Introduction

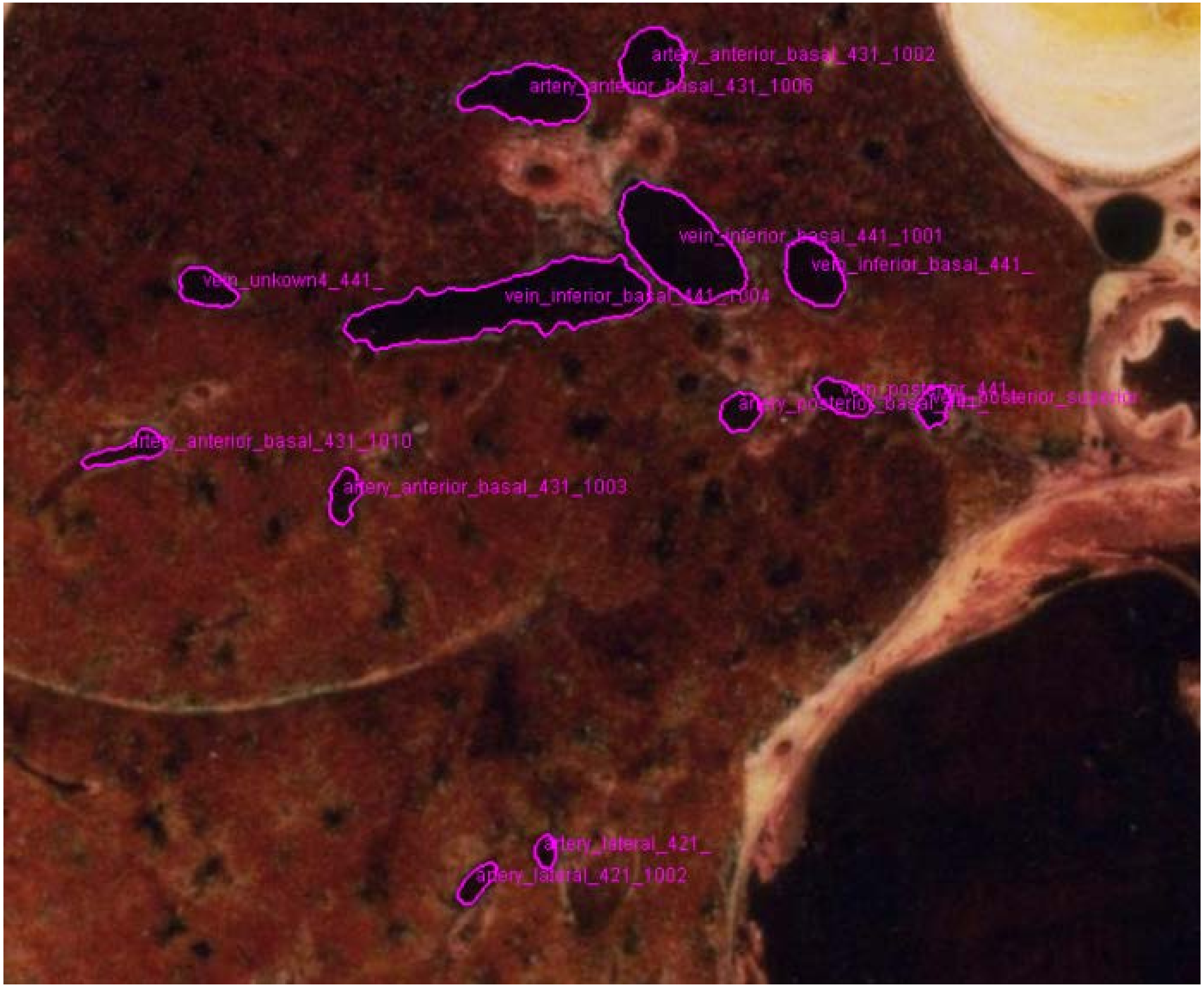
Computer vision is the analysis of digital images by a computer for such applications as:

- Industrial: part localization and inspection, robotics
- Medical: disease classification, screening, planning
- Military: autonomous vehicles, tank recognition
- Intelligence Gathering: face recognition, video analysis
- Security: video analysis
- Science: classification, measurement
- Document Processing: text recognition, diagram conversion

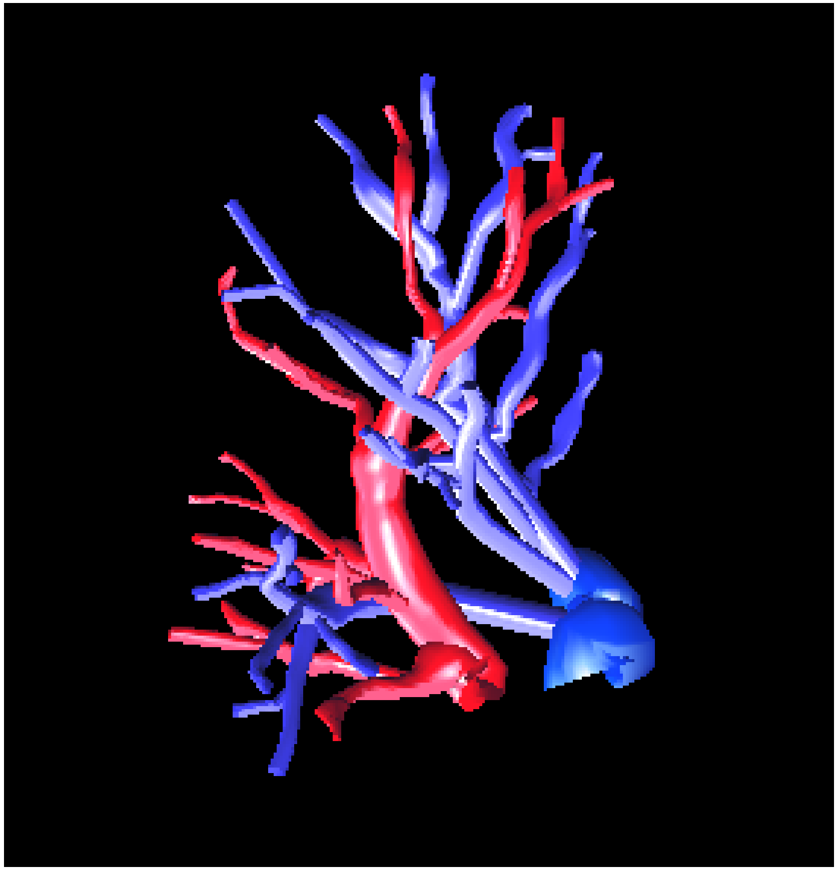
Medical Applications

CT image of a patient's abdomen

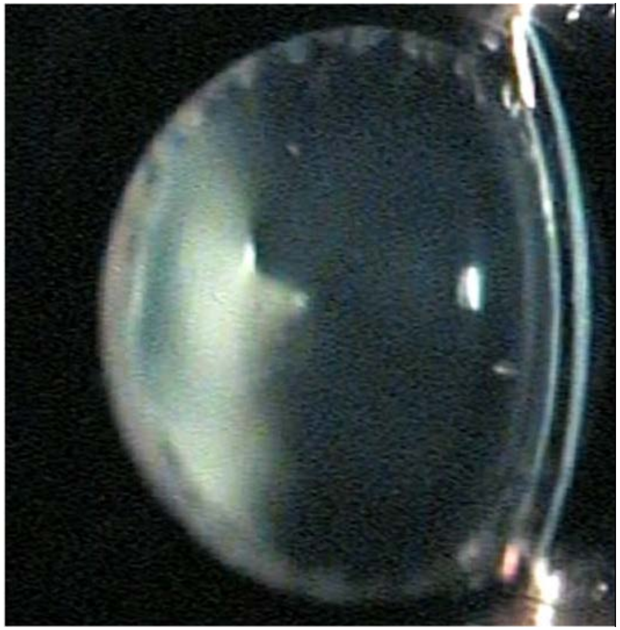
Visible Man Slice Through Lung

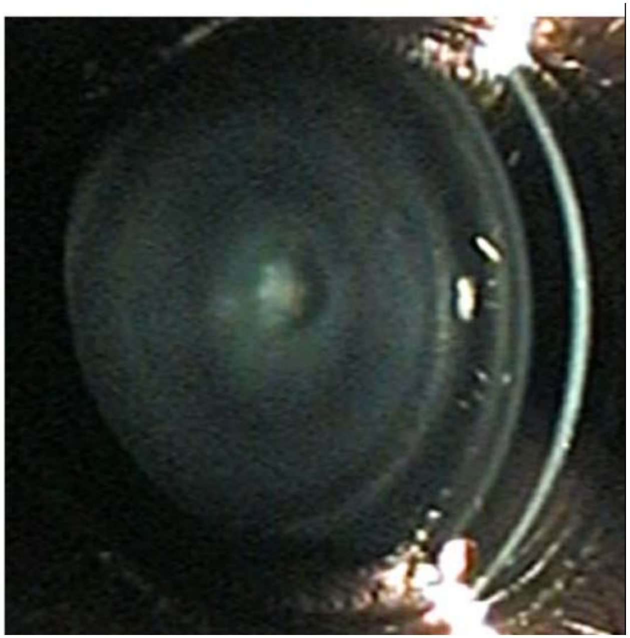


3D Reconstruction of the Blood Vessel Tree



CBIR of Mouse Eye Images for Genetic Studies





Robotics

- 2D Gray-tone or Color Images

"Mars" rover

- 3D Range Images

What am I?

