Applications

Presented by

Sherin Aly

Most slides Courtesy to the papers' authors

Photo Tourism:
Exploring Photo Collections in 3D

Noah Snavely
Steven M. Seitz
University of Washington
Richard Szeliski
Microsoft Research

15,464

37,383

76,389

Google Demo

The goal of Photo Tourism is to use location information to build a better visualization and photo exploration tool for collections of photos of the same scene.

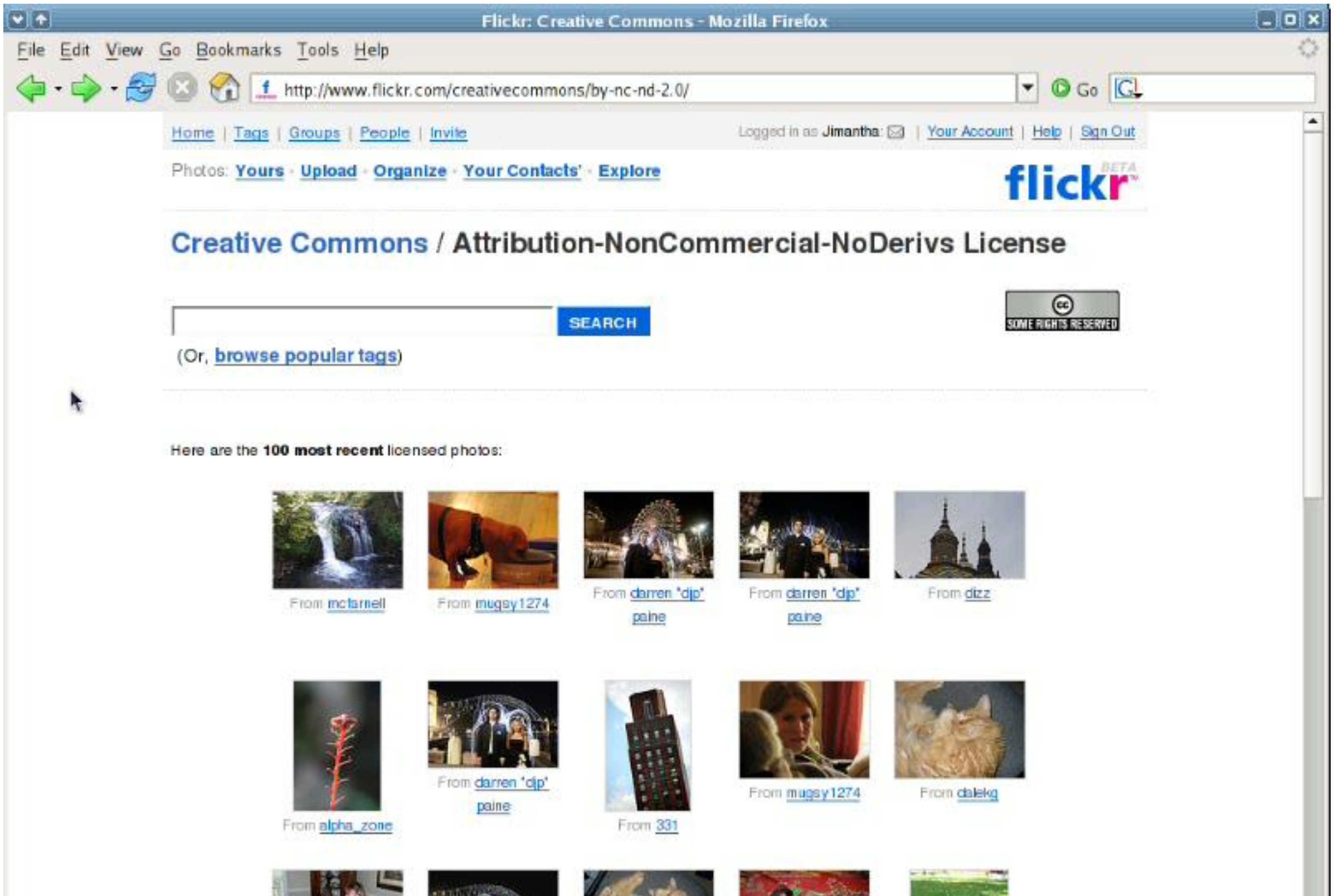
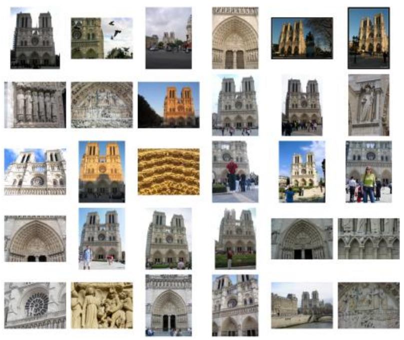


