Content-Based Image Retrieval

ECE P 596
Autumn 2019

Linda Shapiro

Content-Based Image Retrieval

- Queries
- Commercial Systems
- Retrieval Features
- Indexing in the FIDS System
- Lead-in to Object Recognition

Content-based Image Retrieval (CBIR)

Searching a large database for images that *match* a query:

- What kinds of databases?
- What kinds of queries?
- What constitutes a match?
- How do we make such searches efficient?

Applications

- Art Collections
  e.g. Fine Arts Museum of San Francisco
- Medical Image Databases
  CT, MRI, Ultrasound, The Visible Human
- Scientific Databases
  e.g. Earth Sciences
- General Image Collections for Licensing
  Corbis, Getty Images
- The World Wide Web
  Google, Microsoft, etc

What is a query?

- an image you already have
- a rough sketch you draw
- a symbolic description of what you want
  e.g. an image of a man and a woman on a beach

Some Systems You Can Try

- Corbis ~~sells~~ sold high-quality images for use in advertising, marketing, illustrating, etc. Corbis was sold to a Chinese company, but

- Getty images now provides the image sales.

- http://www.gettyimages.com/search/2/image?excludenudity=true&sort=best

Google Image

- Google Images http://www.google.com/imghp

Try the camera icon.

Microsoft Bing

- http://www.bing.com/

Try Visual Search

Problem with Text-Based Search

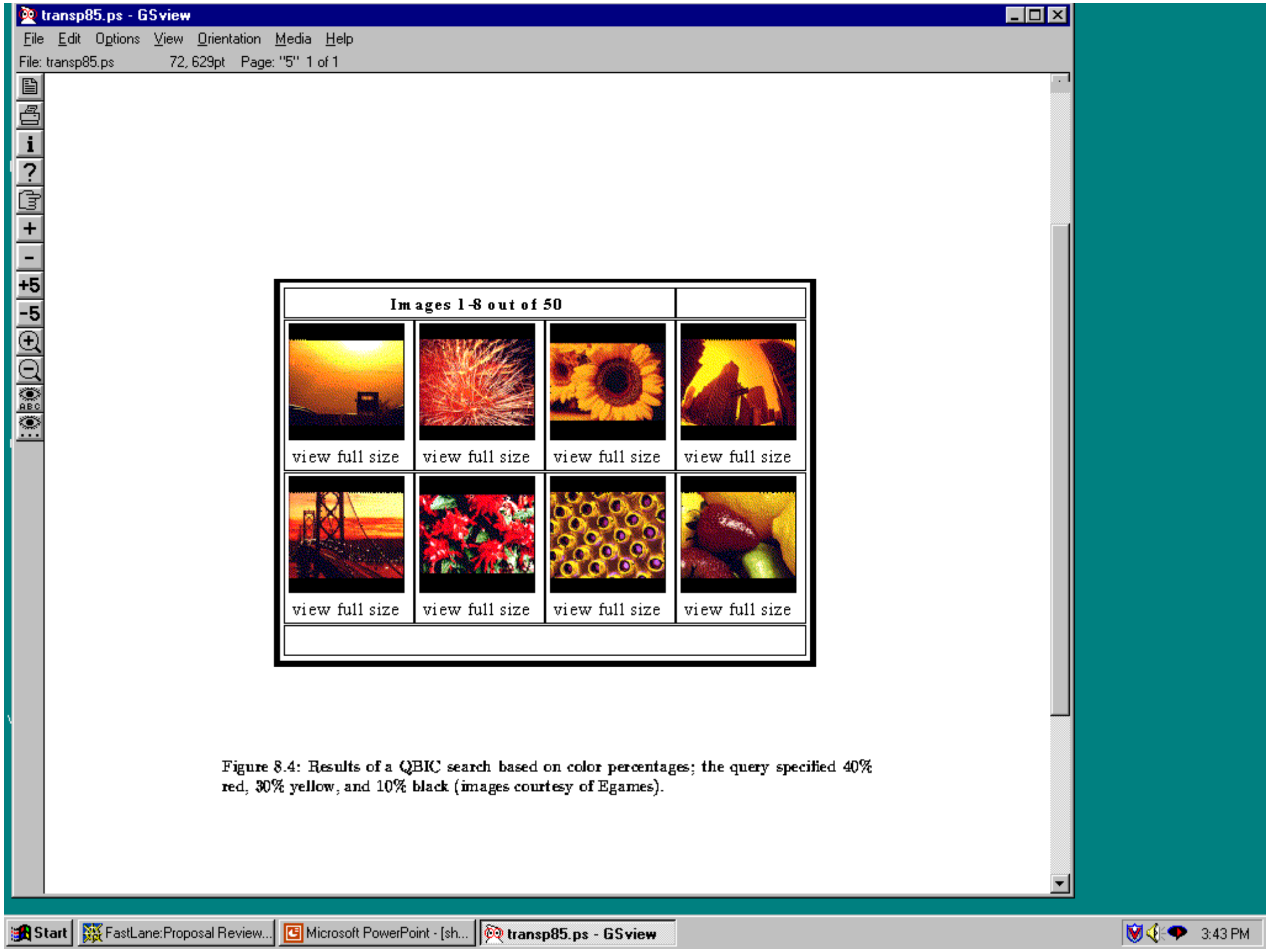
- Retrieval for pigs for the color chapter of my book
- Small company (was called Ditto)
- Allows you to search for pictures from web pages

Features

- Color (histograms, gridded layout, wavelets)
- Texture (Laws, Gabor filters, local binary pattern)
- Shape (first segment the image, then use statistical
  or structural shape similarity measures)
- Objects and their Relationships

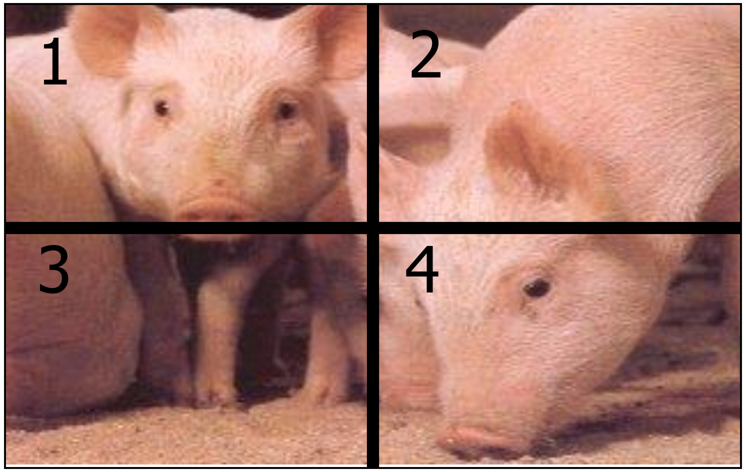
This is the most powerful, but you have to be able to recognize the objects!

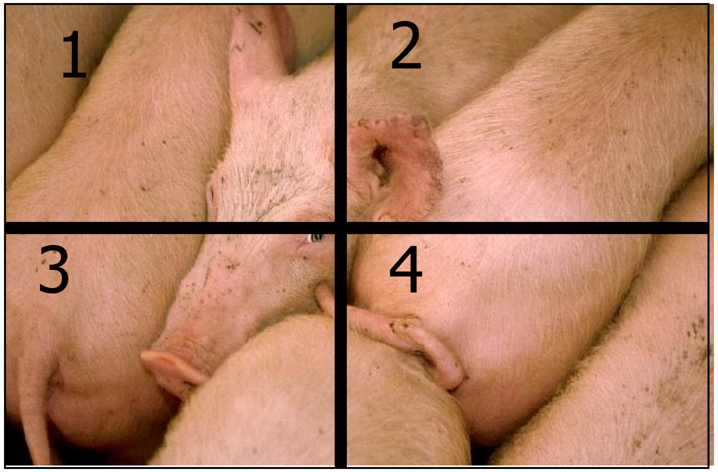
Color Histograms



Gridded Color

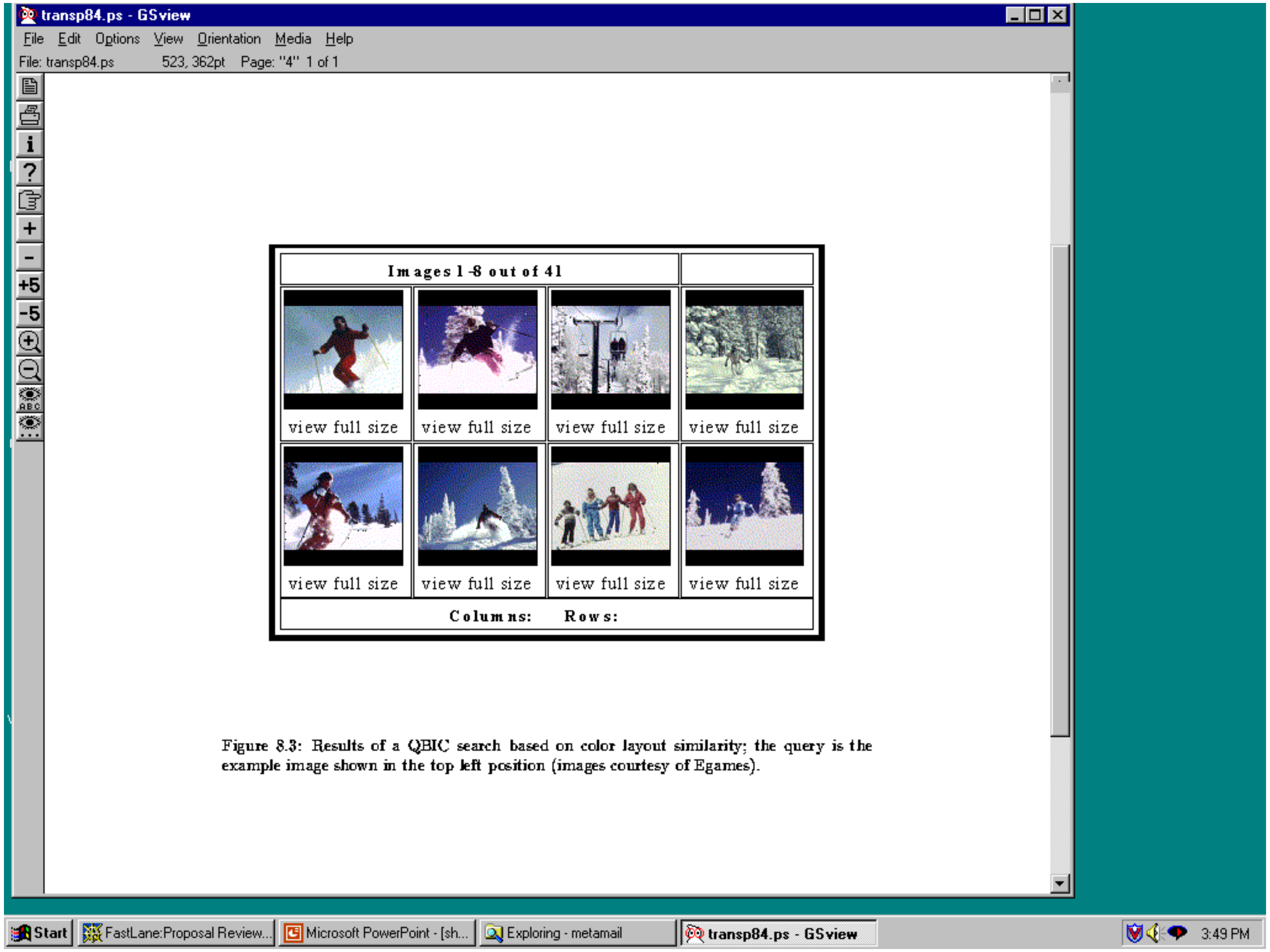
Gridded color distance is the sum of the color distances in each of the corresponding grid squares.





What color distance would you use for a pair of grid squares?

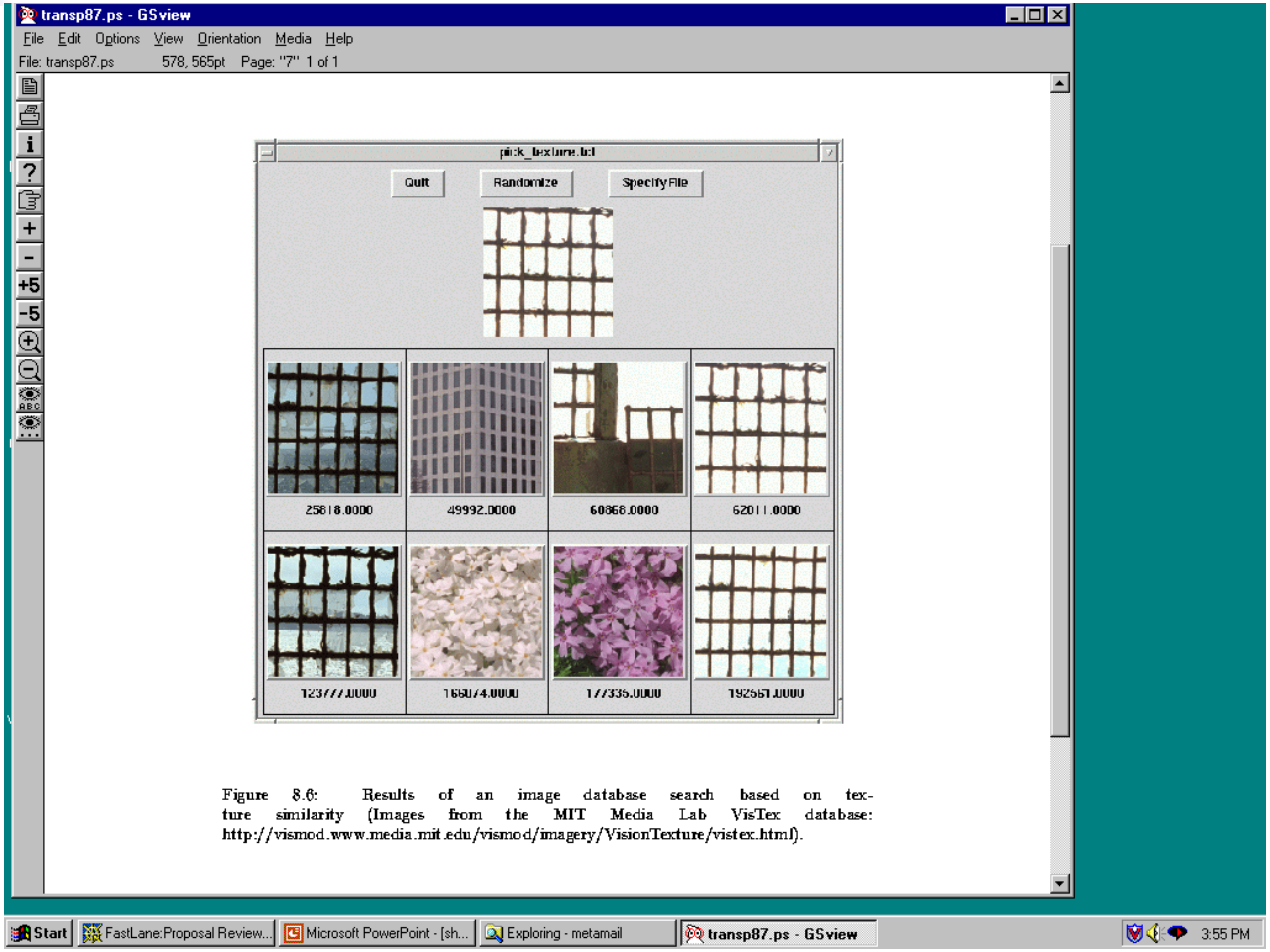
Color Layout (IBM's Gridded Color)



Texture Distances

- Pick and Click (user clicks on a pixel and system retrieves images that have in them a region with similar texture to the region surrounding it.
- Gridded (just like gridded color, but use texture).
- Histogram-based (e.g. compare the LBP histograms).