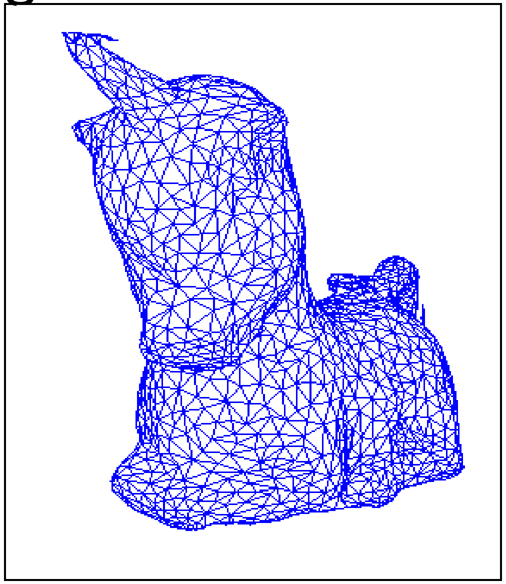


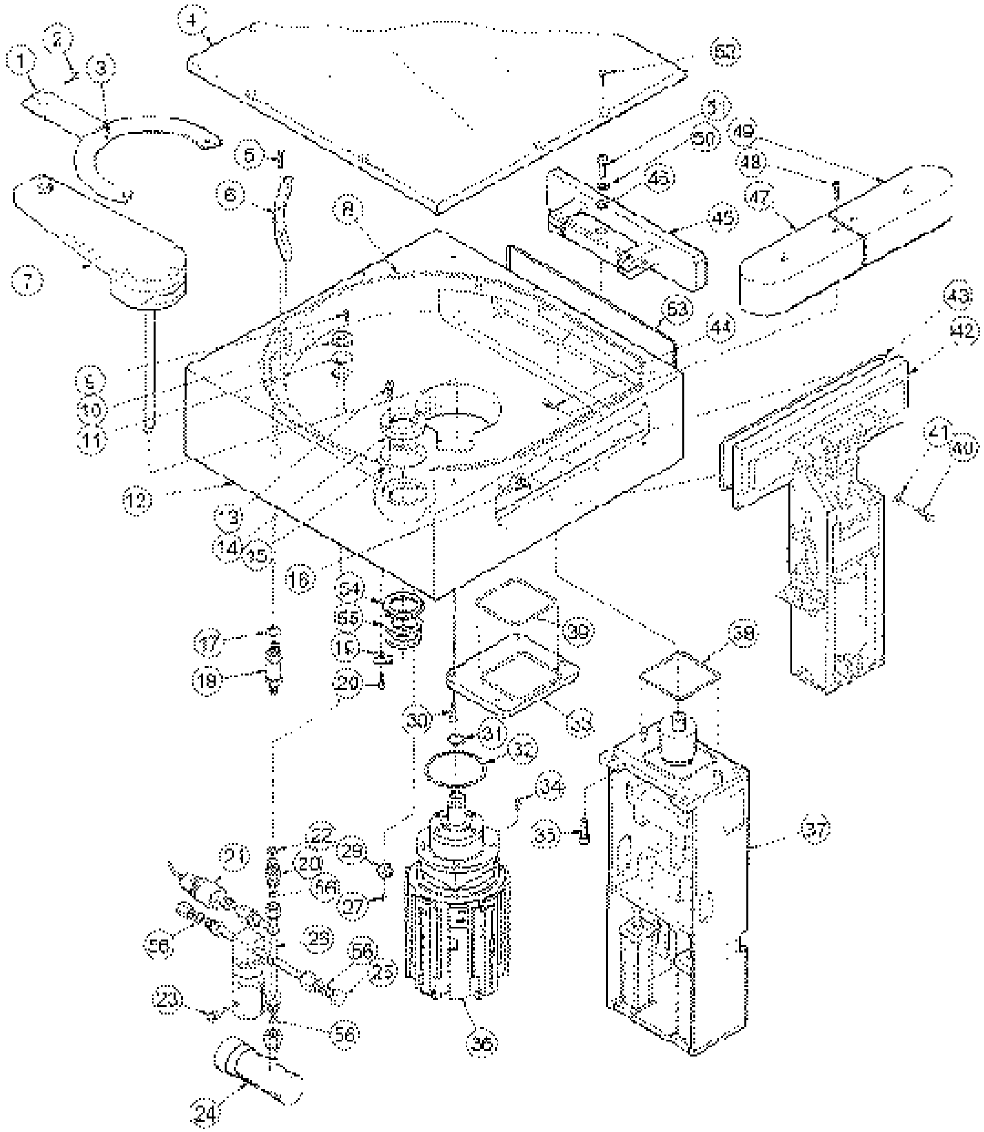
Image Databases:

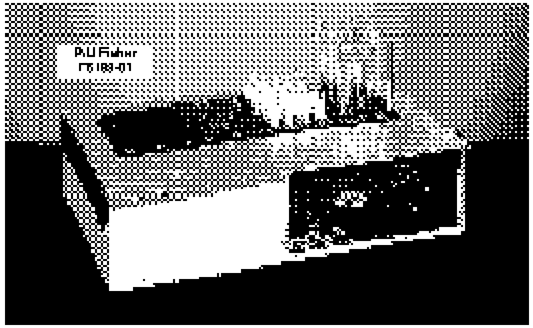
Images from my Ground-Truth collection.

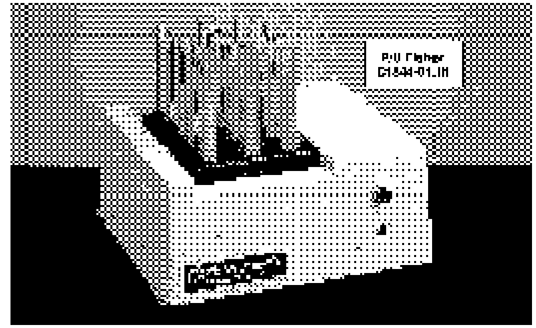


- Retrieve all images that have trees.
- Retrieve all images that have buildings.
- Retrieve all images that have antelope.

Documents:







Model 145 Isotemp® Dry Bath Incubator

- *Holds 1 to 4 heating blocks with choice of 11 well sizes*
- *Maintains every sample to within ±0.1° C of temperature*

Unique sample wells are shaped so that aluminum blocks deliver same amount of heat to all parts of the sample tube. No temperature gradient—neither too hot at the bottom nor too cold at the top—that may invalidate tests in racks with drilled cylindrical wells. Sample tubes rest on insulating plugs to prevent localized heating. A low watt-density heater is mounted on a thick aluminum heat distribution plate in the floor of the bath. Plate is 3/8" thick (9.5 mm). Dry bath minimizes cleanup problems because tubes stay dry.

Ambient to 125° C (257° F) with ±0.1° C control. Dial temperature control graduated from 25° to 125° C. Ideal for enzyme reactions, inactivation of sera, Rh studies, blood crossmatching and rhogam determinations. Dimensions: 9 L x 15 1/4 W x 4" H (23 x 39 x 10 cm). With line cord and plug. **Heating blocks sold separately** (see lower right).

Electrical Requirements	Cat. No.	Each
120V 50/60 Hz, 300W (CSA approved)	11-715-100	[illegible]
240V 50/60 Hz, 600W	11-715-101	[illegible]

[illegible] 37° C
[illegible] Model

Model 147 Isotemp® Dry Bath

- *Holds single heating block with choice of 11 well sizes*

Similar to Model 145, but with 3/8" thick (9.5 mm) plate. Ideal for labs with smaller volumes of enzyme and cholesterol assays, Rh studies, and dry incubators. Screwdriver-adjusted temperature control between ambient and 95° C (203° F). Observe thermometer placed in unused sample tube to set; adjust control through hole in front panel. Maintains set temperature with consistency and uniformity of ±0.5° C.

Supplied with strong nylon case, thermostatically controlled heater, heat indicator lamp, line cord and plug, and instructions. Dimensions: 8 L x 6 3/4 W x 3" H (15 x 17 x 8 cm). CSA approved. **Heating blocks sold separately** (see below).

Electrical Requirements	Cat. No.	Each
120V 50/60 Hz, 120W	11-715-102	223.50

Interchangeable Heating Blocks for Isotemp® Dry Baths

For Models 145 and 147 Dry Baths. Composed of black anodized aluminum alloy. Chemical resistant. Dimensions: 4 L x 3 W x 2" H (10 x 7.5 x 5 cm).

The 11-715-123 block provides a safe, dry bath alternative for warming the paraffin of tissue processors. Avoids hazardous use of burners around flammable histological solvents.

The 11-715-120 block is specifically designed to hold twenty 9.5 mm Roche Diagnostics Placentex® pregnancy test tubes. This special shallow well block is similar to the other blocks with 10 mm holes, but sample wells are only 1/2" deep (1.3 cm) to meet test requirements. Wells in all other blocks are 1 3/4" deep (4.4 cm).

Tube Size, mm	Wells/Block	Cat. No.	Each
6	35	11-715-105	71.[illegible]
10	20	11-715-107	71.[illegible]
9.5	20 (shallow)	11-715-120	71.[illegible]
12	12	11-715-109	71.[illegible]
[illegible]	12	11-715-121	71.[illegible]
13	12	11-715-111	71.[illegible]
15	12	11-715-113	71.[illegible]
16	9	11-715-123	71.[illegible]
18	12	11-715-115	71.[illegible]
20	6	11-715-117	71.[illegible]
25	5	11-715-119	71.[illegible]

[illegible]
[illegible]

Incu-Block® Partial Immersion Thermometers

For all standard heating blocks and water baths. Critical temperatures (25°, 30°, 37°, 56° C) are marked with arrows. Available with shatter-proof, contamination-proof Teflon® coating. Total length: 175 mm. Immersion: 35 mm.

Range, °C	Div., °C	Teflon-Coated	Cat. No.	Each
25-57	0.5	No	14-992	45.84
25-57	0.5	Yes	14-993	45.15

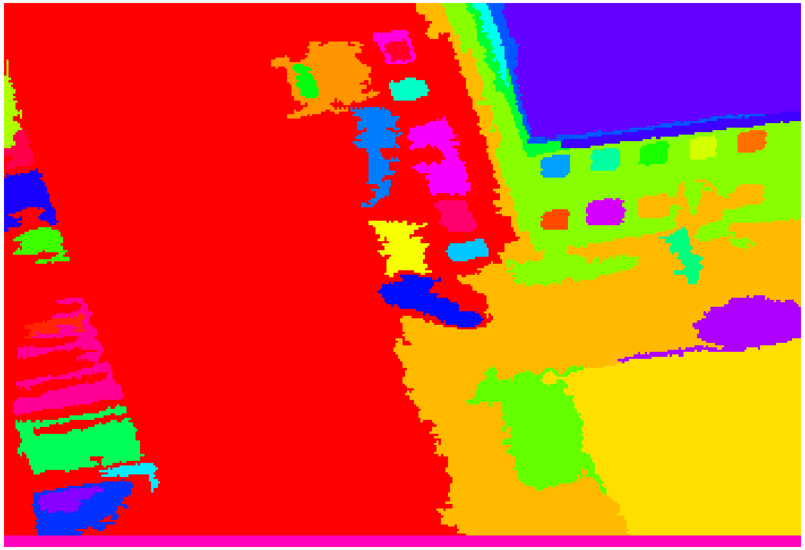
More Thermometers

For more thermometers, including digital types,

see page 952

Surveillance: Object and Event Recognition in Aerial Videos

Original Video Frame



Color Regions

Structure Regions

Digital Image Terminology:

0	0	0	0	1	0	0
0	0	1	1	1	0	0
0	1	95	96	94	93	92
0	0	92	93	93	92	92
0	0	93	93	94	92	93
0	1	92	93	93	93	93
0	0	94	95	95	96	95

pixel (with value 94)

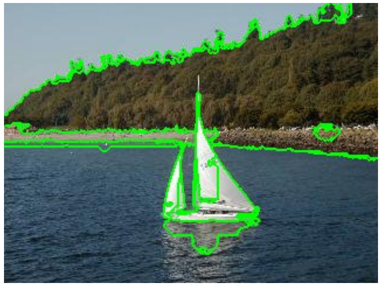
its 3x3 neighborhood

region of medium intensity

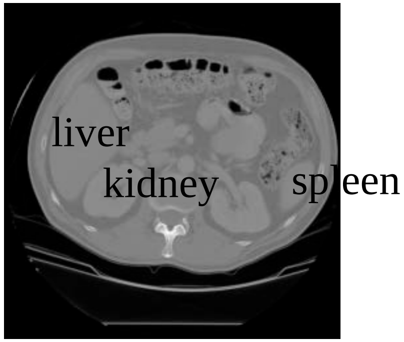
resolution (7x7)

- binary image – 0's and 1's
- gray-scale (or gray-tone) image – 0 to 255
- color image – (R,G,B) at each pixel
- multi-spectral image – multiple values per pixel
- range image – depth value at each pixel
- labeled image – result of processing and labeling