Photo Tourism

- A demo
- The system takes the set of photos and automatically determines the relative positions and orientations from which each photo was taken.
- We can then load the photos into the immersive 3D browser where the user can visualize and explore the photos using spatial relationships.

Photo Tourism overview



Input photographs
-Internet
-Personal collection

Scene geometry reconstruction

- Relative camera positions and orientations
- Point cloud for scene geometry

Sparse correspondence

Photo Explorer

Related work

- Image-based modeling (recover scene geometry)

Debevec, *et al.*
SIGGRAPH 1996

Schaffalitzky and Zisserman
ECCV 2002

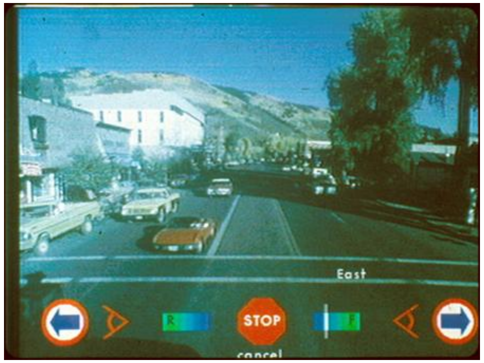
Brown and Lowe
3DIM 2005

automatic structure from motion on unordered sets of images

However, this modeling system is the first to be successfully applied to hundreds of images taken from the Internet.

Related work

- Image-based rendering (depict transitions between images)



Aspen Movie Map
Lippman, *et al.*, 1978

Photorealistic IBR:
Levoy and Hanrahan, SIGGRAPH 1996
Gortler, *et al*, SIGGRAPH 1996
Seitz and Dyer, SIGGRAPH 1996
Aliaga, *et al*, SIGGRAPH 2001
and many others

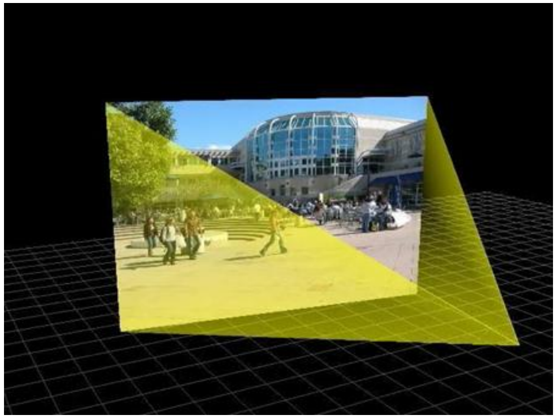
They created an interactive virtual tour of the city of Aspen,

Same exploration features, but which required extensive manual effort to create.