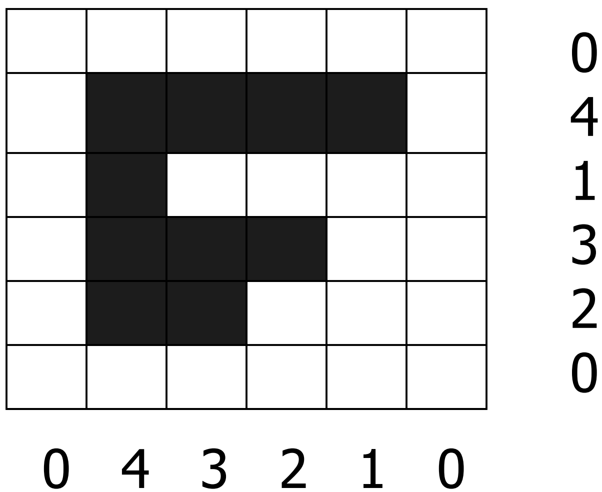
Laws Texture



Shape Distances

- Shape goes one step further than color and texture.
- It requires identification of regions to compare.
- There have been many shape similarity measures suggested for pattern recognition that can be used to construct shape distance measures.

Global Shape Properties: Projection Matching

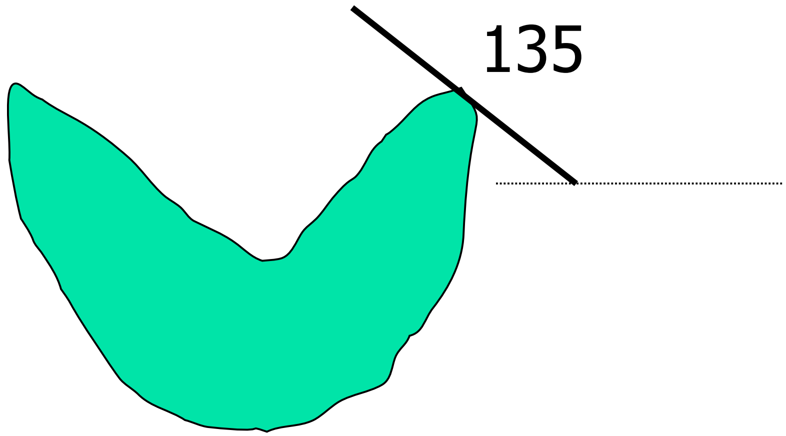


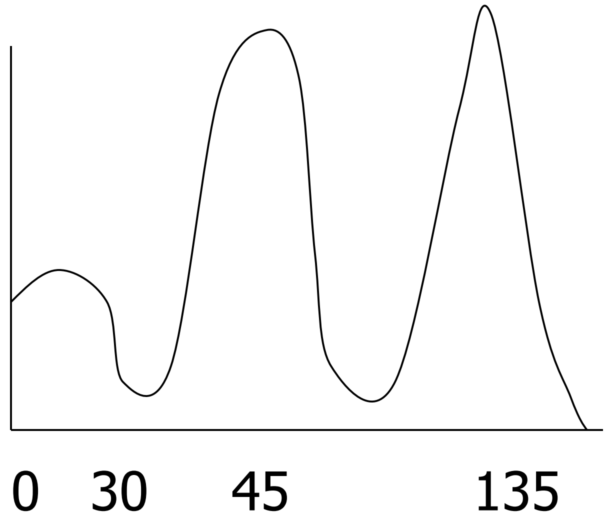
Feature Vector
(0,4,1,3,2,0,0,4,3,2,1,0)

In projection matching, the horizontal and vertical projections form a histogram.

What are the weaknesses of this method? strengths?

Global Shape Properties: Tangent-Angle Histograms





Is this feature invariant to starting point?
Is it invariant to size, translation, rotation?

Boundary Matching

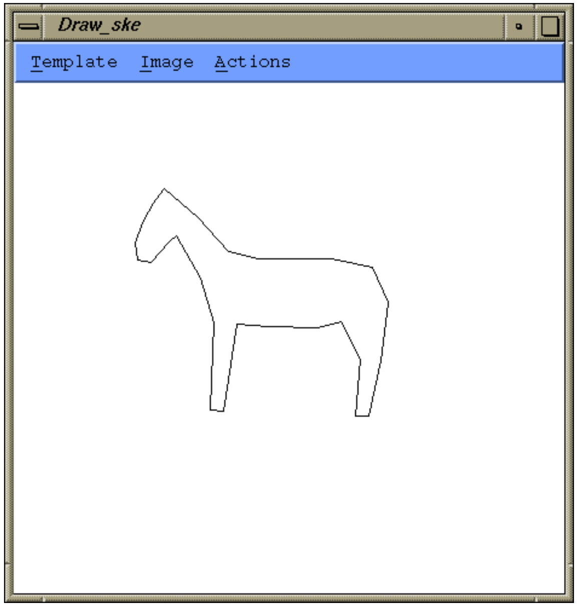
- Fourier Descriptors
- Sides and Angles
- Elastic Matching

  The distance between query shape and image shape has two components:

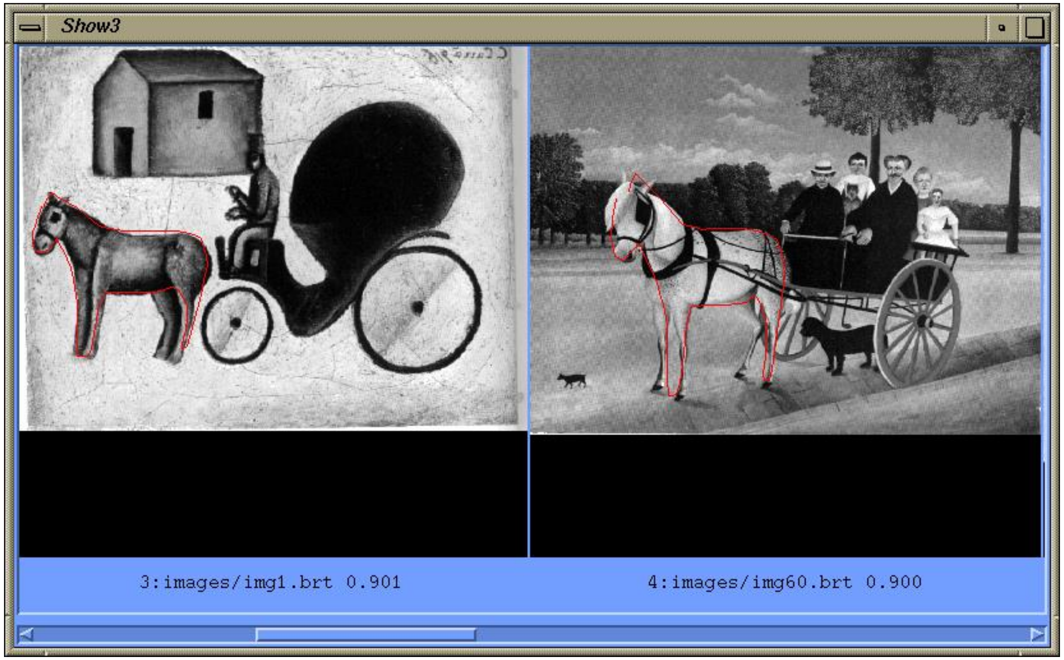
  1. energy required to deform the query shape into one that best matches the image shape

  2. a measure of how well the deformed query matches the image

Del Bimbo Elastic Shape Matching



query



retrieved images

Regions and Relationships

- Segment the image into regions
- Find their properties and interrelationships
- Construct a graph representation with nodes for regions and edges for spatial relationships
- Use graph matching to compare images

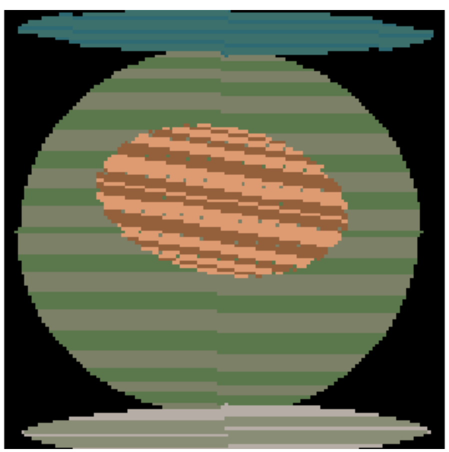
Like what?

Blobworld (Carson et al, 1999)

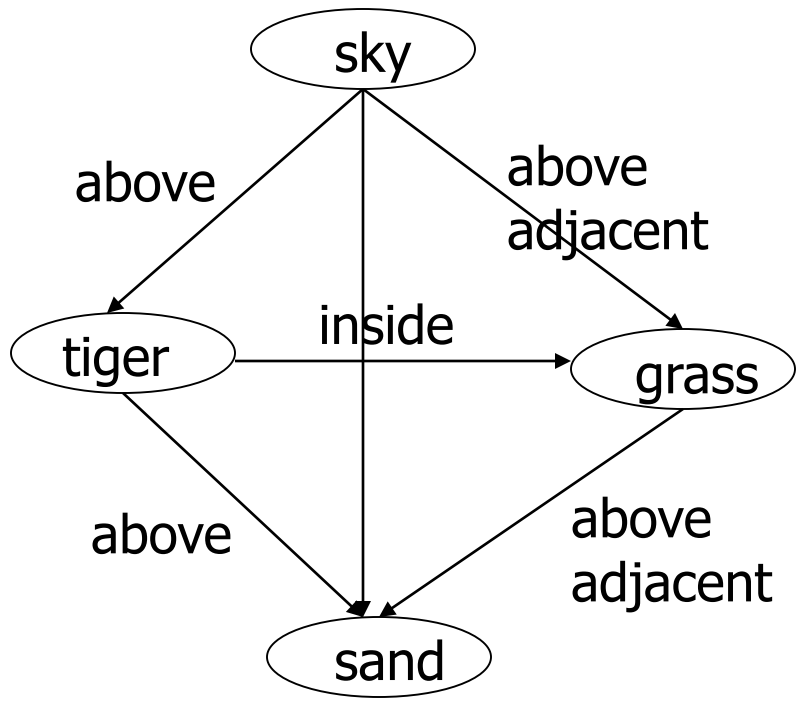
- Segmented the query (and all database images) using EM on color+texture
- Allowed users to select the most important region and what characteristics of it (color, texture, location)
- Asked users if the background was also important

Tiger Image as a Graph (motivated by Blobworld)