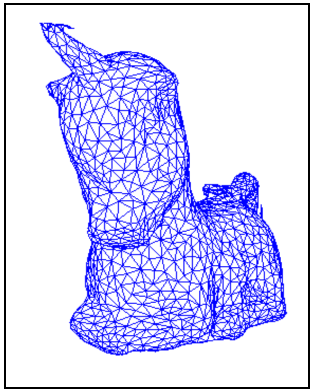
Goals of Image and Video Analysis



- Segment an image into useful regions
- Perform measurements on certain areas



- Determine what object(s) are in the scene
- Calculate the precise location(s) of objects
- Visually inspect a manufactured object



- Construct a 3D model of the imaged object
- Find "interesting" events in a video

The Three Stages of Computer Vision

- low-level

  image → image

- mid-level

  image → features

- high-level (the intelligent part)

  features → analysis

Low-Level

sharpening

blurring

Low-Level

original image

Canny edge operator →

edge image

Mid-Level (Lines and Curves)

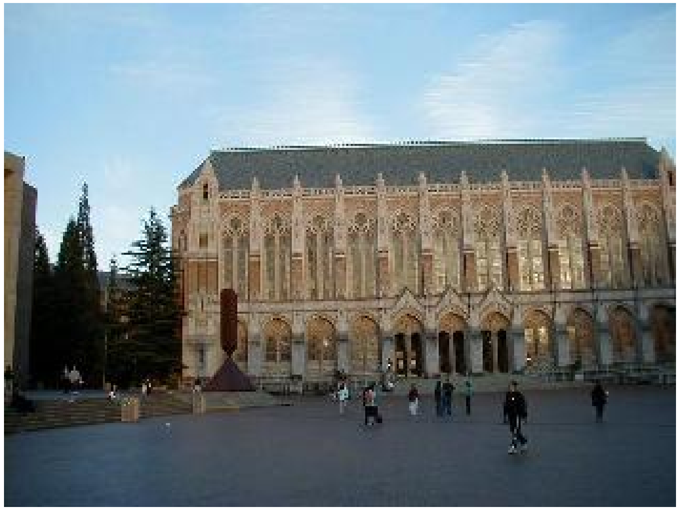
edge image

ORT line & circle extraction →

data structure

circular arcs and line segments

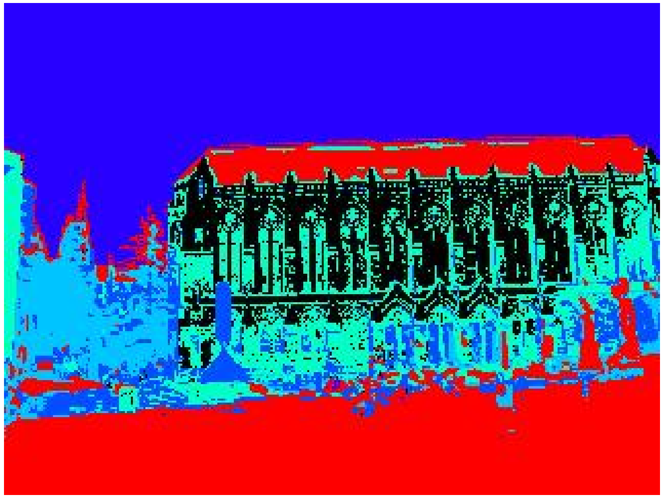
Mid-level (Regions)



original color image

K-means clustering

(followed by connected component analysis)



regions of homogeneous color

data structure

Low- to High-Level

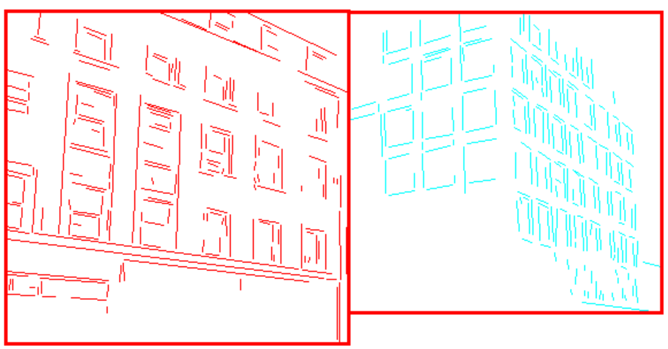
low-level

edge image

mid-level

consistent line clusters

high-level



Building Recognition

Filtering Operations Use Masks

- Masks operate on a neighborhood of pixels.
- A mask of coefficients is centered on a pixel.
- The mask coefficients are multiplied by the pixel values in its neighborhood and the products are summed.
- The result (response) goes into the corresponding pixel position in the output image.

36	36	36	36	36
36	36	45	45	45
36	45	45	45	54
36	45	54	54	54
45	45	54	54	54

Input Image

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

3x3 Mask (mean filter)

**	**	**	**	**
**	39	**	**	**
**	**	**	**	**
**	**	**	**	**
**	**	**	**	**

Output Image

Comparison: salt and pepper noise

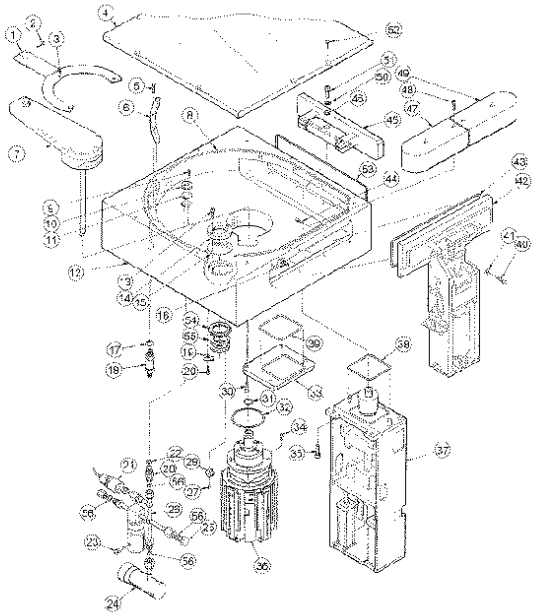
Comparison: Gaussian noise

Lines and Arcs Segmentation

In some image sets, lines, curves, and circular arcs are more useful than regions or helpful in addition to regions.

Lines and arcs are often used in

- object recognition
- stereo matching
- document analysis



Edge Detection

Basic idea: look for a neighborhood with strong signs of change.

Problems:

- neighborhood size
- how to detect change

81	82	26	24
82	33	25	25
81	82	26	24

Differential Operators

Differential operators

- attempt to approximate the gradient at a pixel via masks
- threshold the gradient to select the edge pixels

Example: Sobel Operator

$$Sx = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \qquad Sy = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

On a pixel of the image I

- let gx be the response to Sx
- let gy be the response to Sy

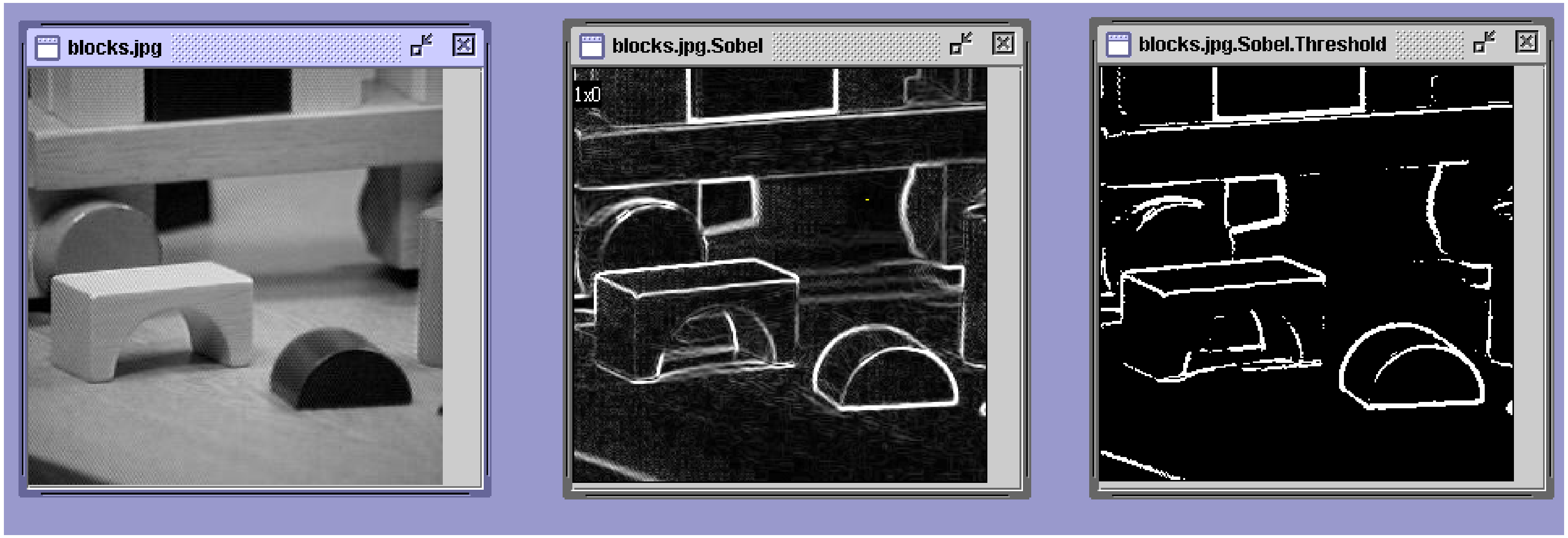
Then the gradient is

$\nabla I = [gx \;\; gy]^T$

and $g = (gx^2 + gy^2)^{1/2}$ is the gradient magnitude.

$\theta = \text{atan2}(gy,gx)$ is the gradient direction.

Sobel Operator on the Blocks Image



original image

gradient magnitude

thresholded gradient magnitude

Common Masks for Computing Gradient

- Sobel:

-1	0	1
-2	0	2
-1	0	1

1	2	1
0	0	0
-1	-2	-1

- Prewitt:

-1	0	1
-1	0	1
-1	0	1

1	1	1
0	0	0
-1	-1	-1

- Roberts

0	1
-1	0

1	0
0	-1

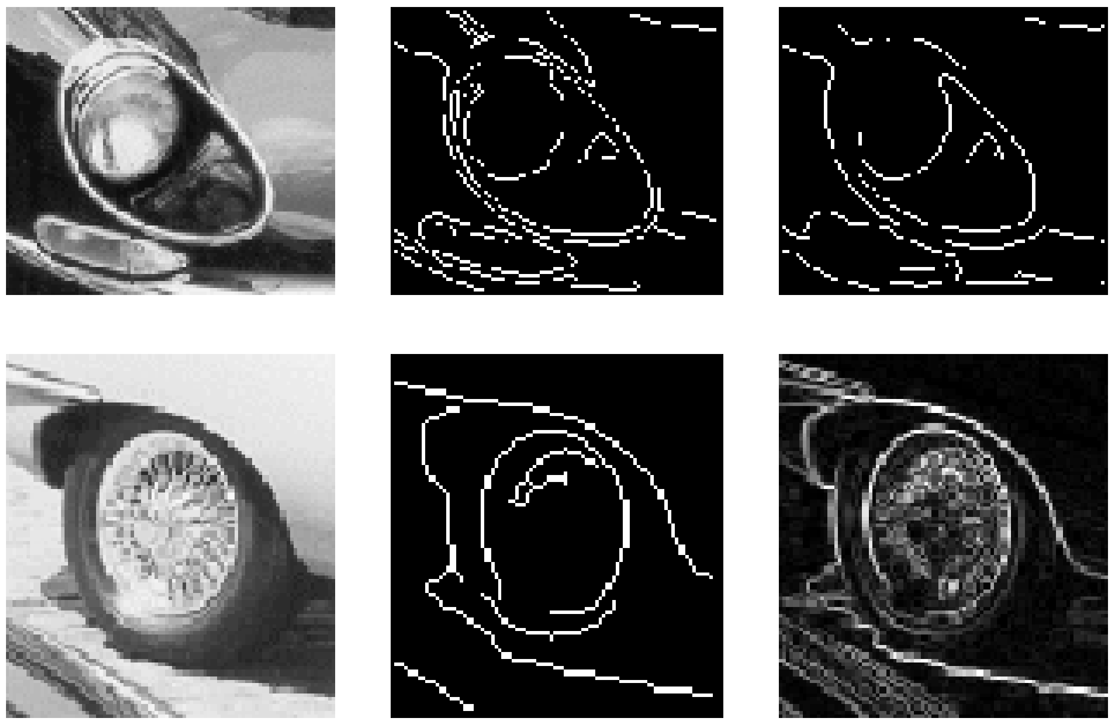
Sx **Sy**

Canny Edge Detector

- Smooth the image with a Gaussian filter with spread $\boldsymbol{\sigma}$**.**
- Compute gradient magnitude and direction at each pixel of the smoothed image.
- Zero out any pixel response $\leq$ the two neighboring pixels on either side of it, along the direction of the gradient.
- Track high-magnitude contours.
- Keep only pixels along these contours, so weak little segments go away.