Related work

- Image browsing and image retreival

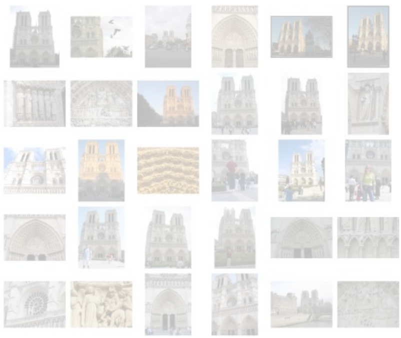
WWMX
Toyama, *et al*,
Int. Conf. Multimedia, 2003
Use location information to organize photos,
No immersive browsing experience



The Realityflythrough project
McCurdy and Griswold
Mobisys 2003
Use GPS to locate images in space
PhotoTourism system don’t use GPS

Video Google
Sivic and Zisserman
ICCV 2003
It allow users to find objects in videos by selecting them.
PT:“3D photo browser” context

Photo Tourism overview



Input photographs

Scene reconstruction

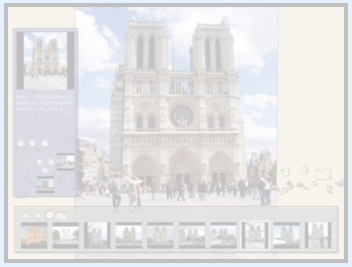
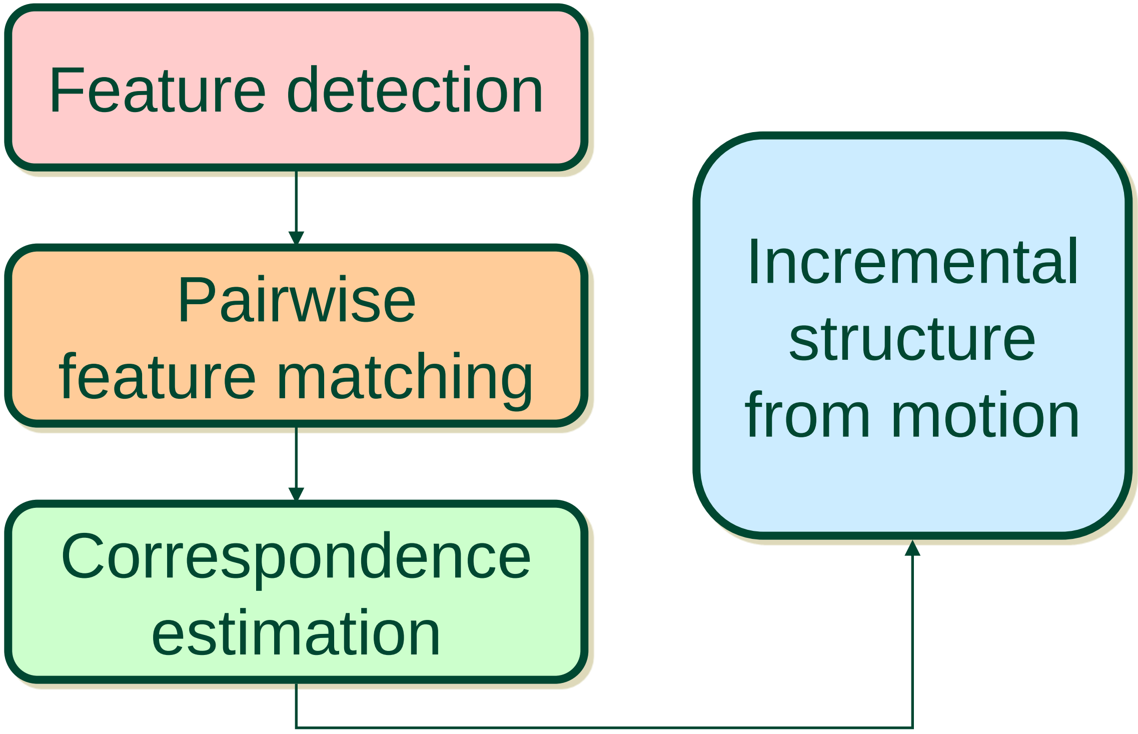