image

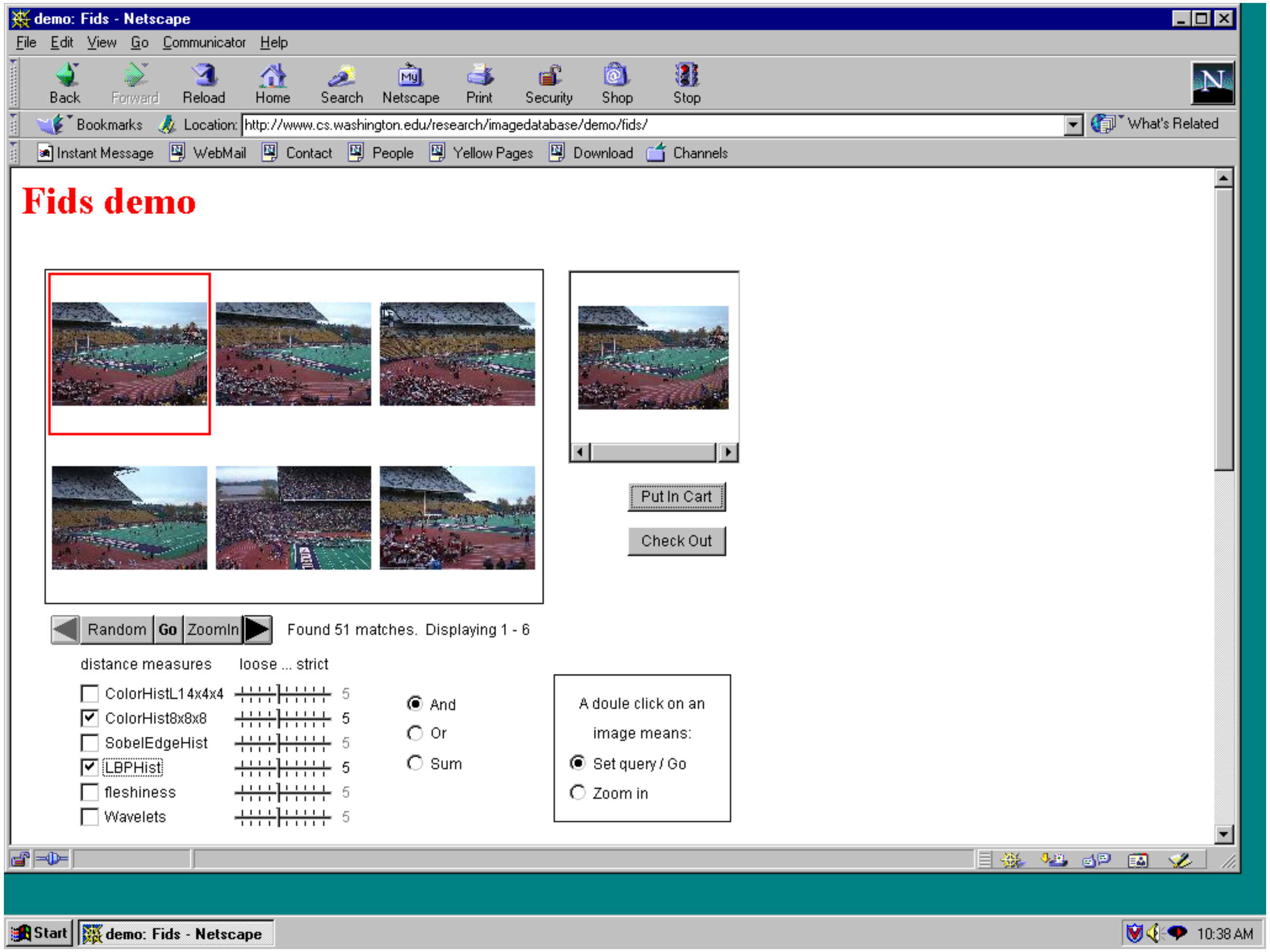


abstract regions



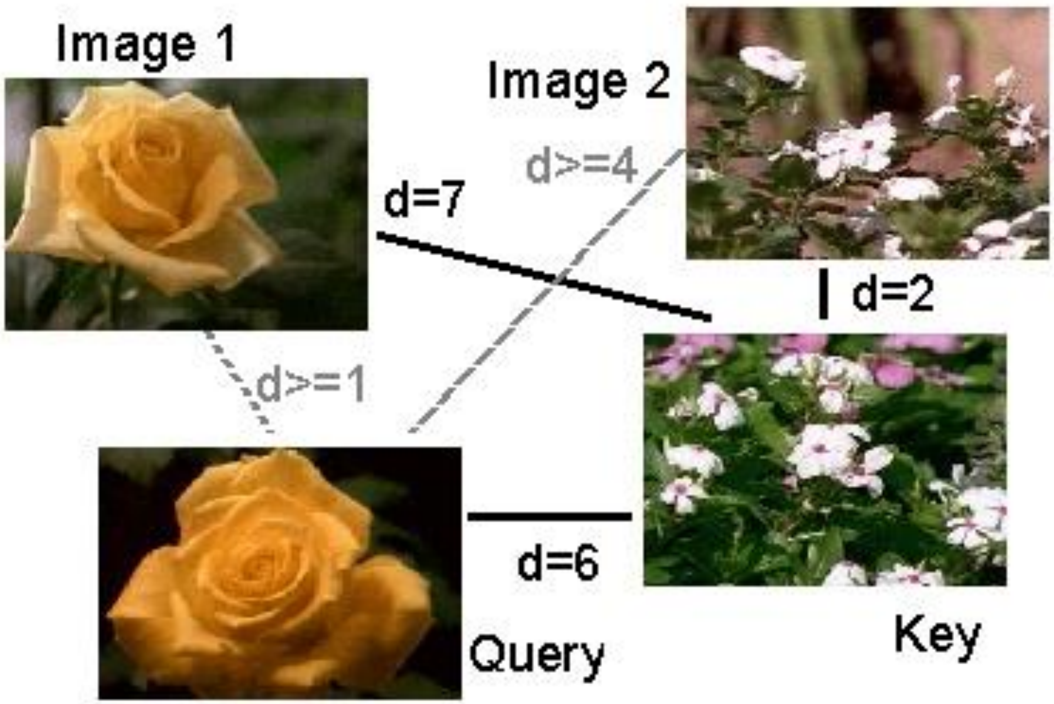
Andy Berman's FIDS System

multiple distance measures
Boolean and linear combinations
efficient indexing using images as keys



Andy Berman's FIDS System:

Use of key images and the triangle inequality for efficient retrieval. d(I,Q) >= |d((I,K) – d(Q,K)|



Andy Berman's FIDS System:

Bare-Bones Triangle Inequality Algorithm

Offline

1. Choose a small set of key images

2. Store distances from database images to keys

Online (given query Q)

1. Compute the distance from Q to each key

2. Obtain lower bounds on distances to database images

3. Threshold or return all images in order of lower bounds

Andy Berman's FIDS System:

Flexible Image Database System: Example

An example from our system using a simple color measure.
images in system: 37,748
threshold: 100 out of 1000
images eliminated: 37,729

Andy Berman's FIDS System:

Bare-Bones Algorithm with Multiple Distance Measures

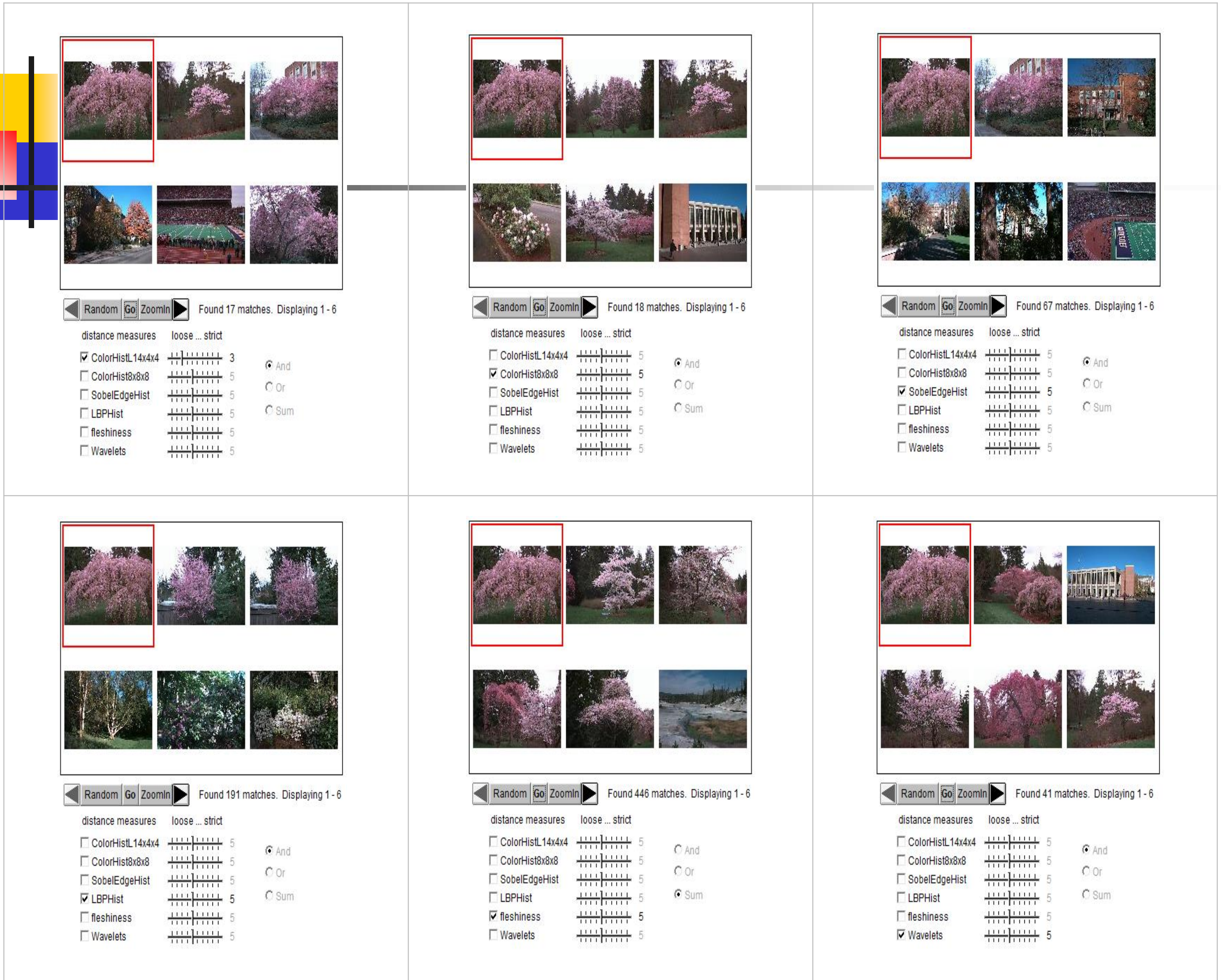
Offline

1. Choose key images for each measure
2. Store distances from database images to keys for all measures

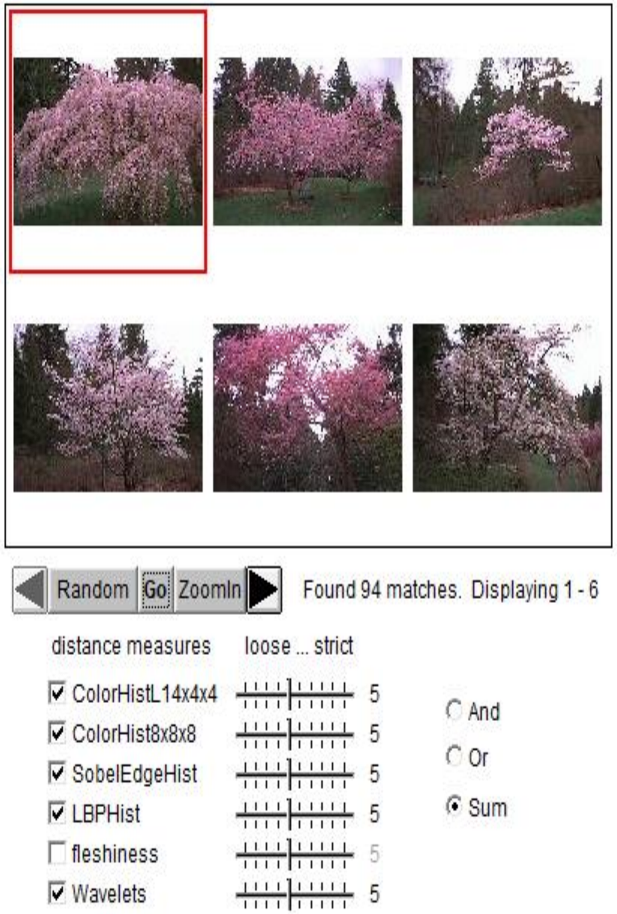
Online (given query Q)

1. Calculate lower bounds for each measure
2. Combine to form lower bounds for composite measures
3. Continue as in single measure algorithm