Canny Examples

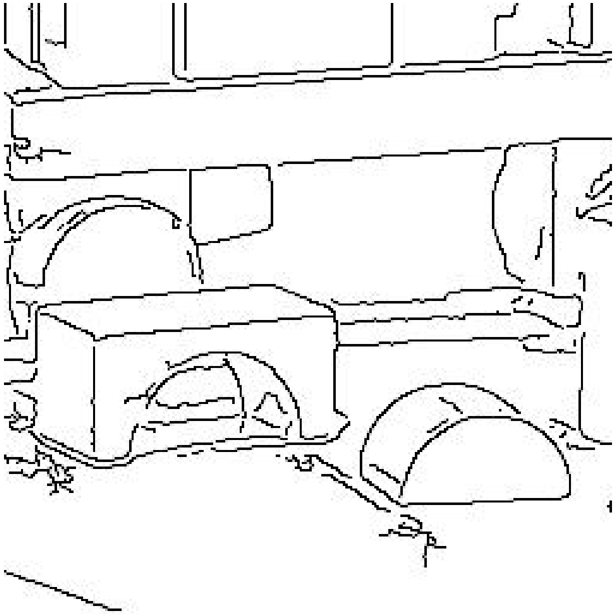
Canny σ=1 Canny σ=4



Canny σ=1 Roberts 2X2

Canny on Kidney Image

Canny on the Blocks image



Canny Characteristics

- The Canny operator gives single-pixel-wide images with good continuation between adjacent pixels

- It is the most widely used edge operator today; no one has done better since it came out in the late 80s. Many implementations are available.

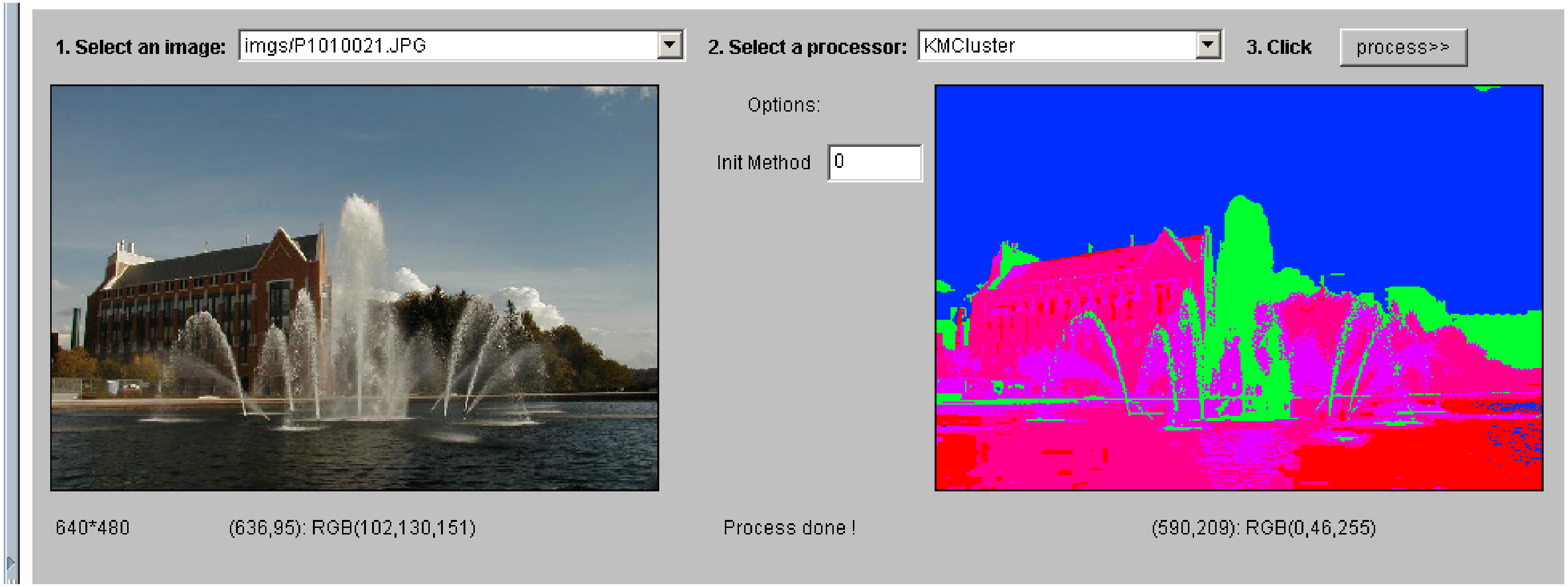
- It is very sensitive to its parameters, which need to be adjusted for different application domains.

Segmentation into Regions

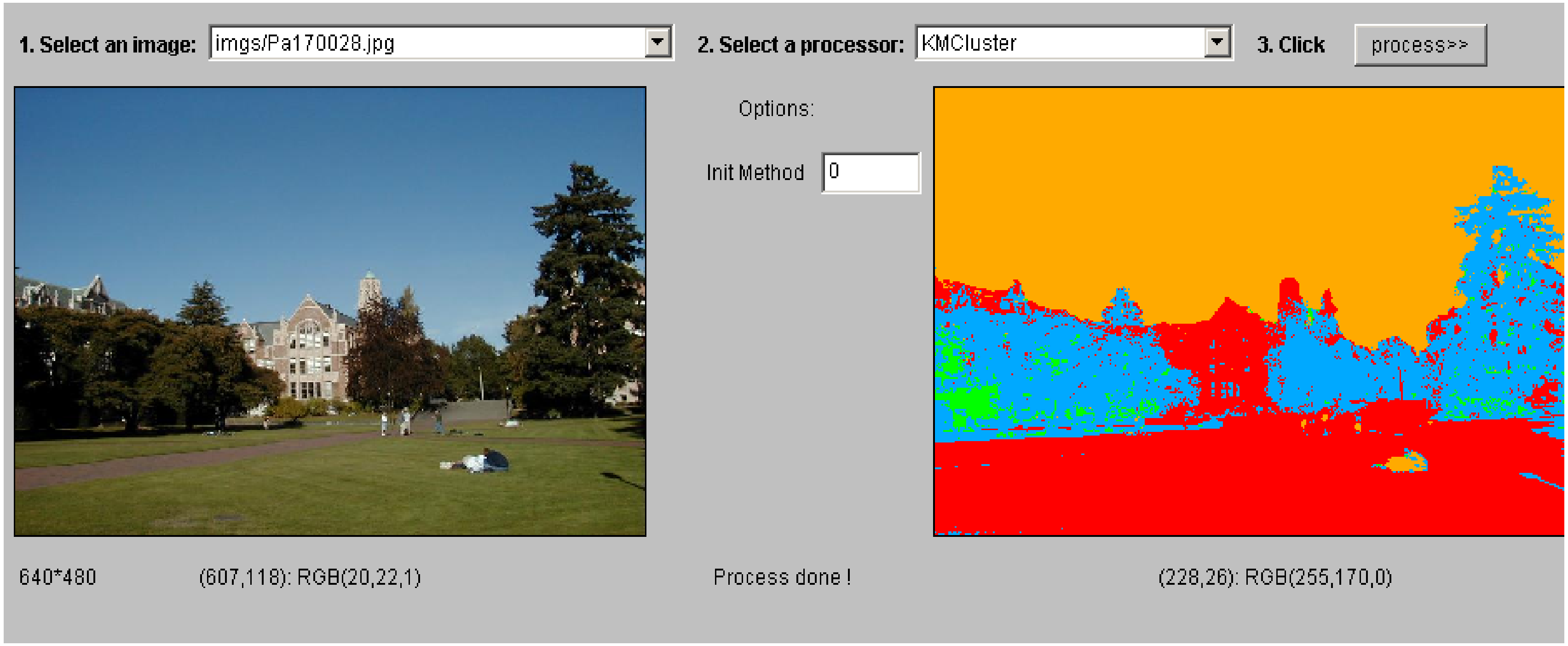
- Instead of looking for 1D features like lines and curves, some processes look for regions.
- The regions must be homogeneous in some attribute such as gray-tone, color, texture,...
- Although “region-growing” was popular in the past, clustering the pixels into subsets has become the best methodology for finding regions.

K-Means Example 1

K-Means Example 2



K-Means Example 3



High-Level Computer Vision

- Detection of classes of objects (faces, motorbikes, trees, cheetahs) in images

- Recognition of specific objects such as George Bush or machine part #45732

- Classification of images or parts of images for medical or scientific applications

- Recognition of events in surveillance videos

- Measurement of distances for robotics

High-level vision uses techniques from AI

- Graph-Matching: A*, Constraint Satisfaction, Branch and Bound Search, Simulated Annealing
- Learning Methodologies: Decision Trees, Neural Nets, SVMs, EM Classifier
- Probabilistic Reasoning, Belief Propagation, Graphical Models

Graph Matching for Object Recognition

- For each specific object, we have a geometric model.
- The geometric model leads to a symbolic model in terms of image features and their spatial relationships.
- An image is represented by all of its features and their spatial relationships.
- This leads to a graph matching problem.

Model-based Recognition as Graph Matching

- Let U = the set of model features.
- Let R be a relation expressing their spatial relationships.
- Let L = the set of image features.
- Let S be a relation expressing their spatial relationships.
- The ideal solution would be a subgraph isomorphism f: U-> L satisfying
- if $(u_1, u_2, ..., u_n) \in R$, then $(f(u_1),f(u_2),...,f(u_n)) \in S$

House Example

2D model

2D image

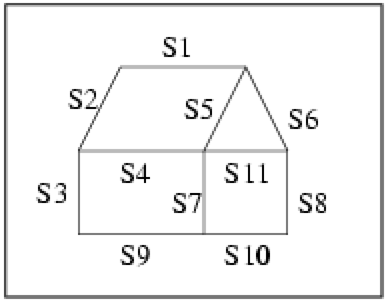


Image 1 **P**

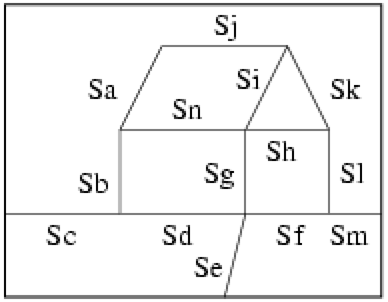


Image 2 **L**

RP and RL are connection relations.