Photo Explorer

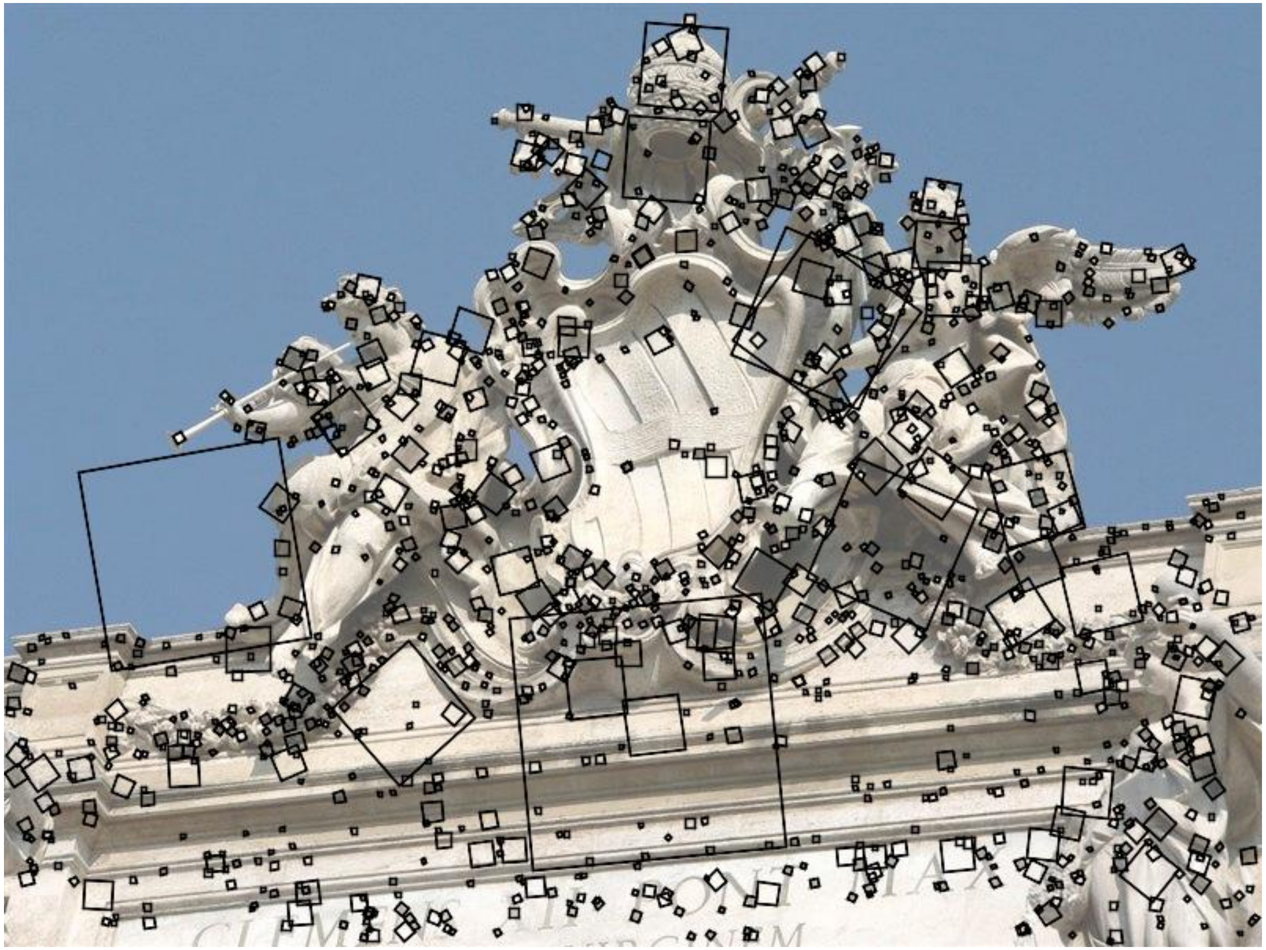
Scene reconstruction

- Automatically estimate
  - position, orientation, and focal length of cameras (i.e. zoom)
  - 3D positions of feature points



Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



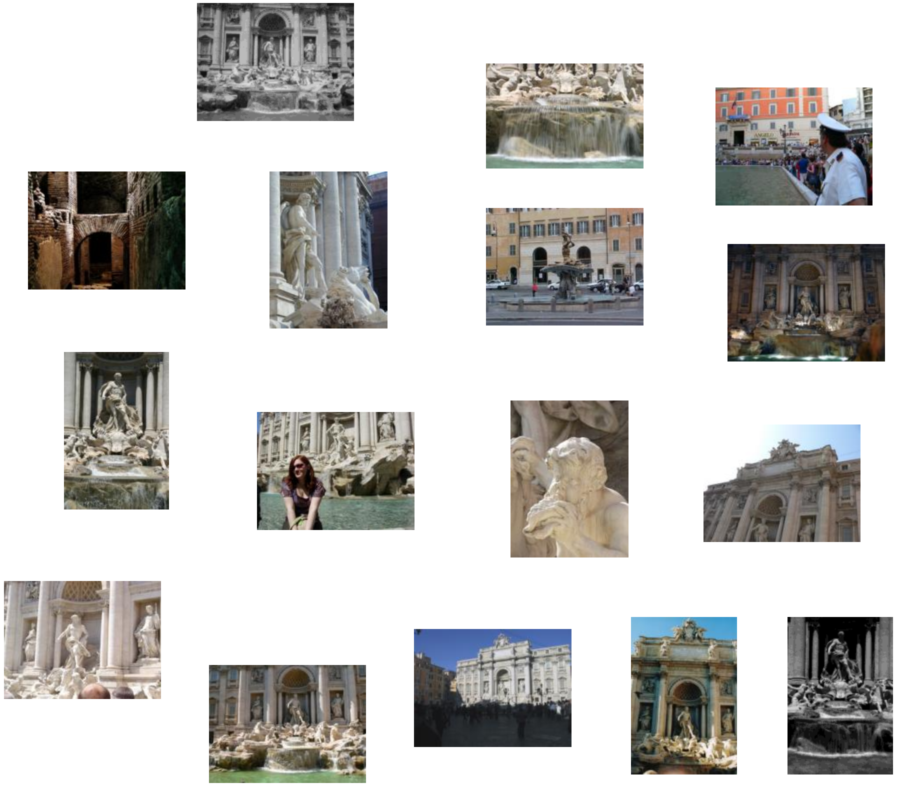
invariant to
-scale
-Rotation
- affine changes in image intensity.

squares are scaled and rotated to reflect the scale and orientation of the features

The Trevi Fountain

Feature detection

Detect features using SIFT [Lowe, IJCV 2004]