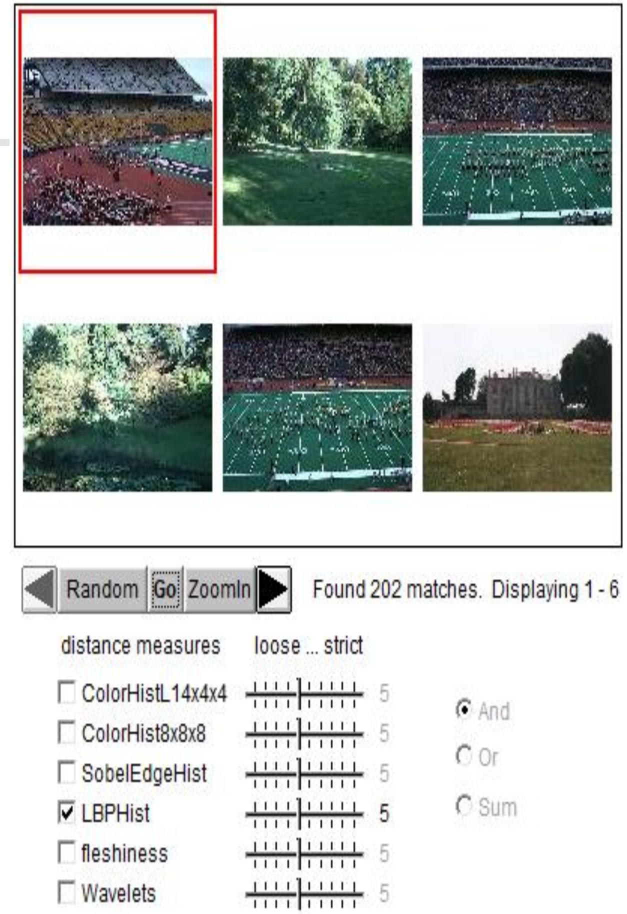
Different Features



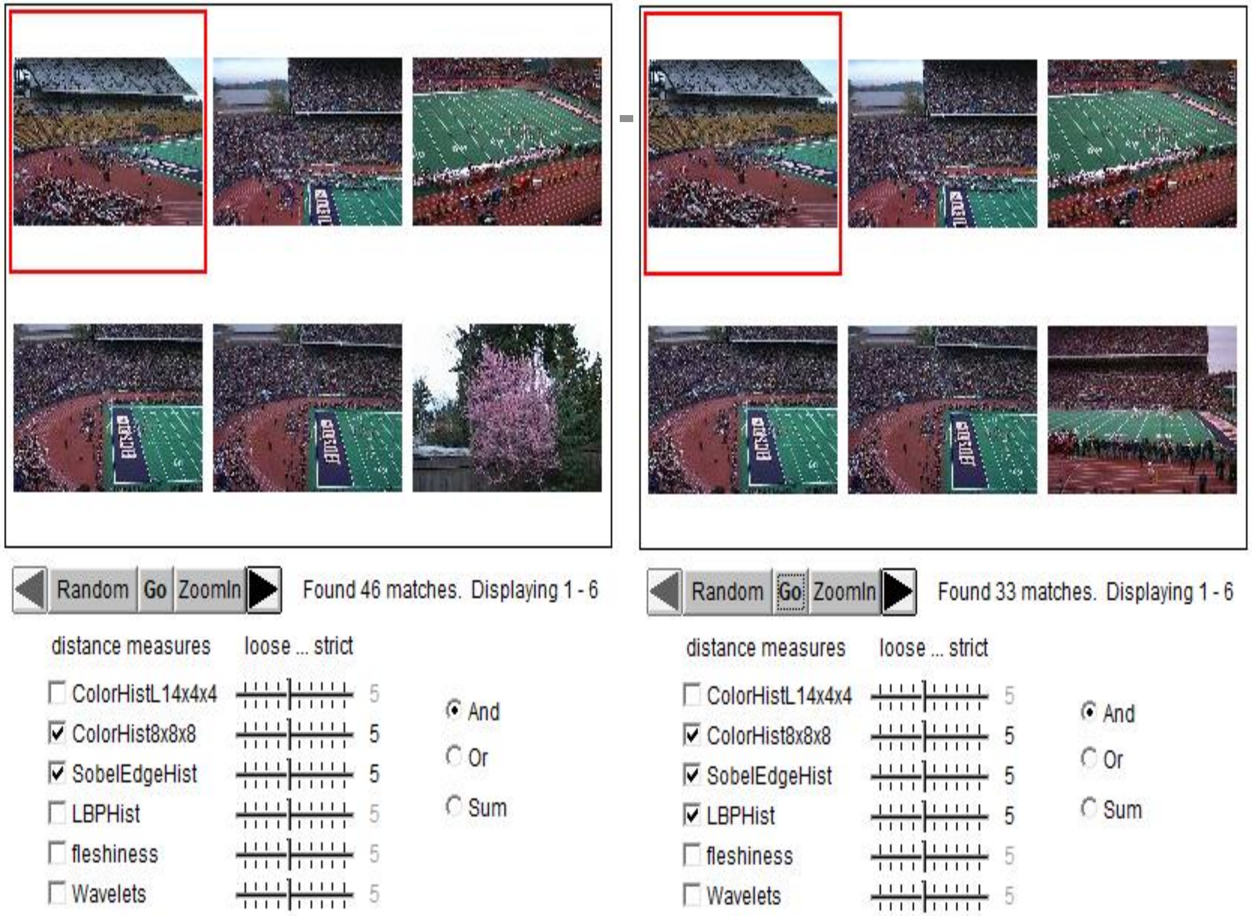
Combined Features



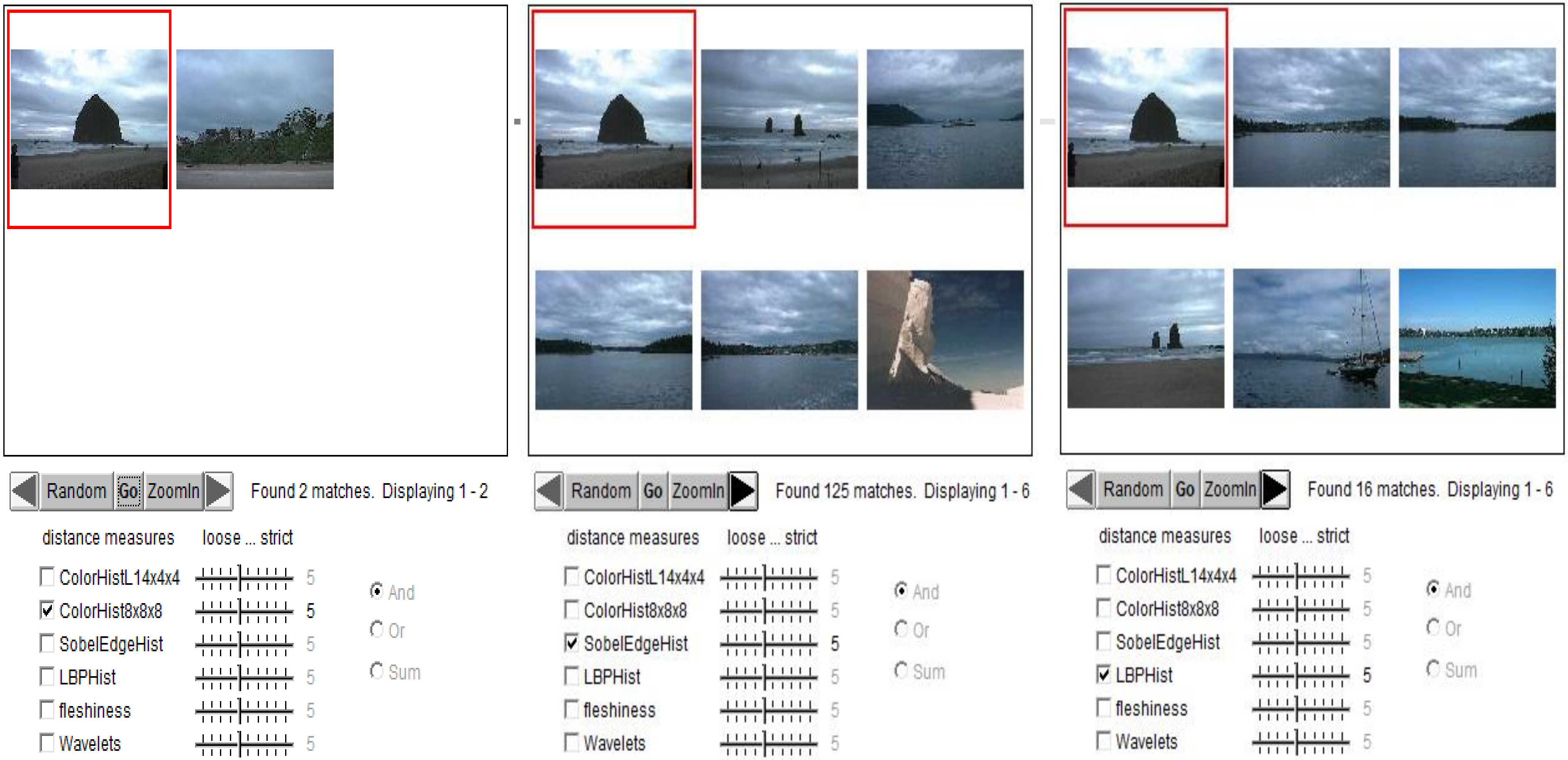
Another example: different features



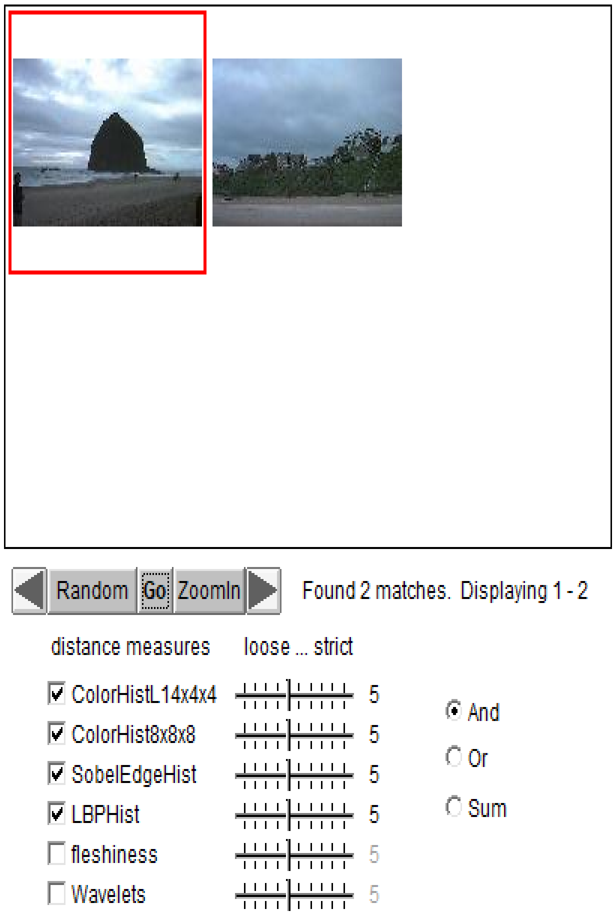
Combined Features

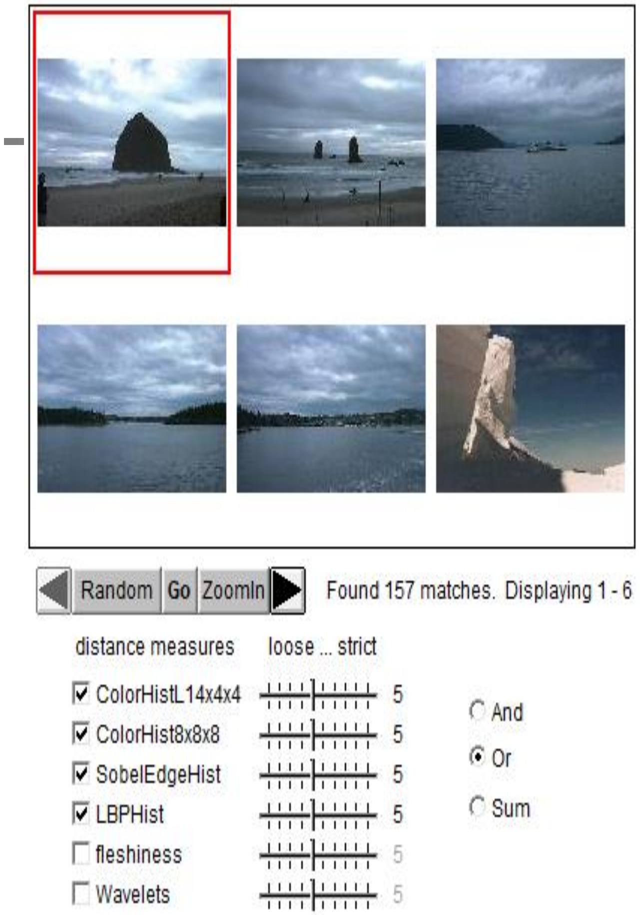


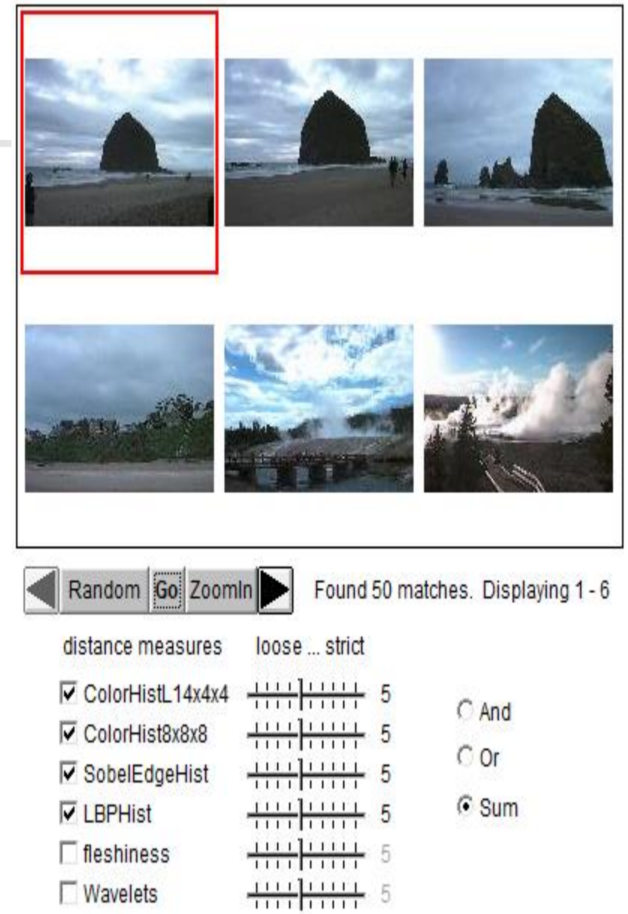
Another example: different features



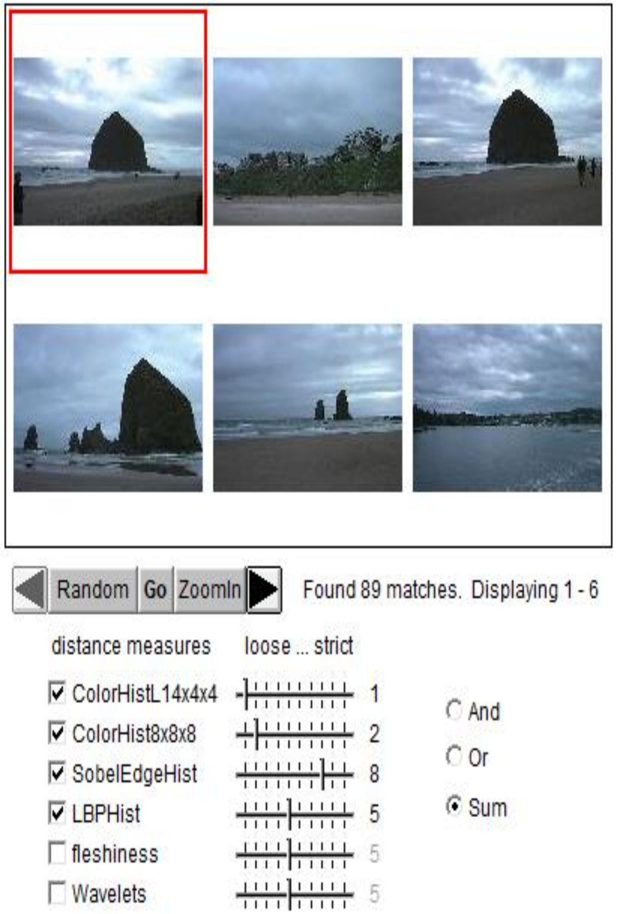
Different ways for combination



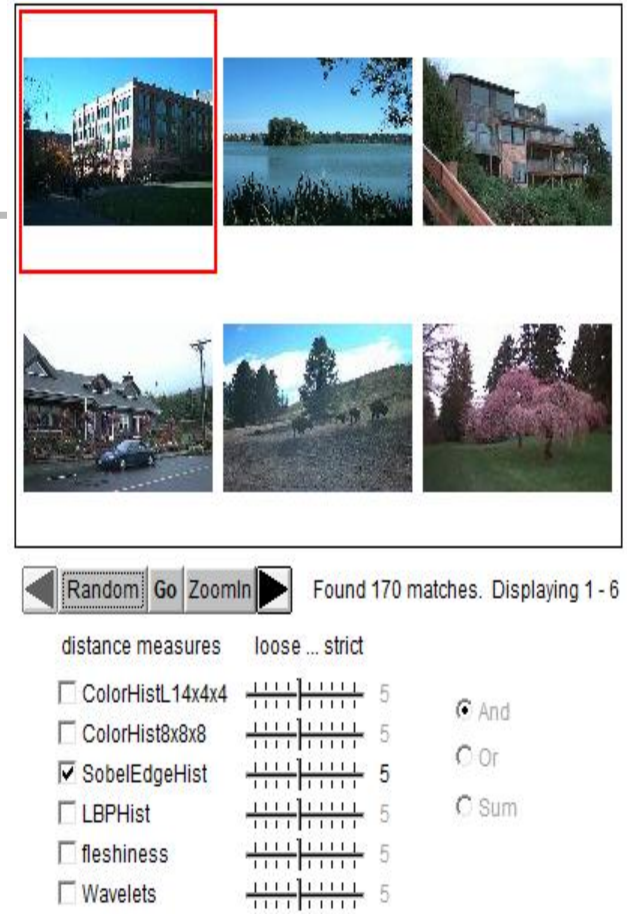


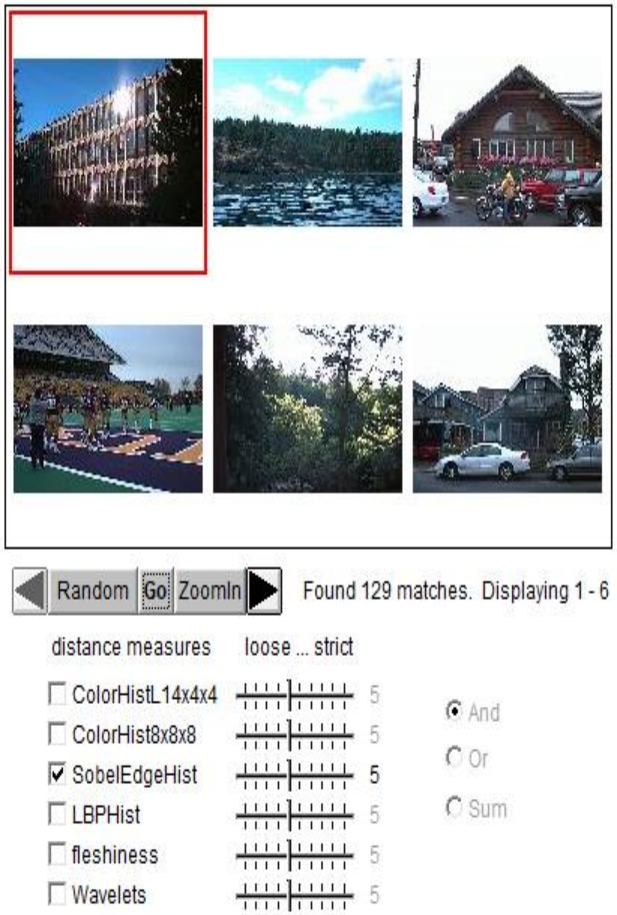


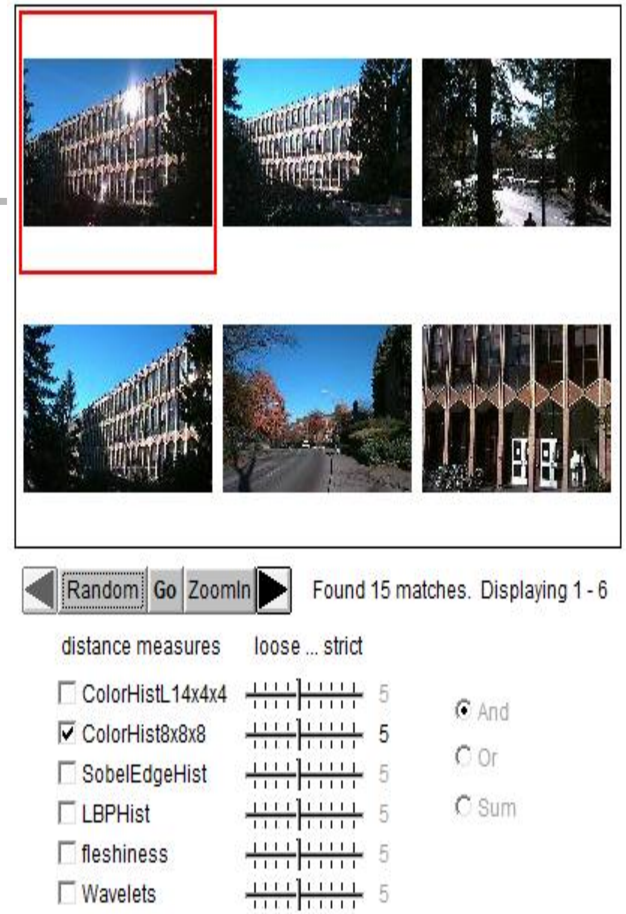
Different weights on features




Random
Go
ZoomIn
Found 170 matches. Displaying 1 - 6
distance measures
loose ... strict
ColorHistL14x4x4
5
ColorHist8x8x8
5
SobelEdgeHist
5
LBPHist
5
fleshiness
5
Wavelets
5
And
Or
Sum


Random
Go
ZoomIn
Found 170 matches. Displaying 1 - 6
distance measures
loose ... strict
ColorHistL14x4x4
5