P = {S1,S2,S3,S4,S5,S6,S7,S8,S9,S10,S11}.

L = {Sa,Sb,Sc,Sd,Se,Sf,Sg,Sh,Si,Sj,Sk,Sl,Sm}.

R_P = { (S1,S2), (S1,S5), (S1,S6), (S2,S3), (S2,S4), (S3,S4), (S3,S9), (S4,S5), (S4,S7), (S4,S11), (S5,S6), (S5,S7), (S5,S11), (S6,S8), (S6,S11), (S7,S9), (S7,S10), (S7,S11), (S8,S10), (S8,S11), (S9,S10) }.

R_L = { (Sa,Sb), (Sa,Sj), (Sa,Sn), (Sb,Sc), (Sb,Sd), (Sb,Sn), (Sc,Sd), (Sd,Se), (Sd,Sf), (Sd,Sg), (Se,Sf), (Se,Sg), (Sf,Sg), (Sf,Sl), (Sf,Sm), (Sg,Sh), (Sg,Si), (Sg,Sn), (Sh,Si), (Sh,Sk), (Sh,Sl), (Sh,Sn), (Si,Sj), (Si,Sk), (Si,Sn), (Sj,Sk), (Sk,Sl), (Sl,Sm) }.

f(S1)=Sj f(S4)=Sn f(S7)=Sg f(S10)=Sf

f(S2)=Sa f(S5)=Si f(S8) = Sl f(S11)=Sh

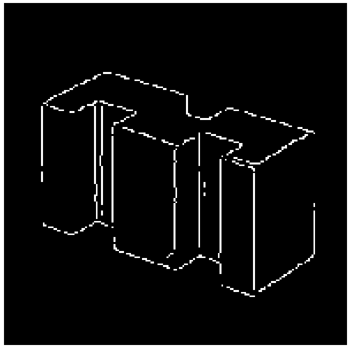
f(S3)=Sb f(S6)=Sk f(S9)=Sd

But this is too simplistic

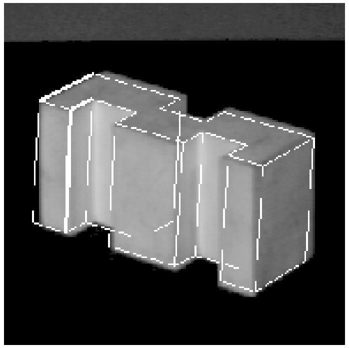
- The model specifies all the features of the object that may appear in the image.
- Some of them don’t appear at all, due to occlusion or failures at low or mid level.
- Some of them are broken and not recognized.
- Some of them are distorted.
- Relationships don’t all hold.

TRIBORS: view class matching of polyhedral objects

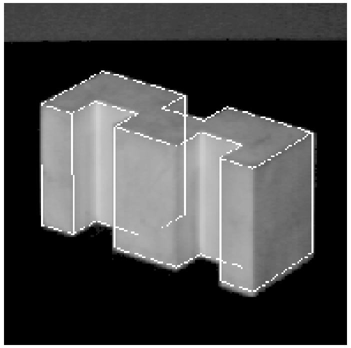
edges from image



model overlayed



improved location



- A view-class is a typical 2D view of a 3D object.

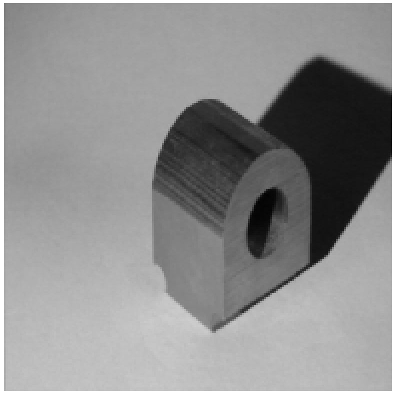
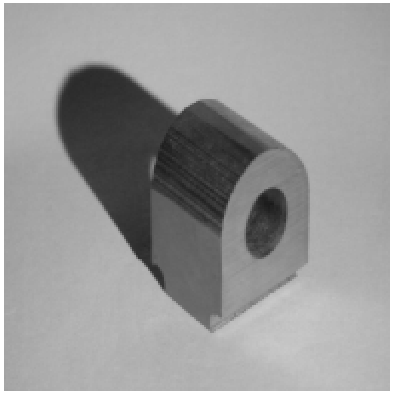
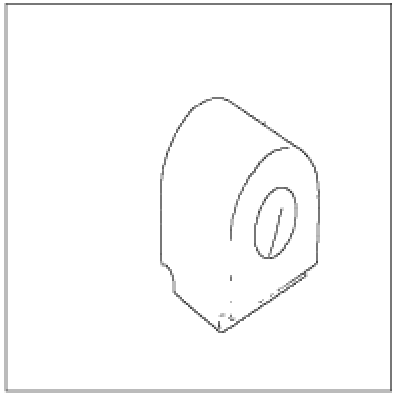
- Each object had 4-5 view classes (hand selected).

- The representation of a view class for matching included:
  - triplets of line segments visible in that class
  - the probability of detectability of each triplet

The first version of this program used iterative-deepening A* search.

RIO: Relational Indexing for Object Recognition

- RIO worked with more complex parts that could have
  - planar surfaces
  - cylindrical surfaces
  - threads

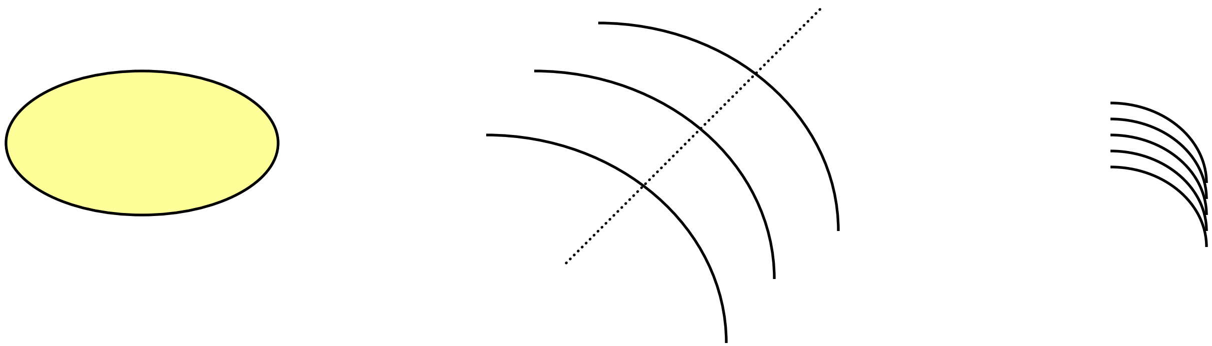
Object Representation in RIO

- 3D objects are represented by a 3D mesh and set of 2D view classes.

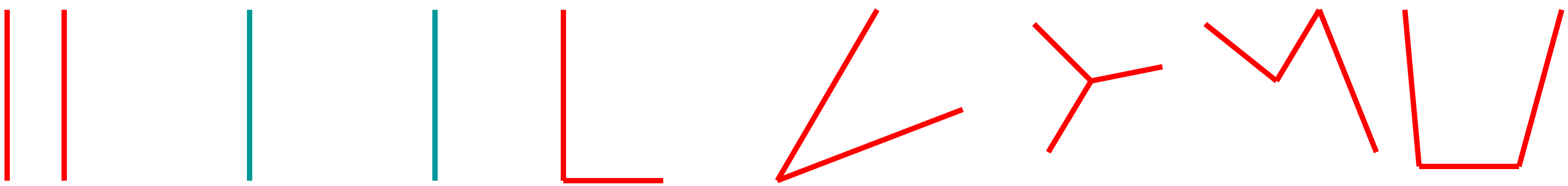
- Each view class is represented by an attributed graph whose nodes are features and whose attributed edges are relationships.

- For purposes of indexing, attributed graphs are stored as sets of 2-graphs, graphs with 2 nodes and 2 relationships.

RIO Features



ellipses coaxials coaxials-multi

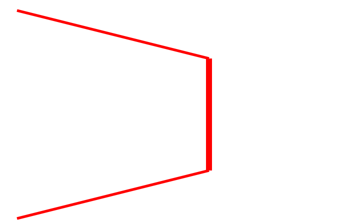


parallel lines close and far

junctions L V Y

triples Z U

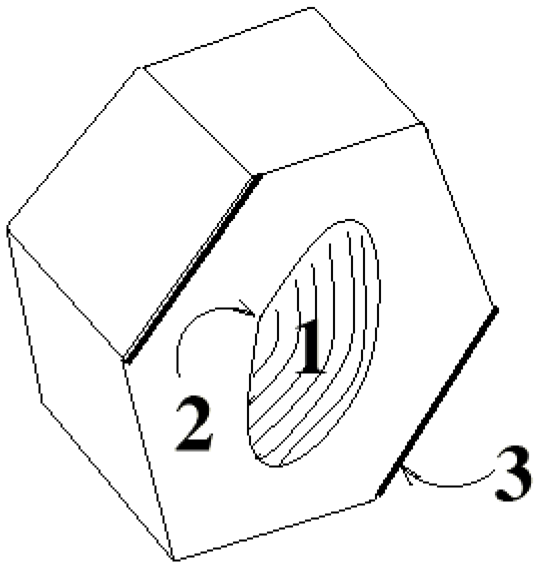
RIO Relationships



- share one arc
- share one line
- share two lines
- coaxial
- close at extremal points
- bounding box encloses / enclosed by

Hexnut Object

MODEL-VIEW



RELATIONS:
a: encloses
b: coaxial

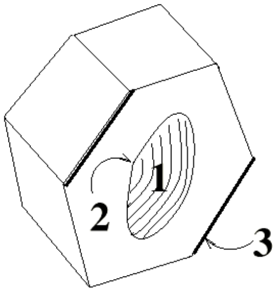
FEATURES:
1: coaxials-multi
2: ellipse
3: parallel lines

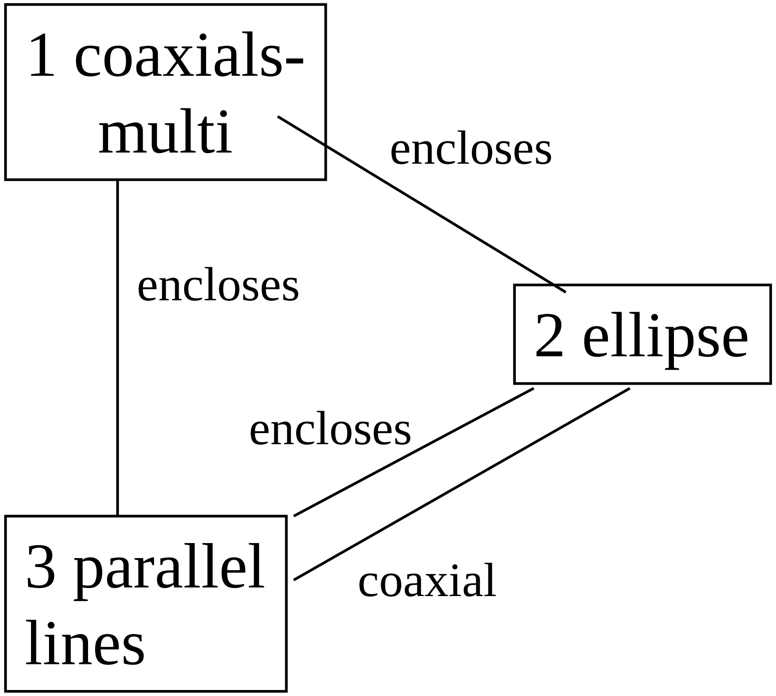
How are 1, 2, and 3 related?

