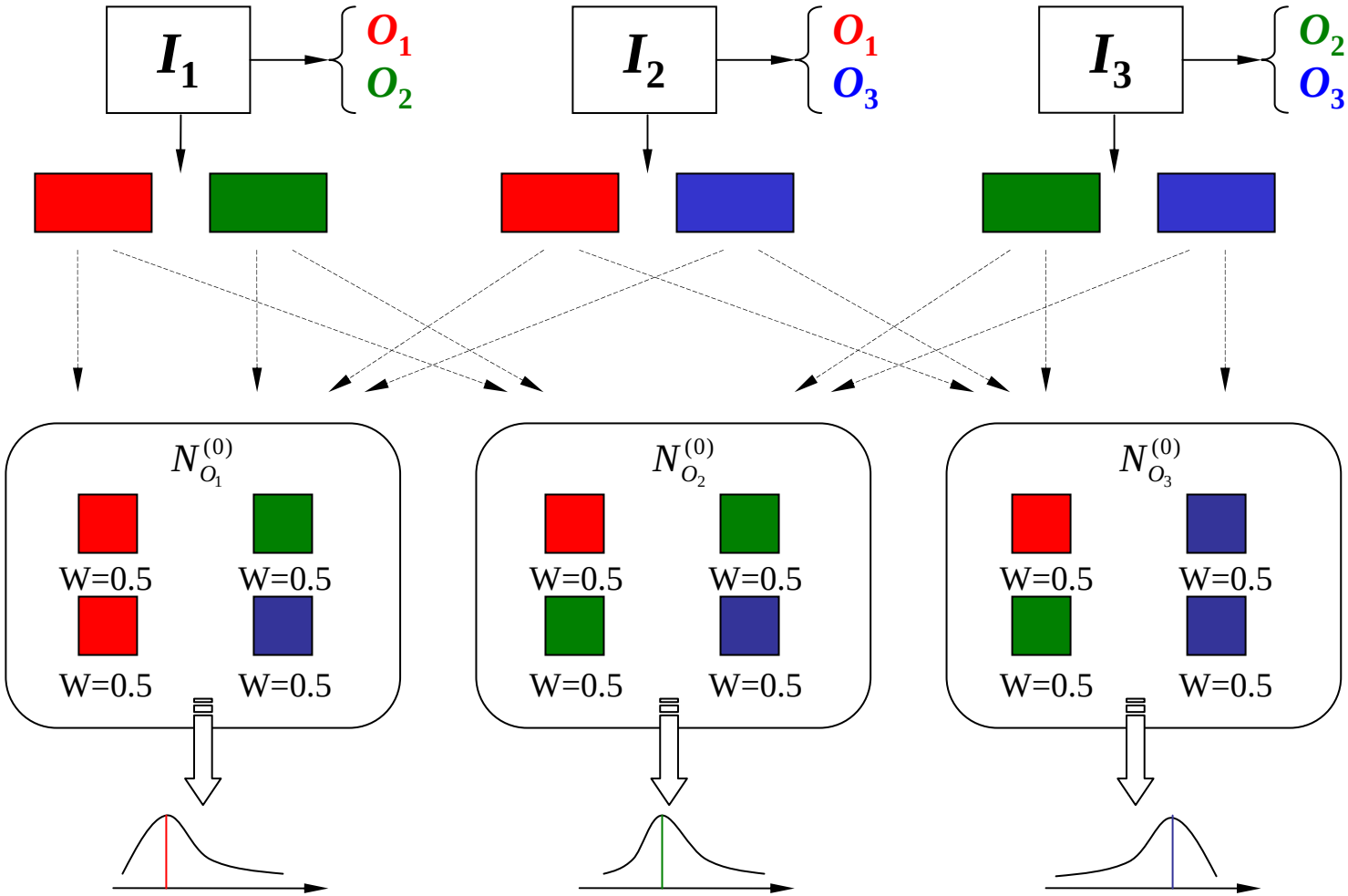


EM Variant

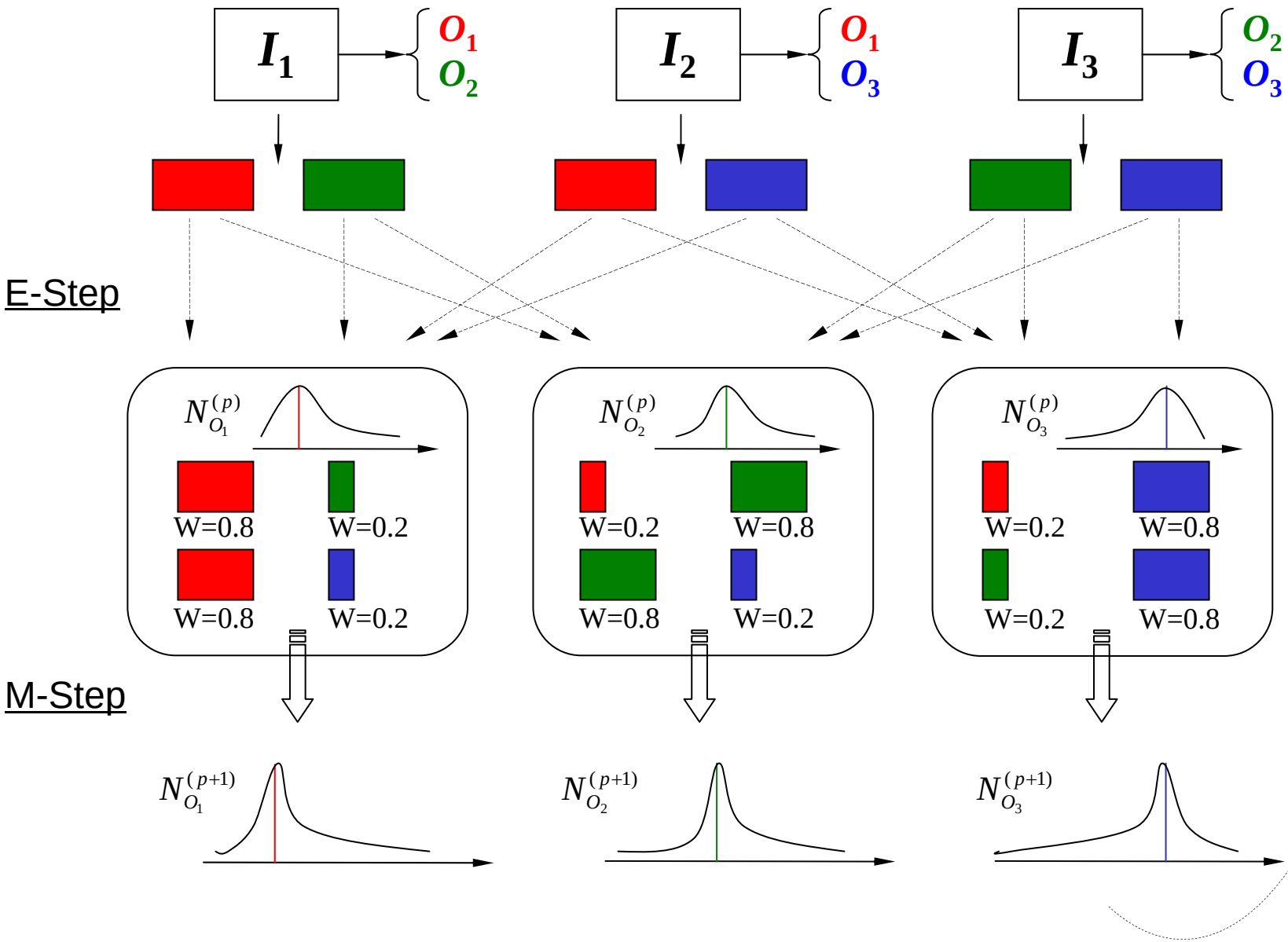
- Fixed components corresponding to the given object labels and fixed component responsibilities corresponding to the frequencies of the corresponding objects in the training data.
- Customized initialization takes advantage of known labels to generates more accurate estimates in the first step.
- Controlled posterior calculation ensures that a feature vector only contributes to the Gaussian components representing objects present in its training image.
- Extra background component absorbs noise.

1. Initialization Step (Example)

Image & description



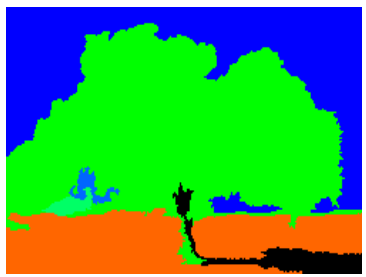
2. Iteration Step (Example)



Recognition

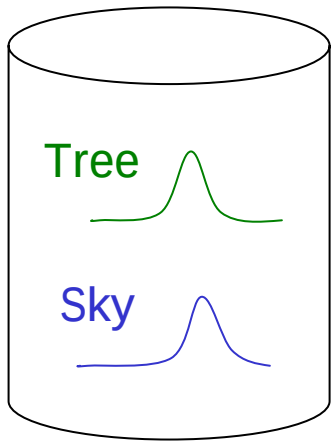
Test Image

Color Regions



compare

Object Model Database



To calculate $p(tree \mid image)$

$$p(tree \mid image) = f\begin{pmatrix} p(\ tree \mid \ ) \\ p(\ tree \mid \ ) \\ p(\ tree \mid \ ) \\ p(\ tree \mid \ ) \end{pmatrix}$$

$$p(o \mid F_I^a) = \underset{r^a \in F_I^a}{f}(p(o \mid r^a))$$

Combining different abstract regions

- Treat the different types of regions independently and combine at the time of classification.

$$p(o \mid \{F_I^a\}) = \prod_a p(o \mid F_I^a)$$

- Form intersections of the different types of regions, creating smaller regions that have both color and texture properties for classification.