ColorHist8x8x8
5
SobelEdgeHist
5
LBPHist
5
fleshiness
5
Wavelets
5
And
Or
Sum


Random
Go
ZoomIn
Found 129 matches. Displaying 1 - 6
distance measures
loose ... strict
ColorHistL14x4x4 5
ColorHist8x8x8 5
SobelEdgeHist 5
LBPHist 5
fleshiness 5
Wavelets 5
And
Or
Sum

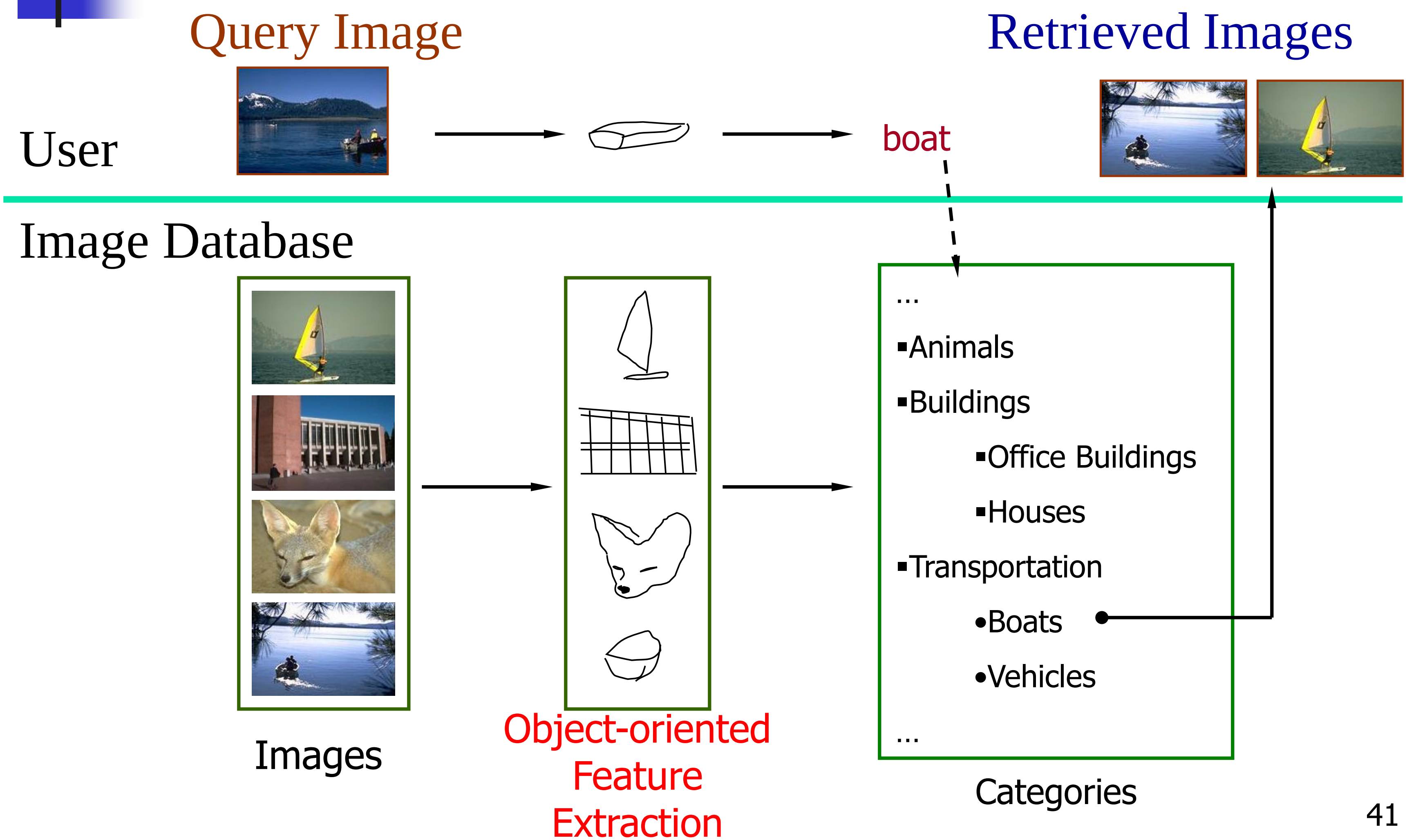

Random
Go
ZoomIn
Found 15 matches. Displaying 1 - 6
distance measures
loose ... strict
ColorHistL14x4x4 5
ColorHist8x8x8 5
SobelEdgeHist 5
LBPHist 5
fleshiness 5
Wavelets 5
And
Or
Sum

Weakness of Low-level Features

- Can’t capture the high-level concepts

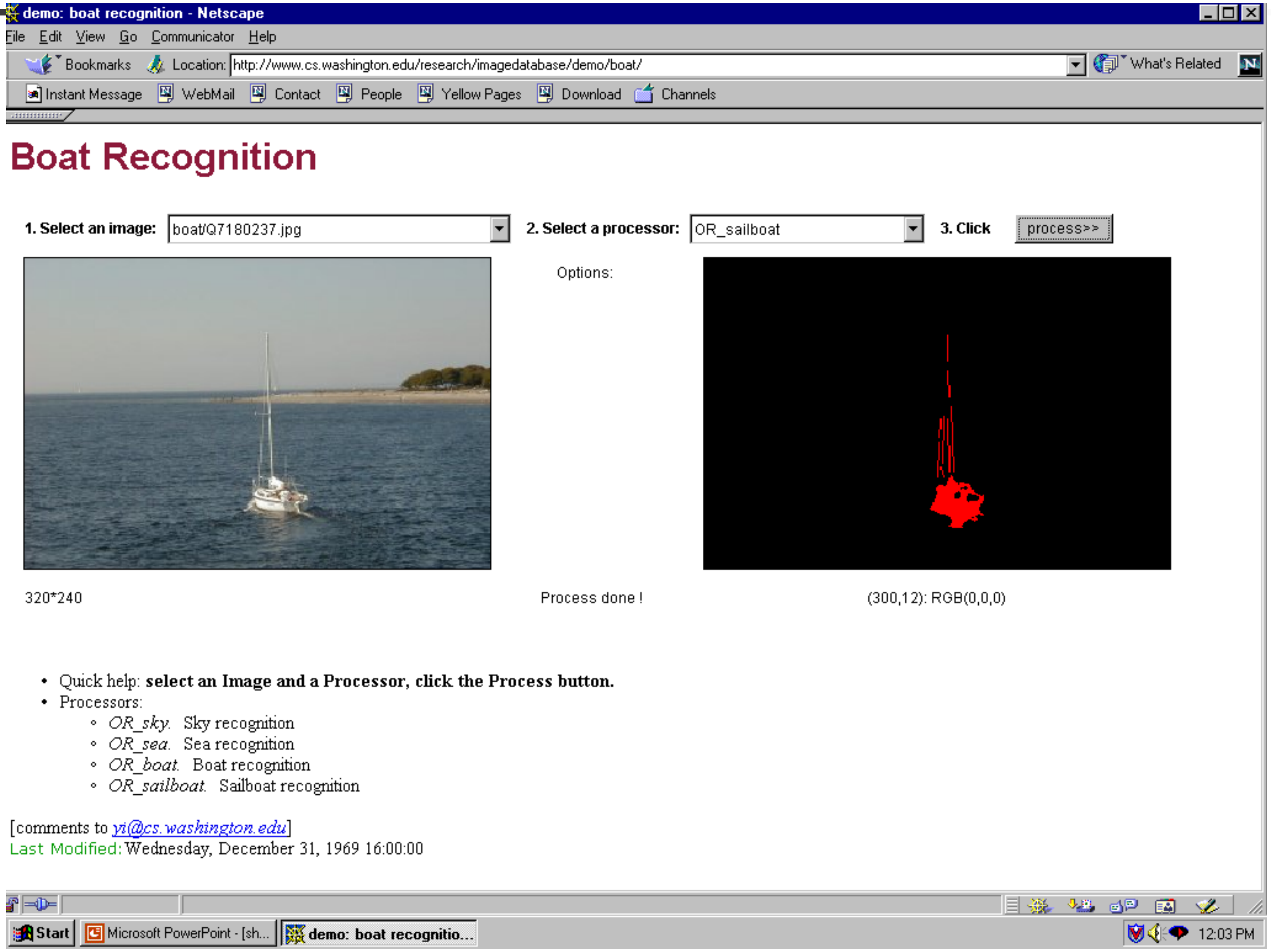
Research Objective: find objects



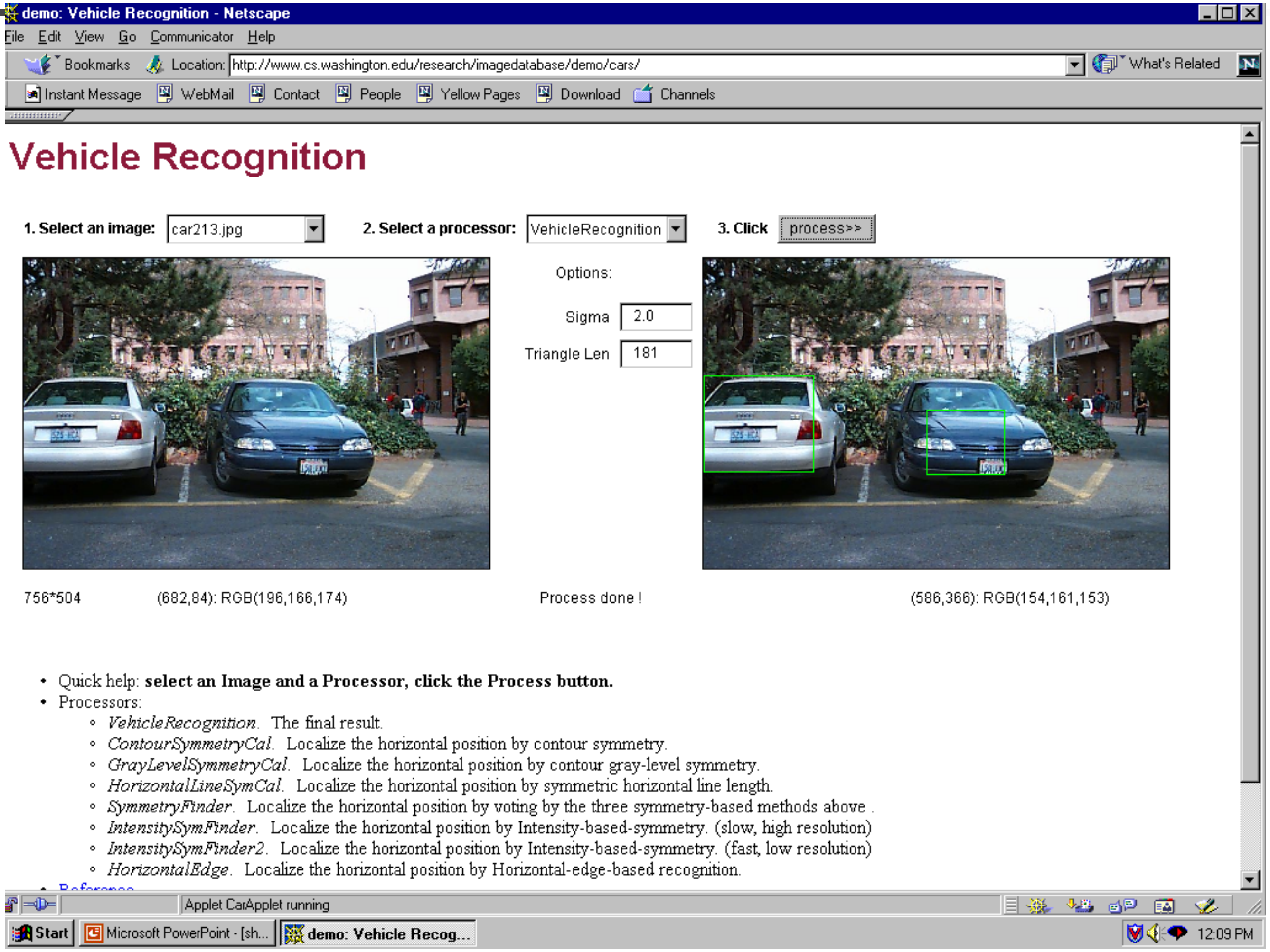
Overall Approach

- Develop object recognizers for common objects
- Use these recognizers to design a new set of both low- and mid-level features
- Design a learning system that can use these features to recognize classes of objects