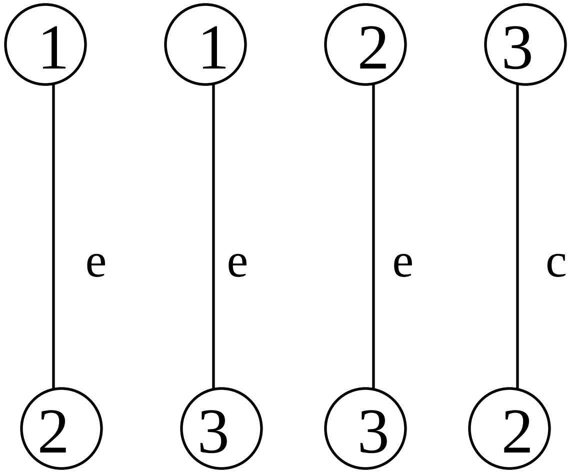
What other features and relationships can you find?

Graph and 2-Graph Representations

MODEL-VIEW







RDF!

Relational Indexing for Recognition

Preprocessing (off-line) Phase

for each model view Mi in the database

- encode each 2-graph of Mi to produce an index
- store Mi and associated information in the indexed bin of a hash table H

Matching (on-line) phase

1. Construct a relational (2-graph) description D for the scene

2. For each 2-graph G of D

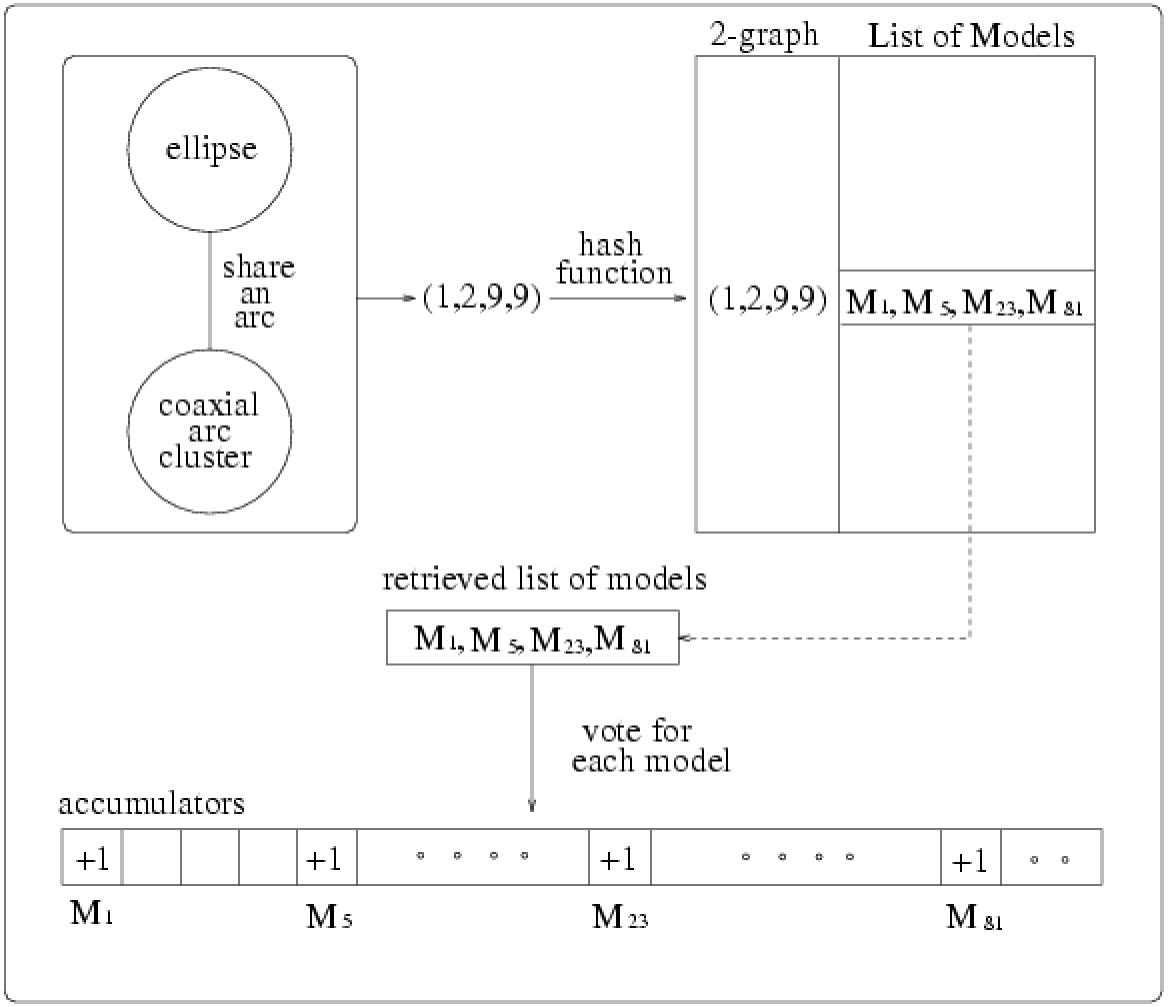
   - encode it, producing an index to access the hash table H

   - cast a vote for each Mi in the associated bin

3. Select the Mi's with high votes as possible hypotheses

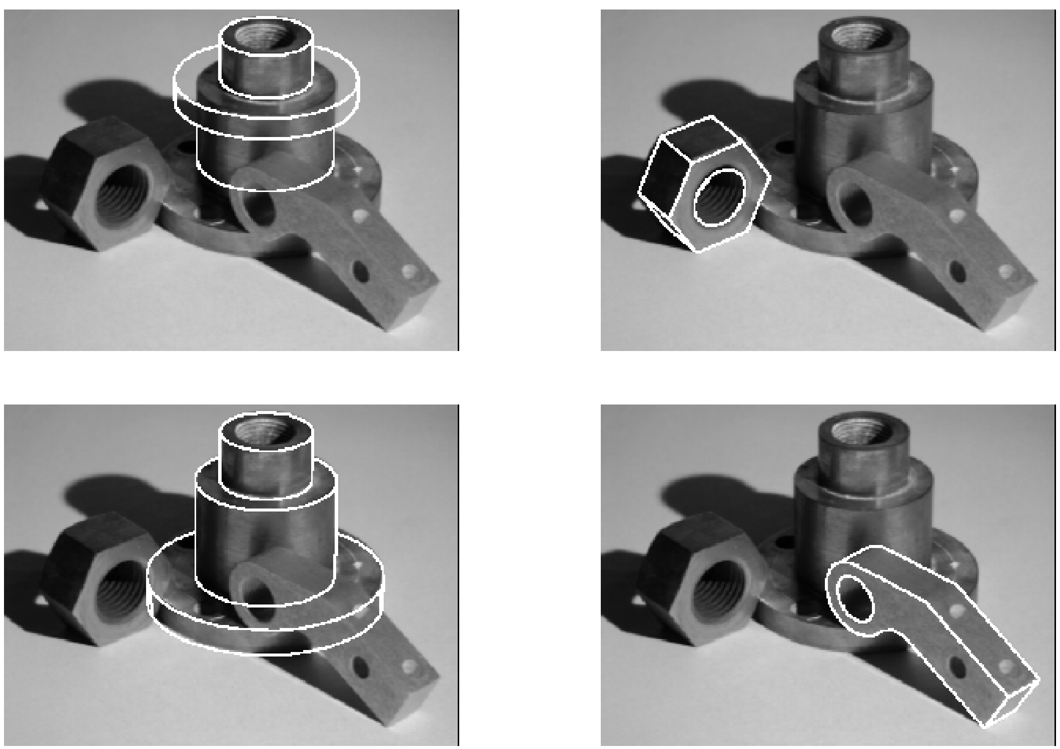
4. Verify or disprove via alignment, using the 3D meshes

The Voting Process



RIO Verifications

incorrect
hypothesis



1. The matched features of the hypothesized object are used to determine its **pose**.

2. The **3D mesh** of the object is used to project all its features onto the image.

3. A **verification procedure** checks how well the object features line up with edges on the image.

Use of classifiers is big in computer vision today.

- 2 Examples:
  - Rowley's Face Detection using neural nets
  - Yi's image classification using EM

Object Detection:
Rowley's Face Finder

1. convert to gray scale
2. normalize for lighting
3. histogram equalization
4. apply neural net(s)
   trained on 16K images

What data is fed to
the classifier?

32 x 32 windows in
a pyramid structure

Object Class Recognition using Images of Abstract Regions

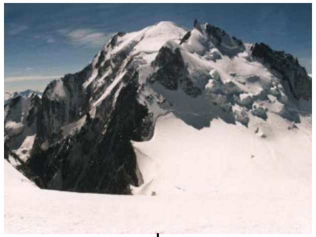
Yi Li, Jeff A. Bilmes, and Linda G. Shapiro
Department of Computer Science and Engineering
Department of Electrical Engineering
University of Washington

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

?

?

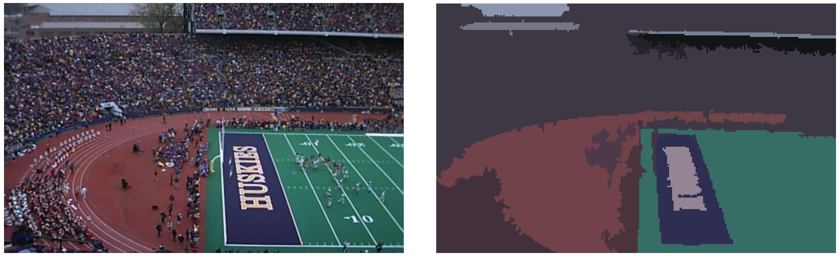
?

?