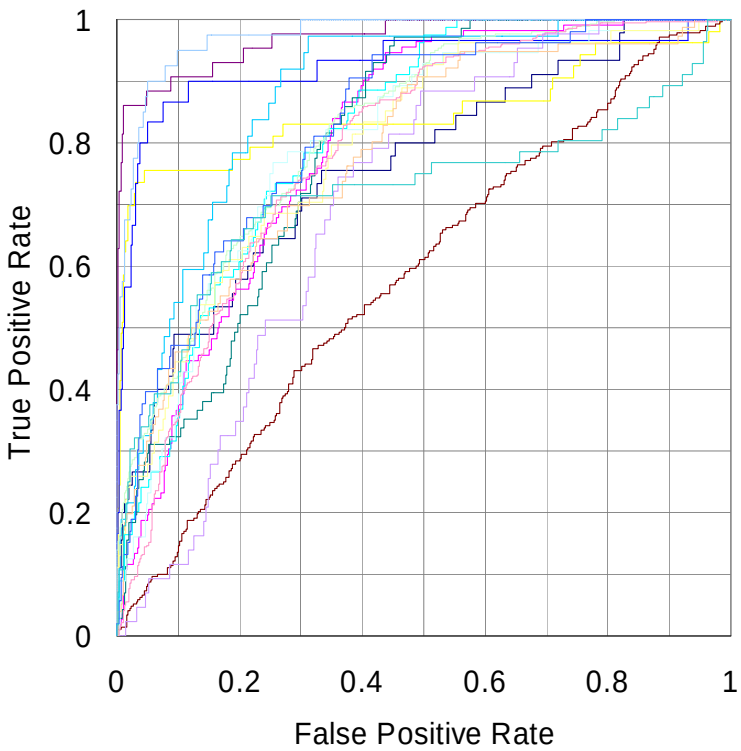
Experiments (on 860 images)

- 18 keywords: mountains (30), orangutan (37), track (40), tree trunk (43), football field (43), beach (45), prairie grass (53), cherry tree (53), snow (54), zebra (56), polar bear (56), lion (71), water (76), chimpanzee (79), cheetah (112), sky (259), grass (272), tree (361).

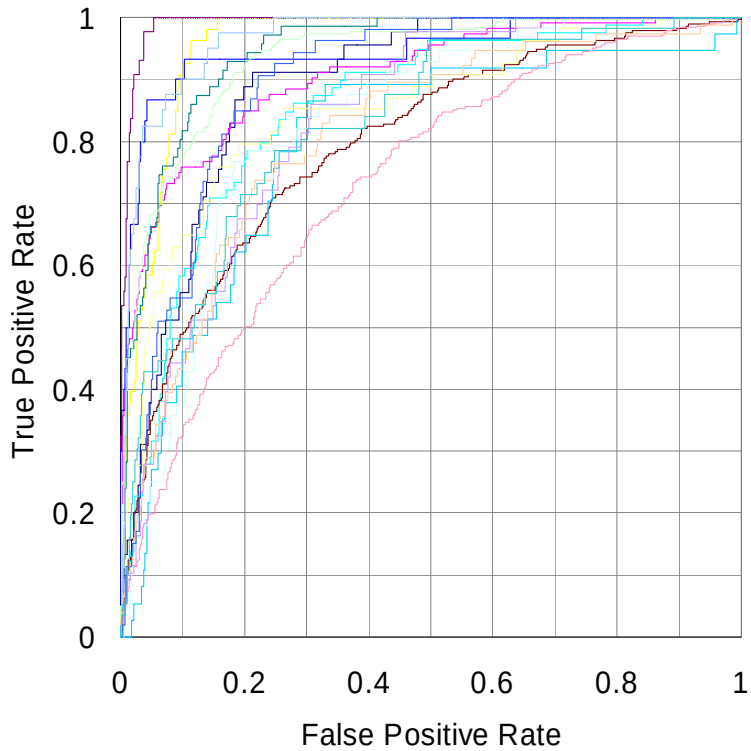
- A set of cross-validation experiments (80% as training set and the other 20% as test set)

- The poorest results are on object classes “tree,” “grass,” and “water,” each of which has a high variance; a single Gaussian model is insufficient.