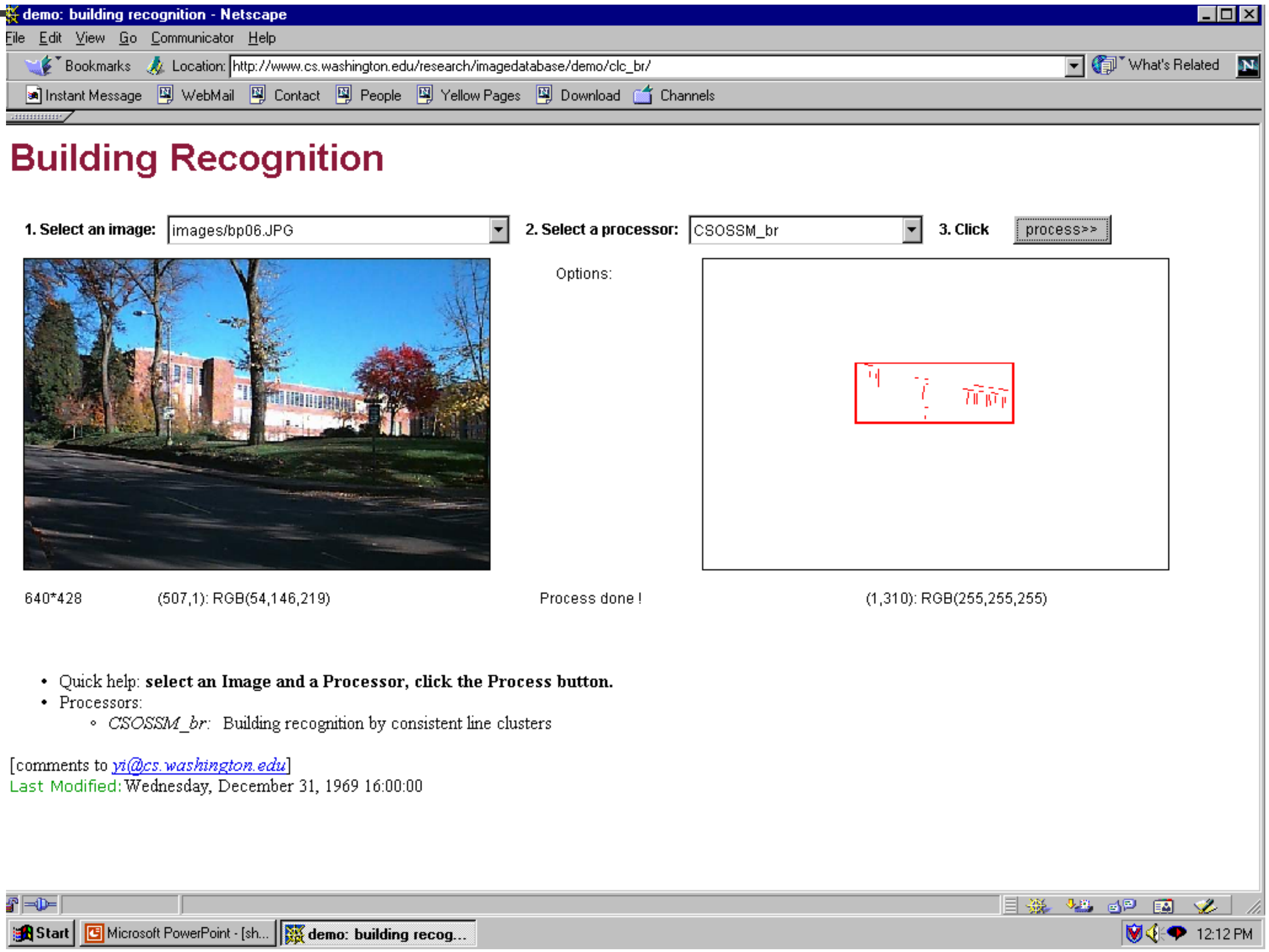
Boat Recognition



Vehicle Recognition



Building Recognition



Building Features: Consistent Line Clusters (CLC)

A **Consistent Line Cluster** is a set of lines that are homogeneous in terms of some line features.

- **Color-CLC**: The lines have the same color feature.
- **Orientation-CLC**: The lines are parallel to each other or converge to a common vanishing point.
- **Spatially-CLC**: The lines are in close proximity to each other.

Color-CLC

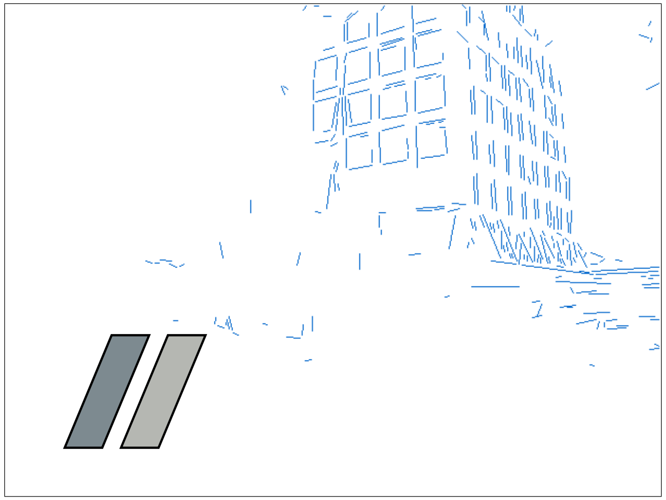
- Color feature of lines: color pair (c_1,c_2)
- Color pair space:

  RGB (256^3*256^3) Too big!

  Dominant colors (20*20)
- Finding the color pairs:

  One line → Several color pairs
- Constructing Color-CLC: use clustering

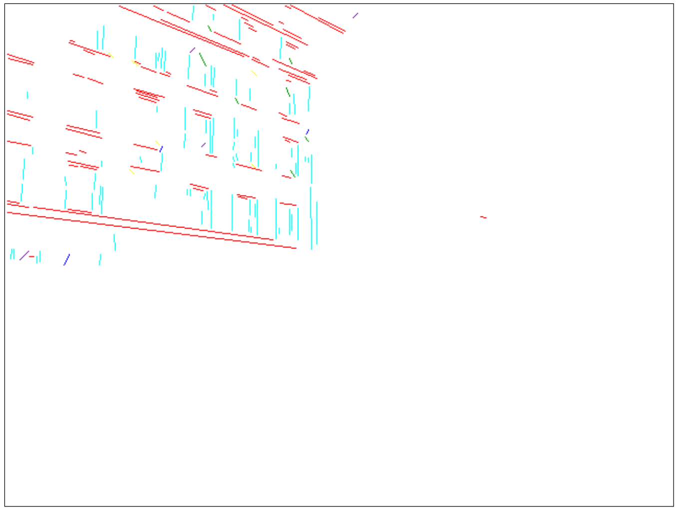
Color-CLC

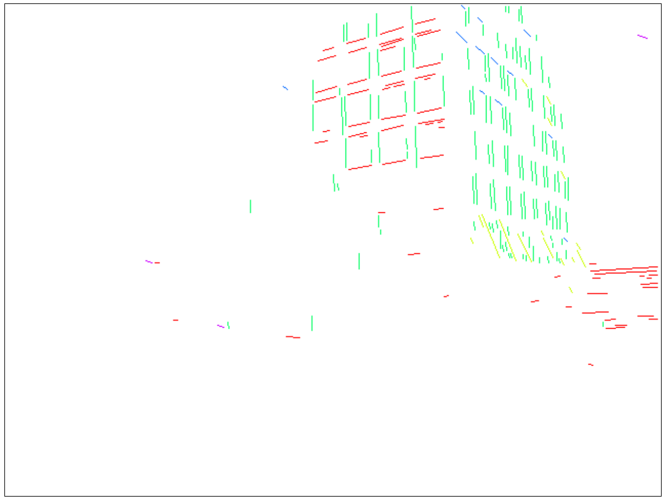


Orientation-CLC

- The lines in an Orientation-CLC are parallel to each other in the 3D world
- The parallel lines of an object in a 2D image can be:
  - Parallel in 2D
  - Converging to a vanishing point (perspective)

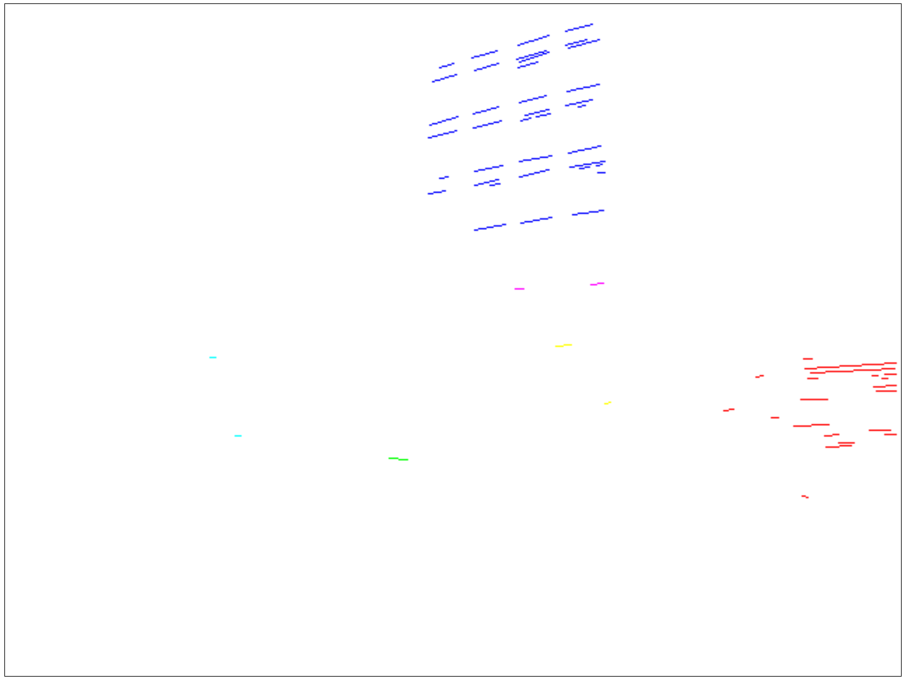
Orientation-CLC





Spatially-CLC

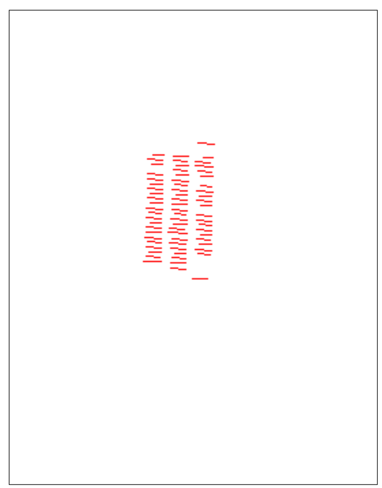
- Vertical position clustering
- Horizontal position clustering

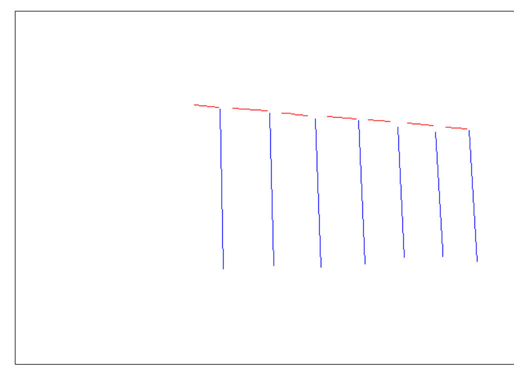


Building Recognition by CLC

Two types of buildings → Two criteria

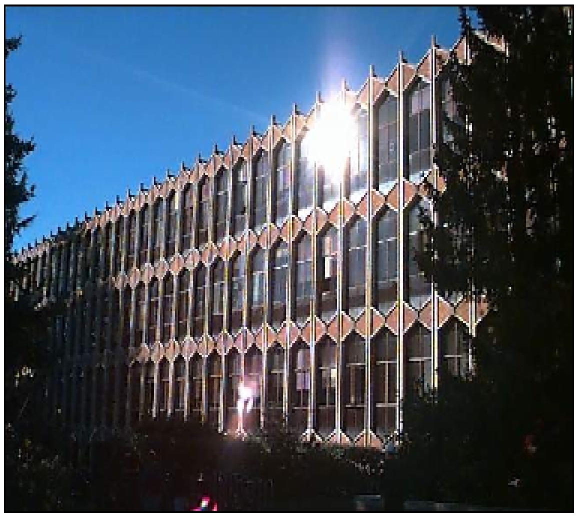
- Inter-relationship criterion
- Intra-relationship criterion



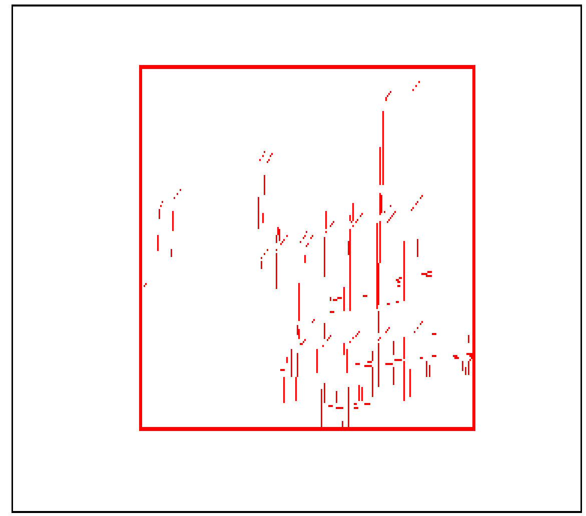


Building Recognition

1. Select an image: images/bp11.JPG 2. Select a processor: CSOSSM_br 3. Click process>>



Options:



640*428 (616,0): RGB(29,30,24) Process done ! (386,402): RGB(255,255,255)

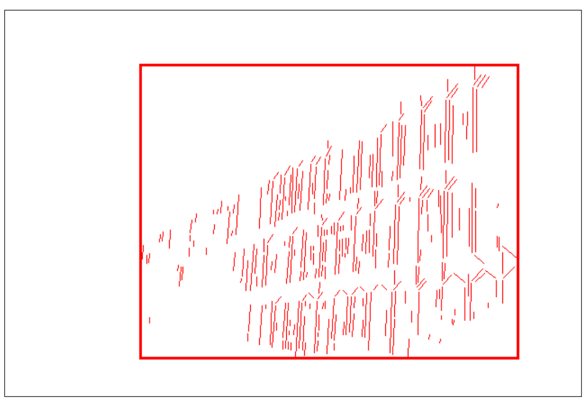
- Quick help: **select an Image and a Processor, click the Process button.**
- Processors:
  - *CSOSSM_br:* Building recognition by consistent line clusters