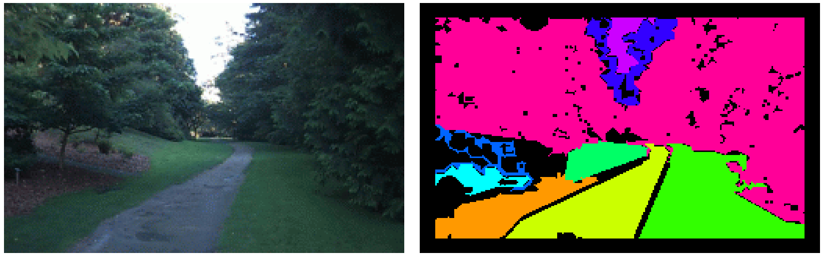
●●●

Image Features for Object Recognition

- Color

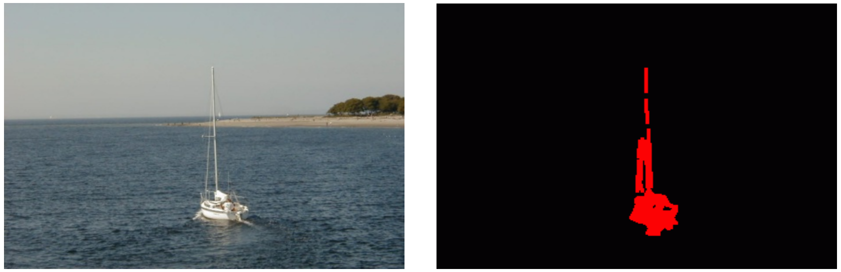


- Texture



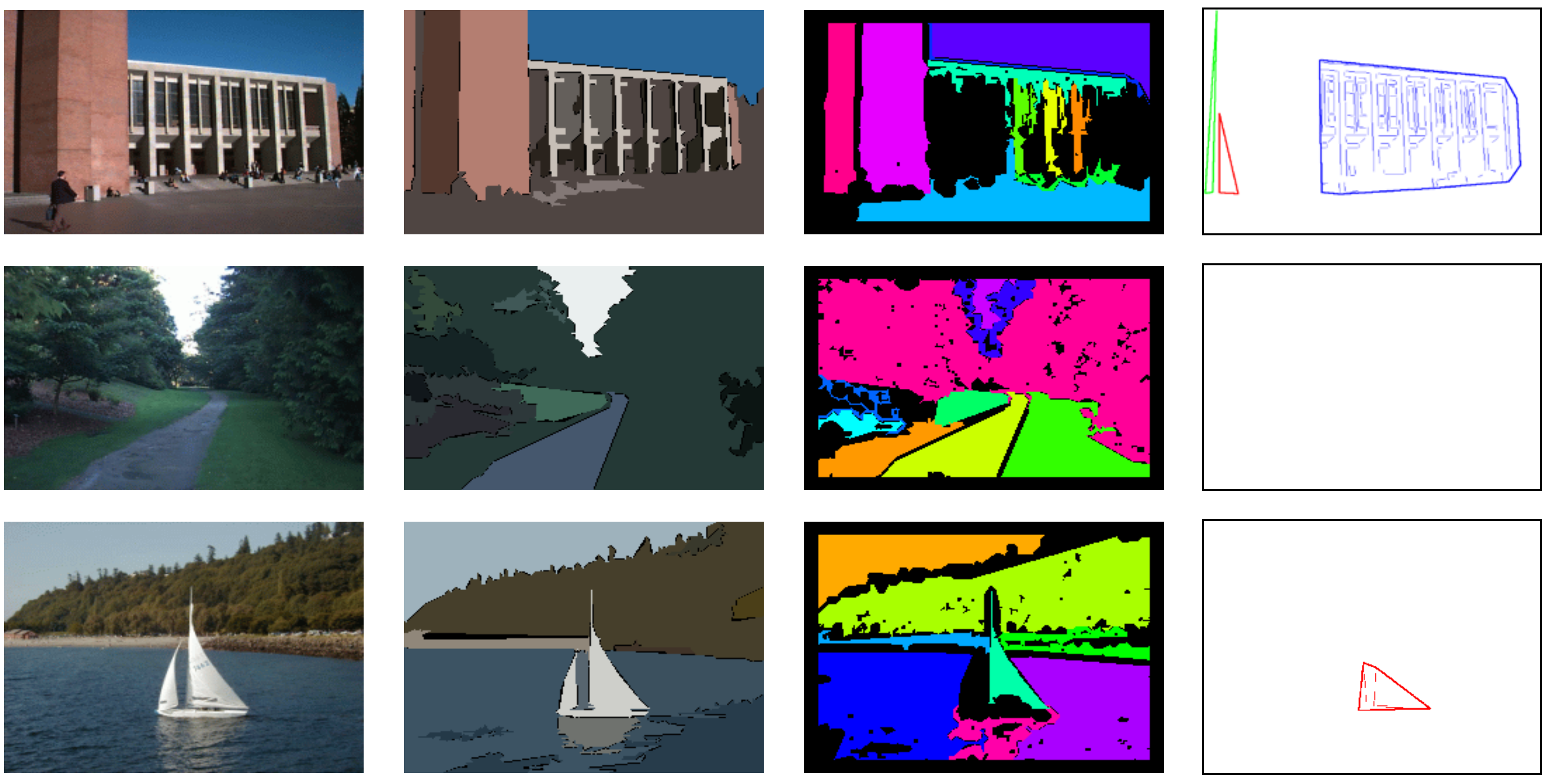
- Structure

- Context



Abstract Regions

Original Images | Color Regions | Texture Regions | Line Clusters



Abstract Regions

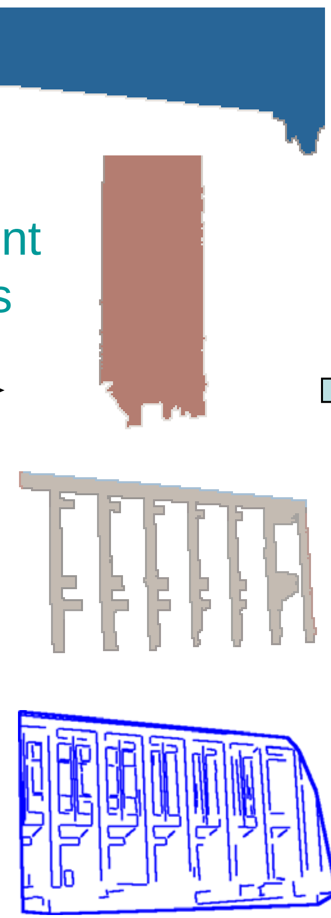
Multiple segmentations whose regions are not labeled; a list of labels is provided for each training image.