ROC Charts



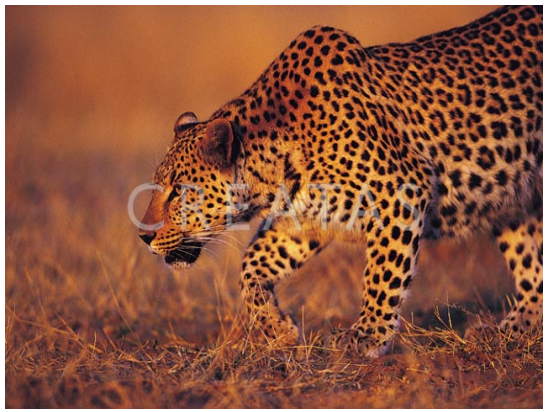
Independent Treatment of Color and Texture

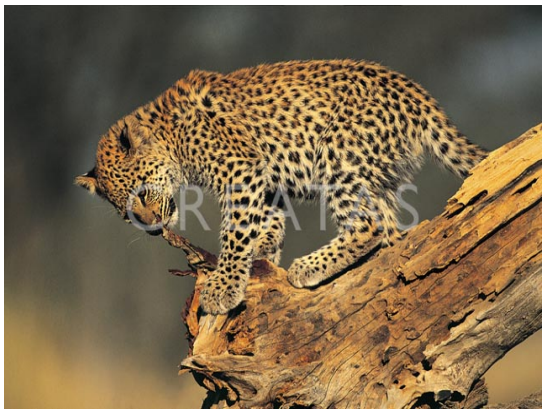


Using Intersections of Color and Texture Regions

Sample Results

cheetah





Sample Results (Cont.)

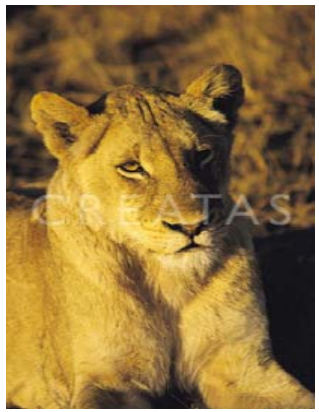
grass

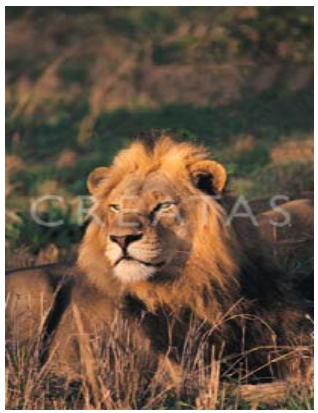
Sample Results (Cont.)

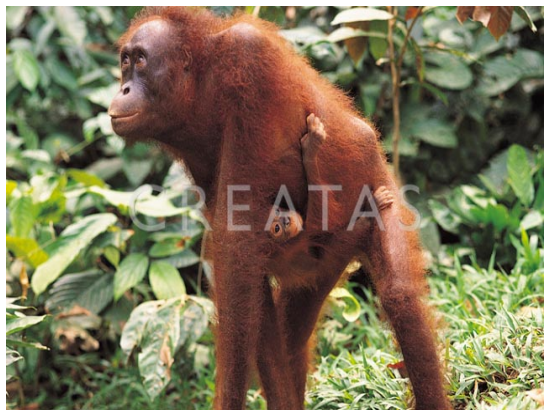
cherry tree

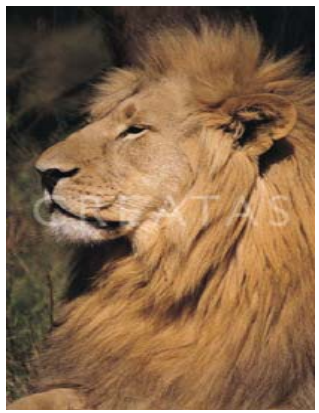
Sample Results (Cont.)

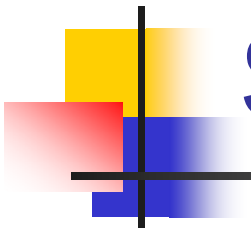
lion





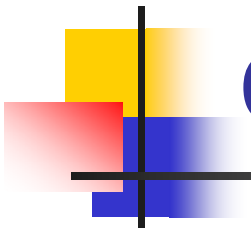






Summary

- Designed a set of abstract region features: color, texture, structure, . . .
- Developed a new semi-supervised EM-like algorithm to recognize object classes in color photographic images of outdoor scenes; tested on 860 images.
- Compared two different methods of combining different types of abstract regions. The intersection method had a higher performance



Current Work on Abstract Regions

- Add more image features
- Investigate other methods for combining different feature types
- Use spatial relationships among regions
- Use Gaussian mixtures for object classes
- Learn spatial configurations of regions