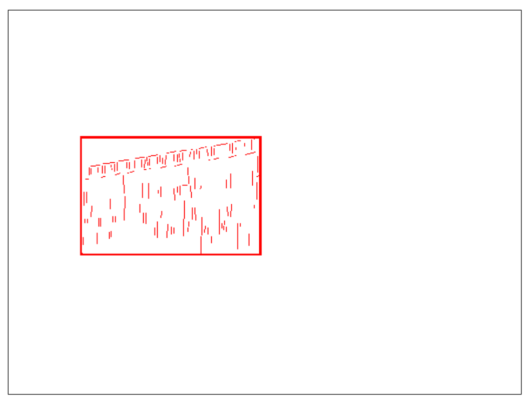
Experimental Evaluation

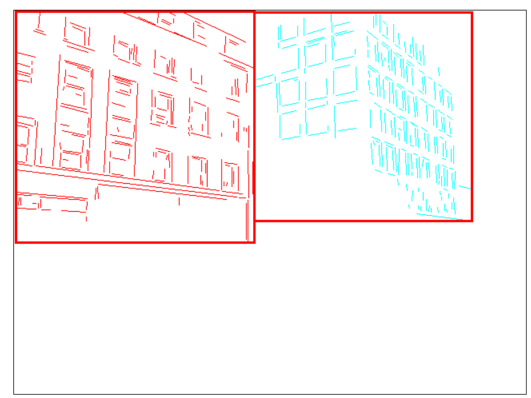
- Object Recognition
  - 97 well-patterned buildings (bp): 97/97
  - 44 not well-patterned buildings (bnp): 42/44
  - 16 not patterned non-buildings (nbnp): 15/16 (one false positive)
  - 25 patterned non-buildings (nbp): 0/25
- CBIR

Experimental Evaluation

Well-Patterned Buildings

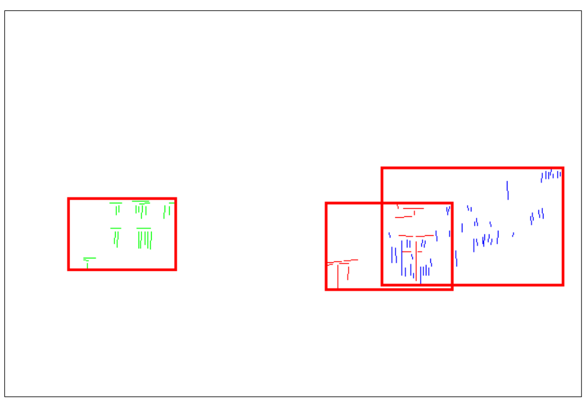


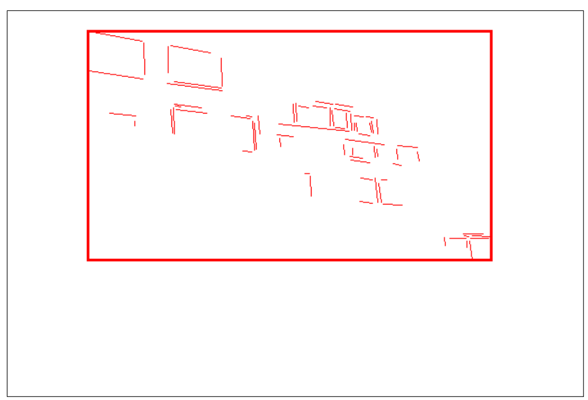




Experimental Evaluation

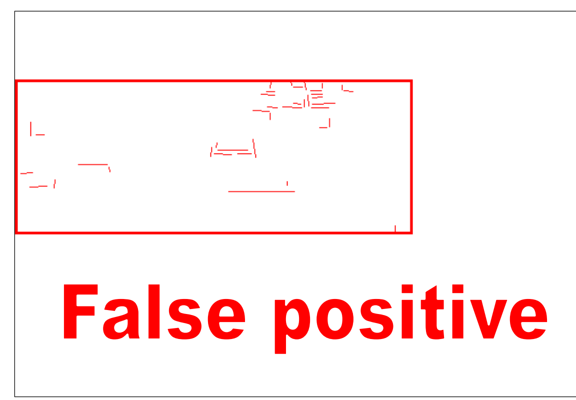
Non-Well-Patterned Buildings





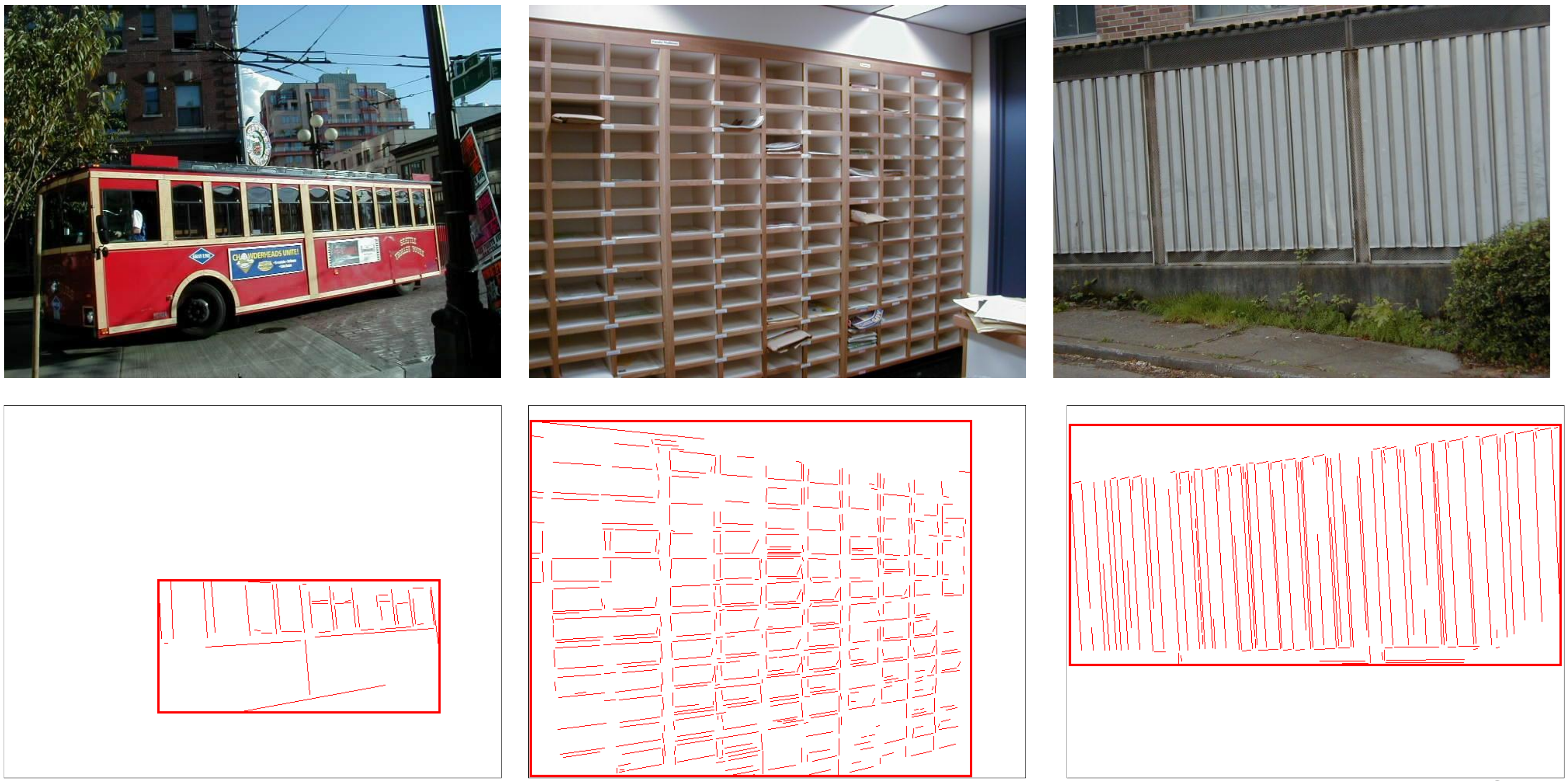
Experimental Evaluation

Non-Well-Patterned Non-Buildings



Experimental Evaluation

Well-Patterned Non-Buildings (false positives)



Experimental Evaluation (CBIR)

	Total Positive Classification (#)	Total Negative Classification (#)	False positive (#)	False negative (#)	Accuracy (%)
Arborgreens	0	47	0	0	100
Campusinfall	27	21	0	5	89.6
Cannonbeach	30	18	0	6	87.5
Yellowstone	4	44	4	0	91.7

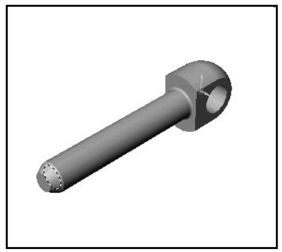
Experimental Evaluation (CBIR)

False positives from Yellowstone

3D Object Retrieval

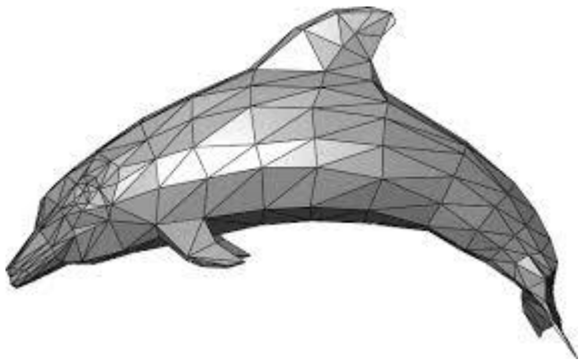
- Given a view of a 3D object
- Retrieve similar 3D objects
- From a database of 3D objects

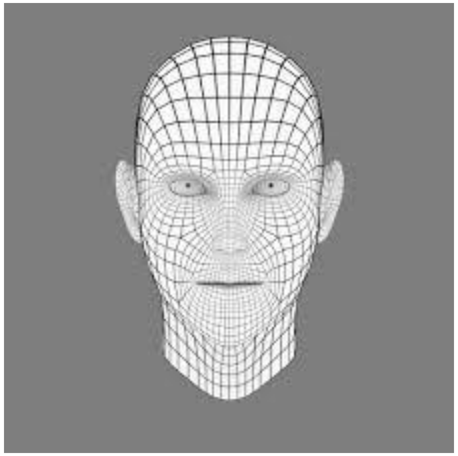
Work of Indriyati Atmosukarto

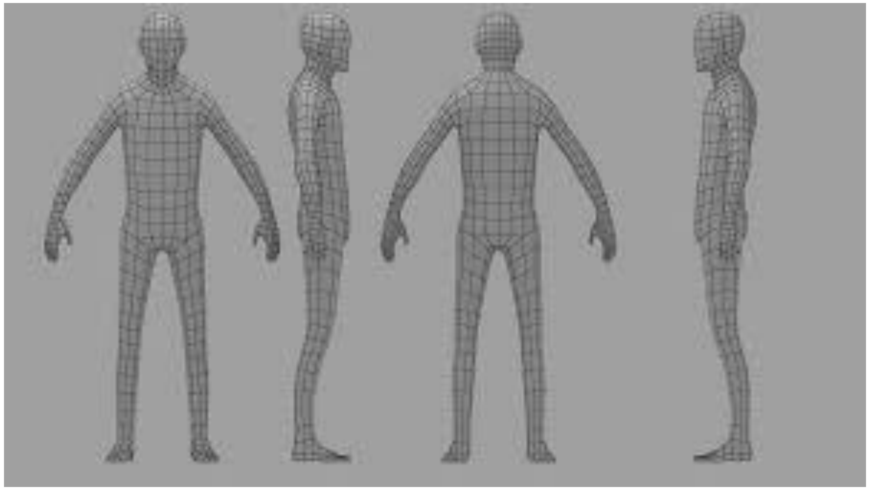



- Increasing number of 3D objects available
- Want to store, index, retrieve 3D objects automatically
- Need to create 3D object descriptor

Object Representation: 3D Mesh

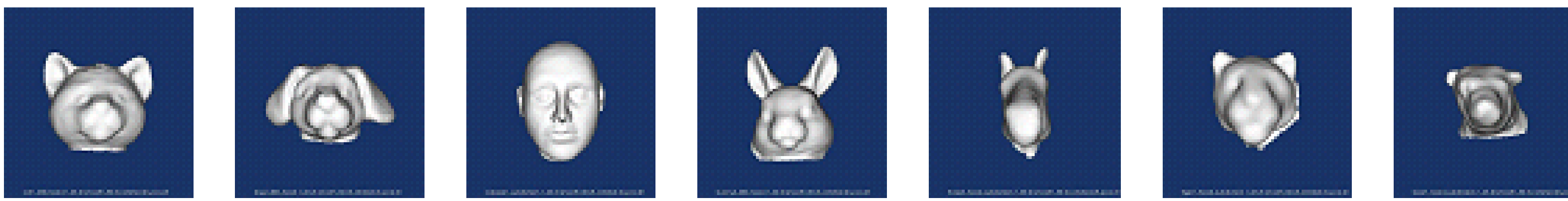






Datasets

- Heads: 375 objects; 7 classes



- SHREC 2008 : 425 objects; 39 classes

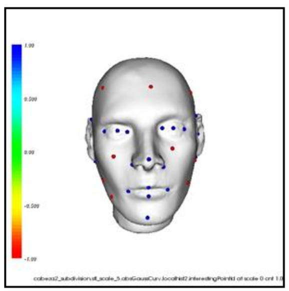
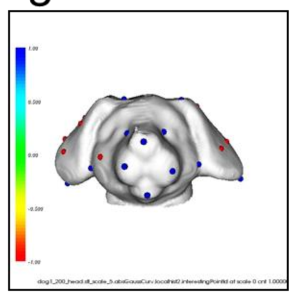
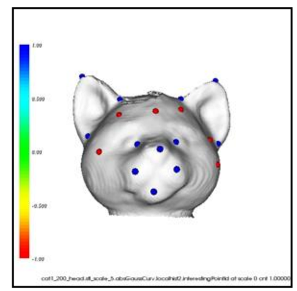
Our first retrieval measure

- Learn to find salient points of the objects
- Use those points to compute a 2D signature in the form of a longitude/latitude map
- Match the maps for retrieval

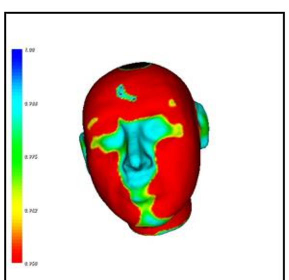
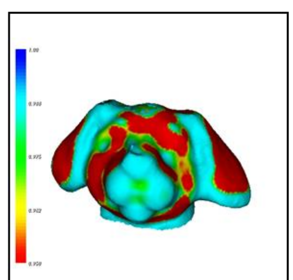
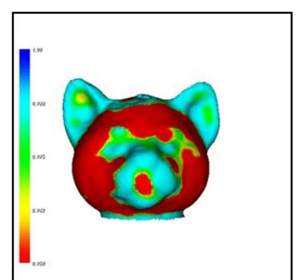
Learning Salient Points

Classifier: SVM. It learns the characteristics of salient points.