image

labels

{sky, building}

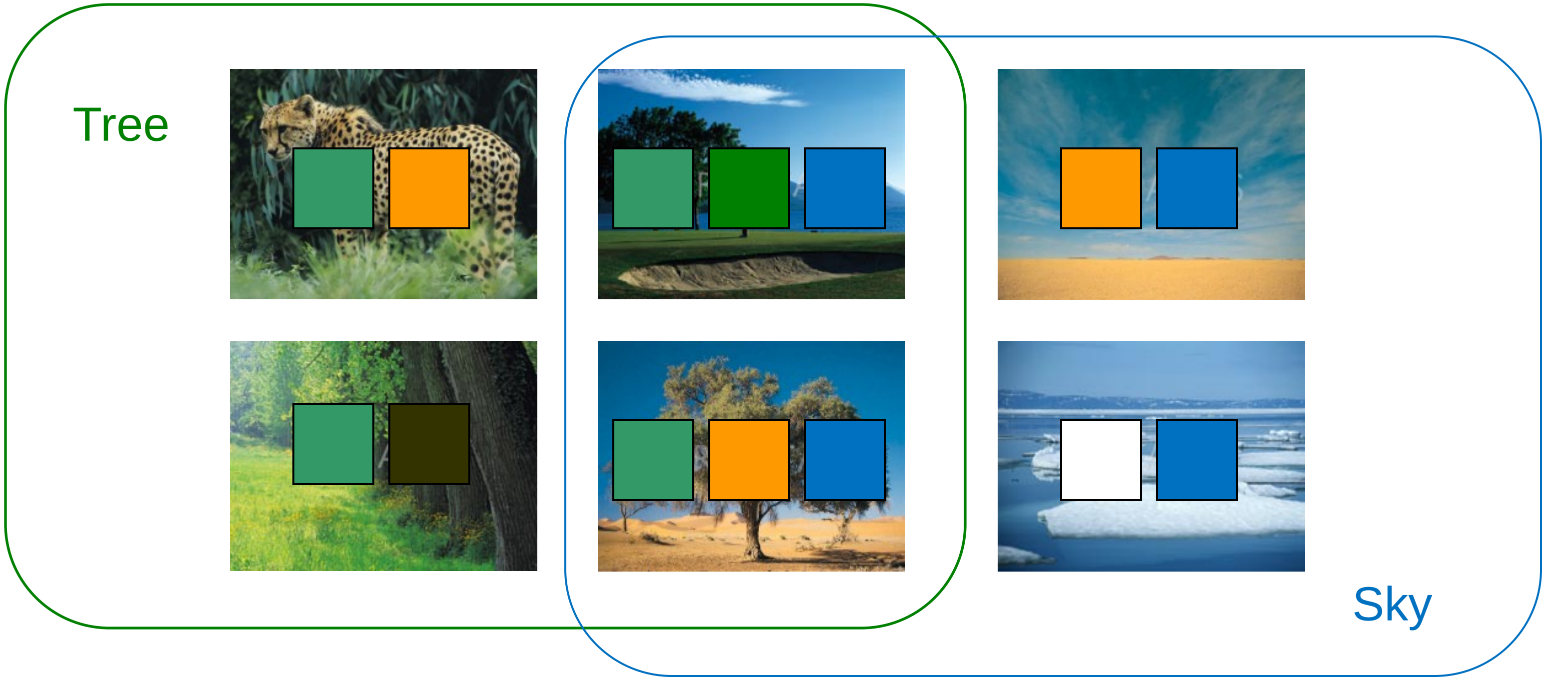
various different segmentations



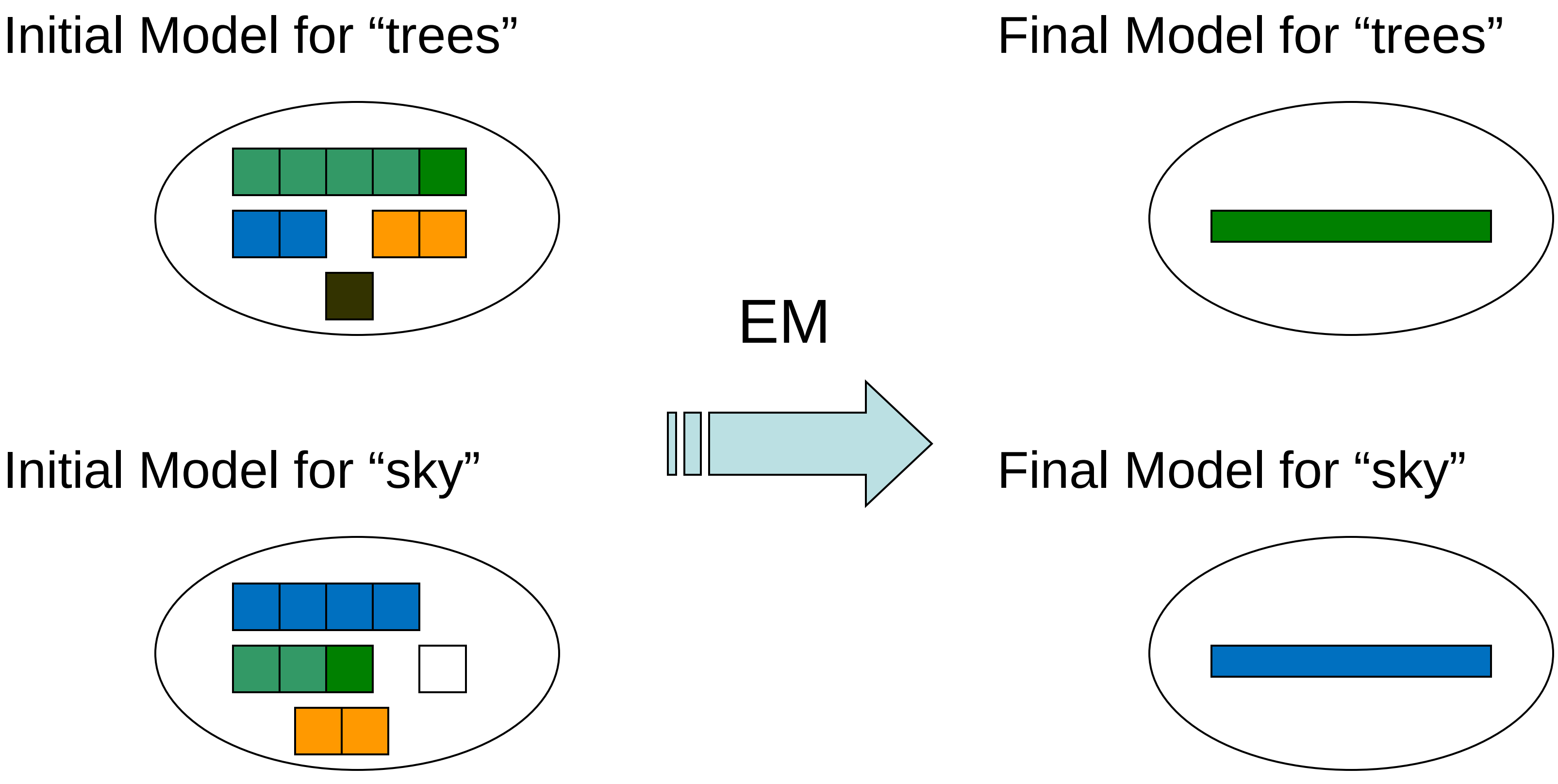
region attributes from several different types of regions

Model Initial Estimation

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



EM Classifier: the Idea



A Better Approach to Combining Different Feature Types

Phase 1:

- Treat each type of abstract region separately

- For abstract region type ***a*** and for object class ***o***, use the EM algorithm to construct clusters that are multivariate Gaussians over the features for type ***a*** regions.

Aggregate Scores for Color

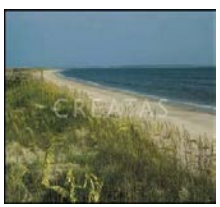
Components

1	2	3	4	5	6	7	8

beach

.93	.16	.94	.24	.10	.99	.32	.00

beach



.66	.80	.00	.72	.19	.01	.22	.02

not beach

.43	.03	.00	.00	.00	.00	.15	.00

Phase 2 Learning

- Let C_i be row i of the training matrix.
- Each such row is a feature vector for the color features of regions of image I_i that relates them to the Phase 1 components.
- Now we can use a second-stage classifier to learn $P(o|I_i)$ for each object class o and image I_i .

Multiple Feature Case

- We calculate separate Gaussian mixture models for each different features type:

- Color: C_i
- Texture: T_i
- Structure: S_i

- and any more features we have (motion).

Now we concatenate the matrix rows from the different region types to obtain a multi-feature-type training matrix and train a neural net classifier to classify images.

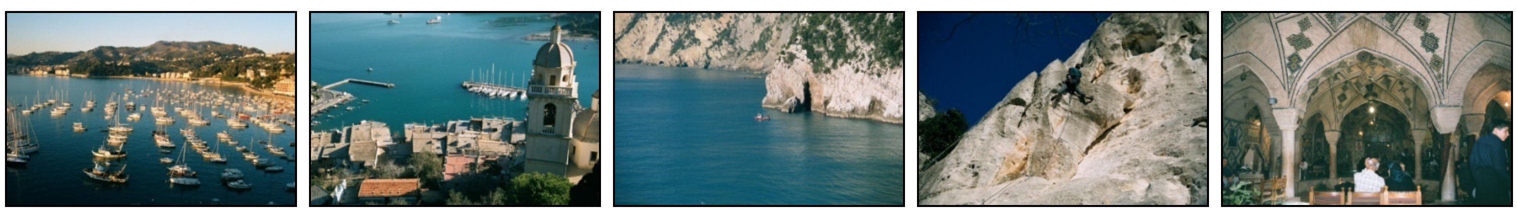
$$\begin{array}{ccccc} \textit{color} & \textit{texture} & \textit{structure} & & \textit{everything} \\ \begin{pmatrix} C_1^+ \\ C_2^+ \\ . \\ . \\ C_1^- \\ C_2^- \\ . \\ . \end{pmatrix} & \begin{pmatrix} T_1^+ \\ T_2^+ \\ . \\ . \\ T_1^- \\ T_2^- \\ . \\ . \end{pmatrix} & \begin{pmatrix} S_1^+ \\ S_2^+ \\ . \\ . \\ S_1^- \\ S_2^- \\ . \\ . \end{pmatrix} & \longrightarrow & \begin{pmatrix} C_1^+ & T_1^+ & S_1^+ \\ C_2^+ & T_2^+ & S_2^+ \\ . & . & . \\ . & . & . \\ C_1^- & T_1^- & S_1^- \\ C_2^- & T_2^- & S_2^- \\ . & . & . \\ . & . & . \end{pmatrix} \end{array}$$

Groundtruth Data Set: Top Results

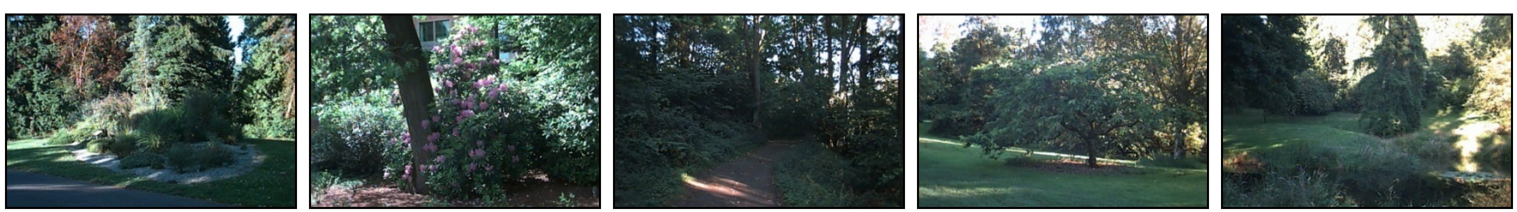
Asian city

Cannon beach

Italy



park



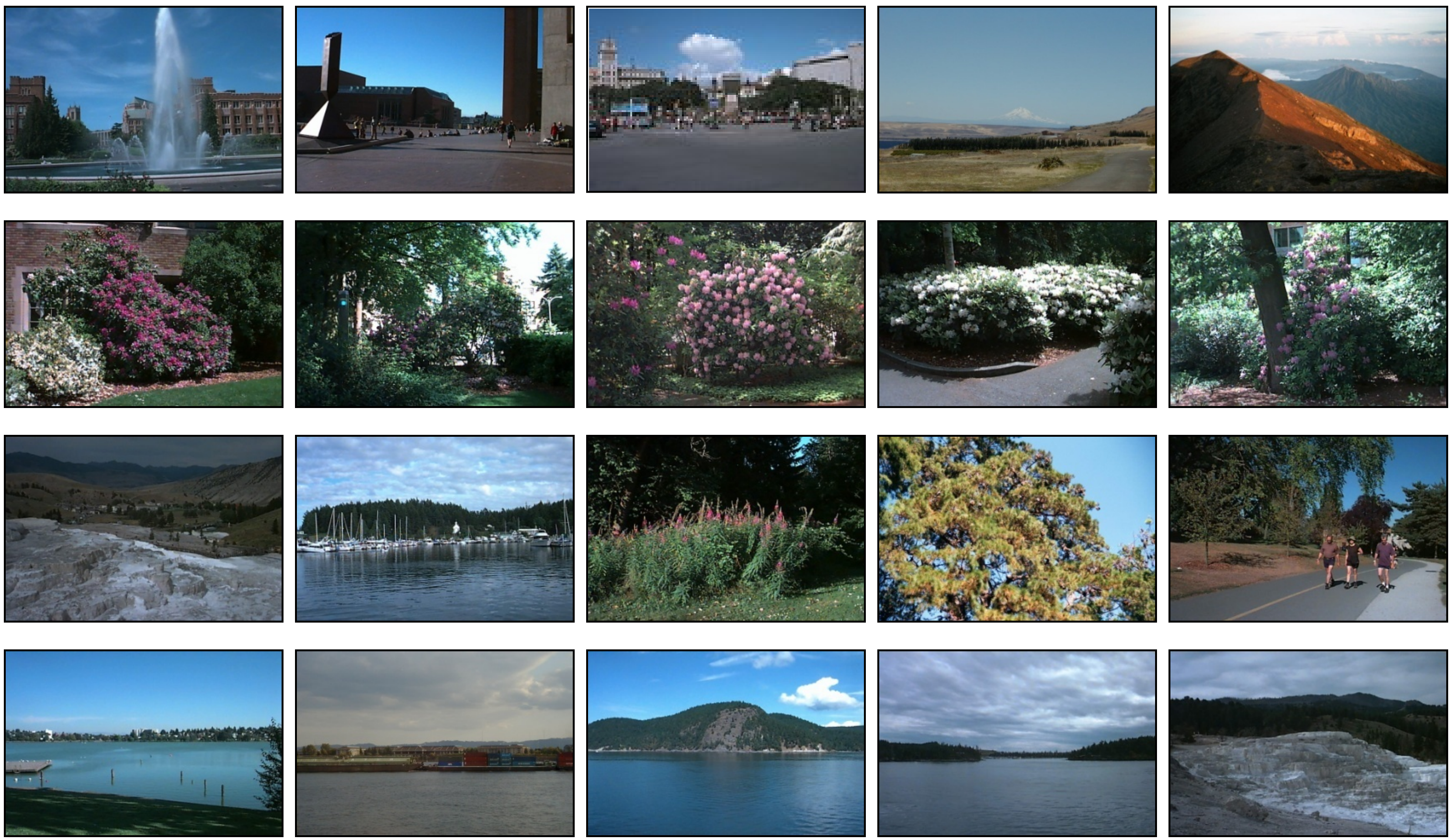
Groundtruth Data Set: Top Results

sky

spring flowers

tree

water



Groundtruth Data Set: Annotation Samples

tree(97.3), **bush**(91.6), **spring flowers**(90.3), **flower**(84.4), park(84.3), **sidewalk**(67.5), **grass**(52.5), **pole**(34.1)

sky(99.8), **Columbia gorge**(98.8), lantern(94.2), **street**(89.2), house(85.8), bridge(80.8), car(80.5), hill(78.3), boat(73.1), pole(72.3), **water**(64.3), mountain(63.8), **building**(9.5)

sky(95.1), **Iran**(89.3), house(88.6), **building**(80.1), boat(71.7), bridge(67.0), **water**(13.5), **tree**(7.7)

Italy(99.9), grass(98.5), **sky**(93.8), rock(88.8), **boat**(80.1), **water**(77.1), Iran(64.2), stone(63.9), bridge(59.6), **European**(56.3), sidewalk(51.1), **house**(5.3)