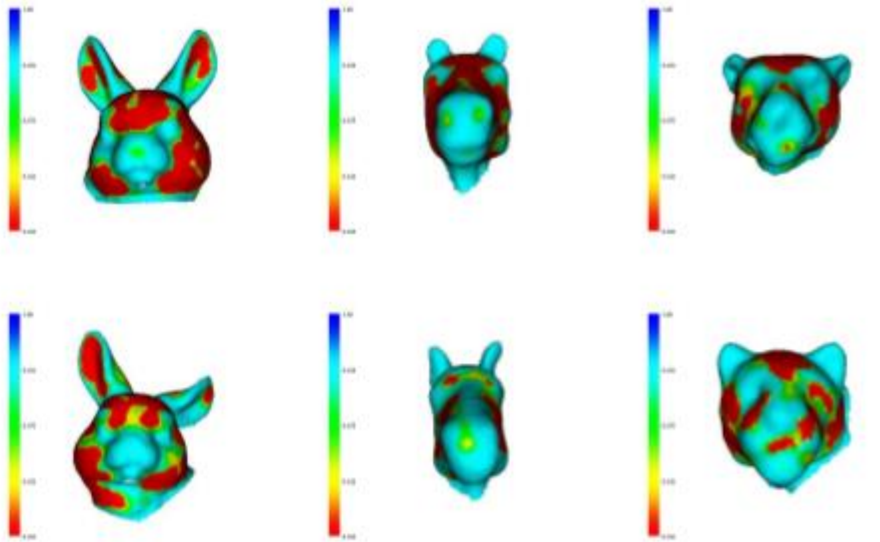
Training: We manually mark salient and non-salient points on a subset of the Heads object. An average of 12 salient points and 12 non-salient points. Histogram of low-level features of each marked points were used for training.

Testing: After training, the classifier is able to label each point of any 3D object as either salient or non-salient and provides a confidence score for its decision.

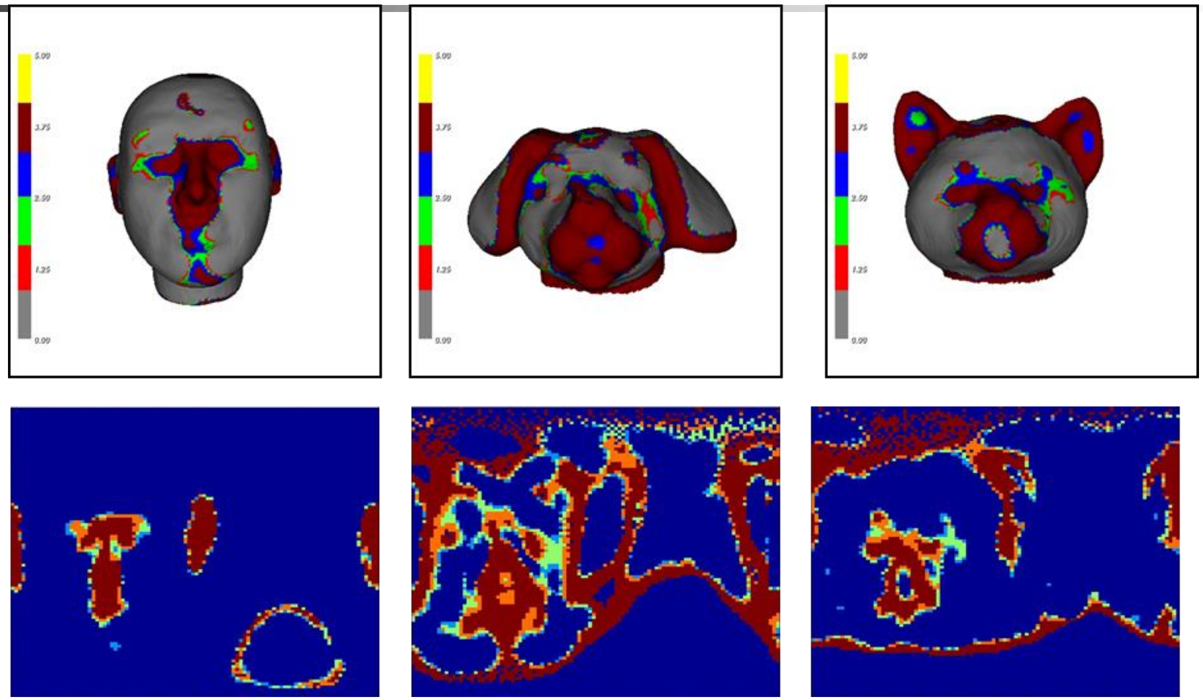
Salient Point Prediction for Heads



(a) (b) (c)

Figure 14: Salient point prediction for (a) rabbit head class, (b) horse head class, and (c) leopard head class from the Heads database. Even though all three classes were not included in the training, the training model was able to predict salient points across the classes.

2D Longitude-Latitude Map Signature



Idea: stretch on the 3D object with the pattern on it to make a 2D map, like making a globe of the world into a 2D map.

Head Retrieval

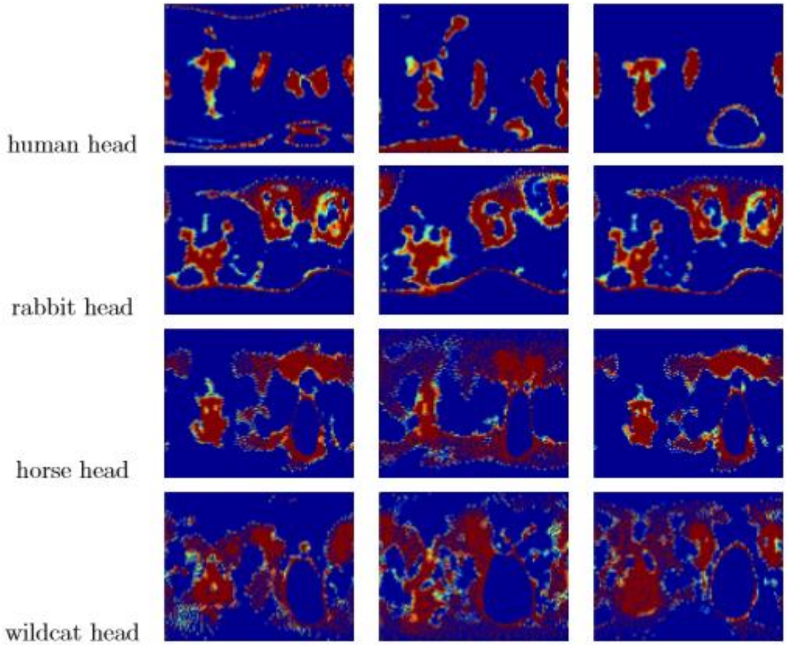
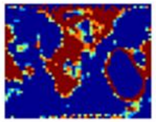
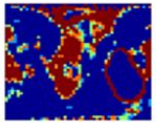
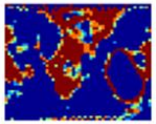
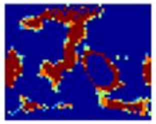
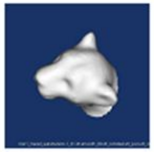
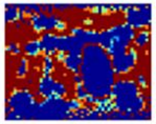


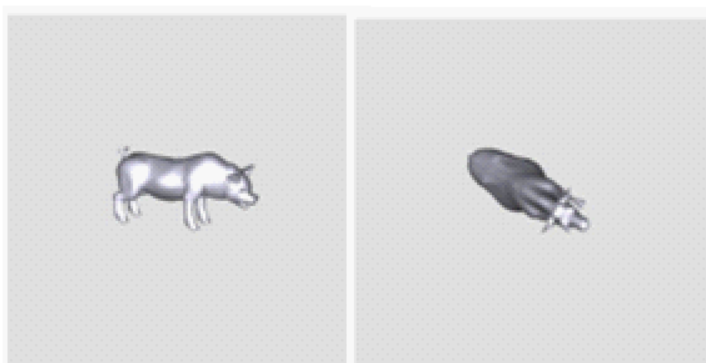
Figure 17: Objects that are similar and belong to the same class will have similar 2D longitude-latitude signature maps.

Rotation-Invariant Retrieval

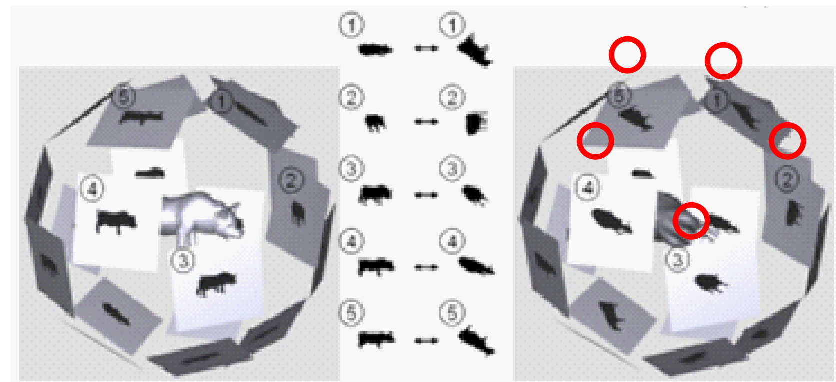
Info	Name	3D Object	Signature
Rank= 1 Dist= 84.33	rabbit_1_30 rotated θ_x=0 θ_y=300 θ_z=300		
Rank= 2 Dist= 84.34	rabbit_1_27 rotated θ_x=0 θ_y=300 θ_z=300		
Rank= 3 Dist 84.59	rabbit_1_26 rotated θ_x=0 θ_y=300 θ_z=300		
...	...	...	...
Rank= 15 Dist= 94.91	horse_1_28 rotated θ_x=0 θ_y=300 θ_z=300		
...	...	...	...
Rank= 105 Dist= 114.9	lion_31 rotated θ_x=0 θ_y=300 θ_z=300		

Related Work in SHREC Competition

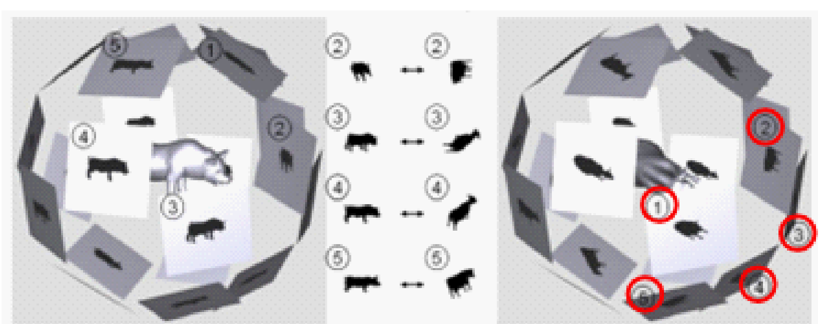
- Light Field Descriptor [Chen et al., 2003]



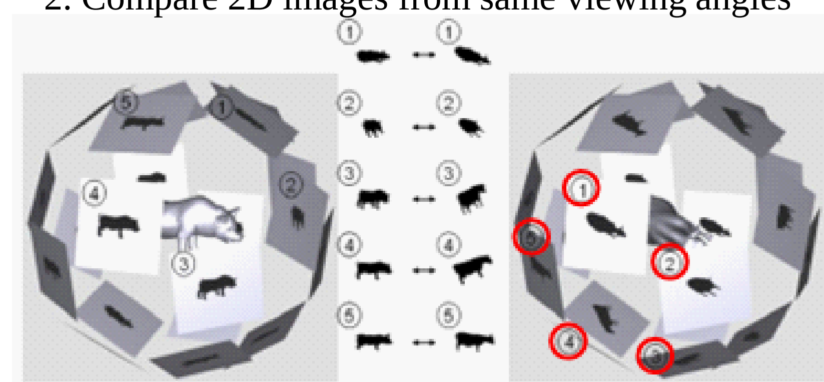
1. Given two 3D models rotated randomly



2. Compare 2D images from same viewing angles



3. Compare 2D images from another angle



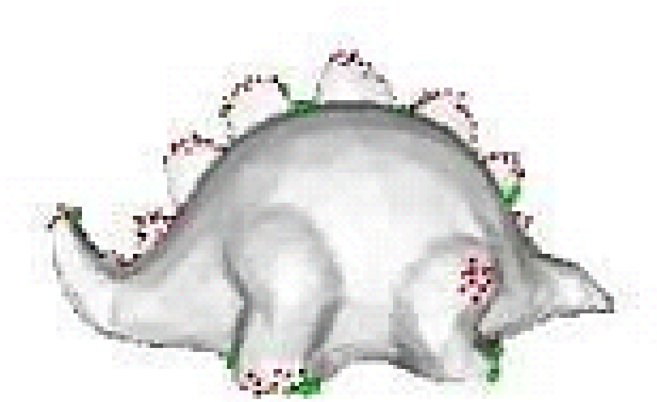
4.Best match = Rotation of camera position with best similarity

New objective of our work: join them and beat them

- Select 2D salient views (instead of all views) to describe 3D object
  - Learn salient points
  - Select a subset of 2D salient views
- Retrieval using view-based similarity measure
  - Use the subset of views so faster than theirs

Selecting Salient Views

- Improve LFD by selecting salient views
- Salient views are discernible and most useful in describing 3D object
- Salient points appear on contour of object
  - Surface normal vector $\perp$ camera view point



Experimental Results

- Comparison to LFD per class

No	Class	# Objects	Avg # distinct salient views	Max distinct salient views score	LFD score
1	human-diff-pose	15	12.33	0.113	0.087
2	monster	11	12.14	0.196	0.169
3	dinosaur	6	12.33	0.185	0.169
4	4-legged-animal	25	12.24	0.274	0.186
5	hourglass	2	11.50	0.005	0.001
6	chess-pieces	7	12.14	0.085	0.085
7	statues-1	19	12.16	0.267	0.250
8	statues-2	1	13.00	0.000	0.000
9	bed-post	2	12.00	0.124	0.008
10	statues-3	1	12.00	0.000	0.000

- Average score:
  - 0.121 (DSV) vs 0.098 (LFD)

Conclusion

- Salient 2D views to speed up LFD
- Similar performance to LFD while rendering fewer views
  - LFD: 100 views
  - Our method DSV: 12 views (10%)
- Achieve 15-fold speed up in feature extraction time