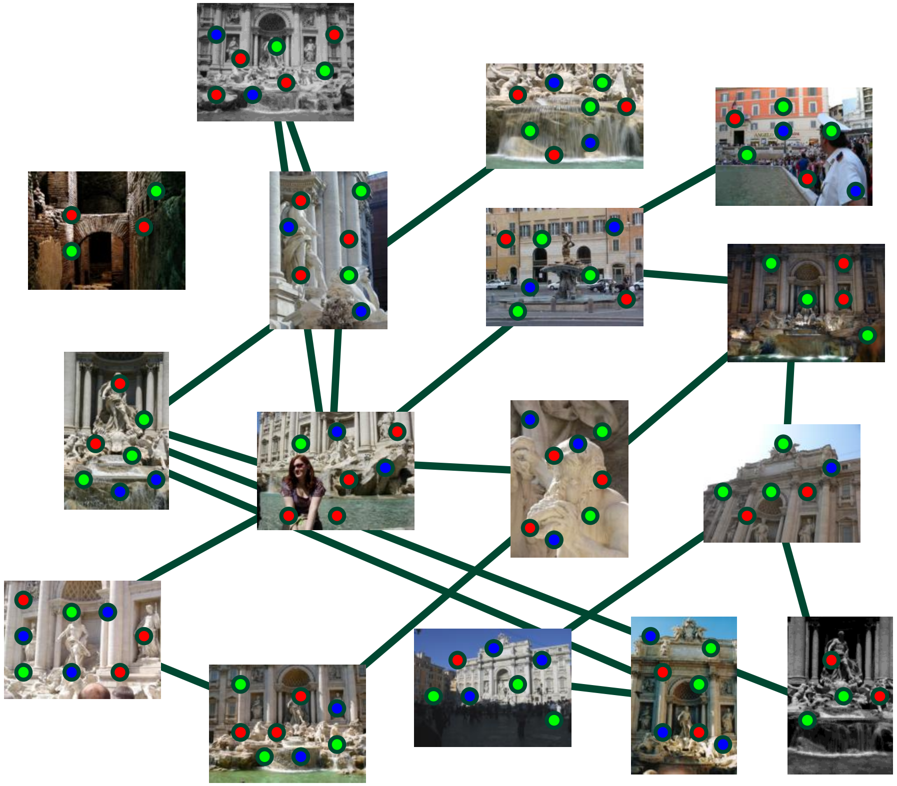


Feature detection

Detect features using SIFT [Lowe, IJCV 2004]

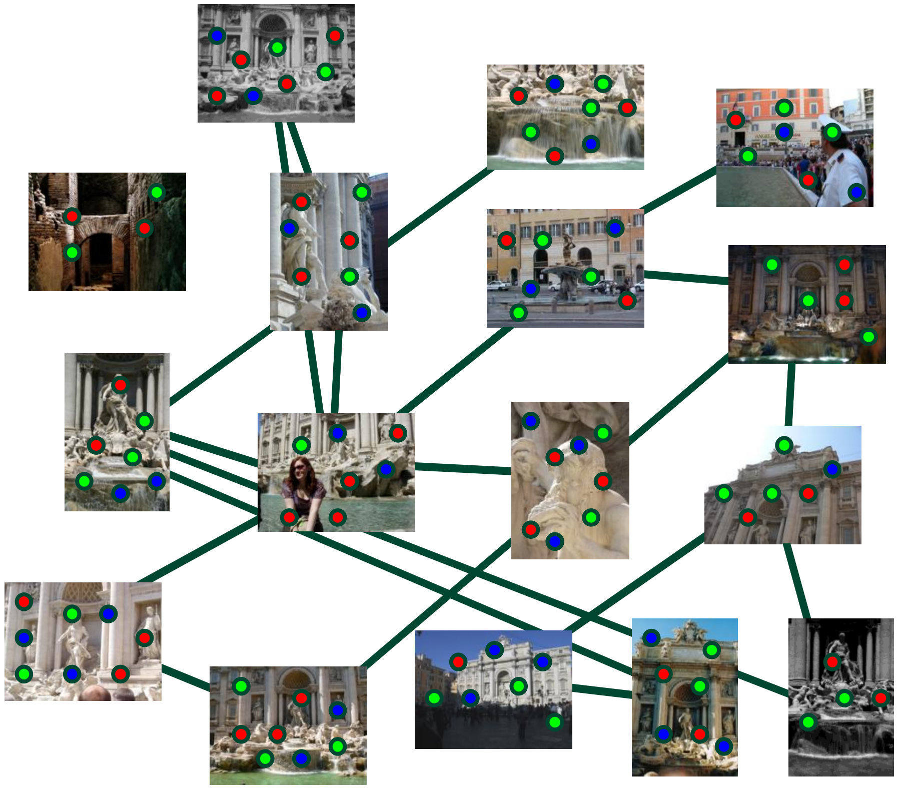
Feature matching

Match features between each pair of images

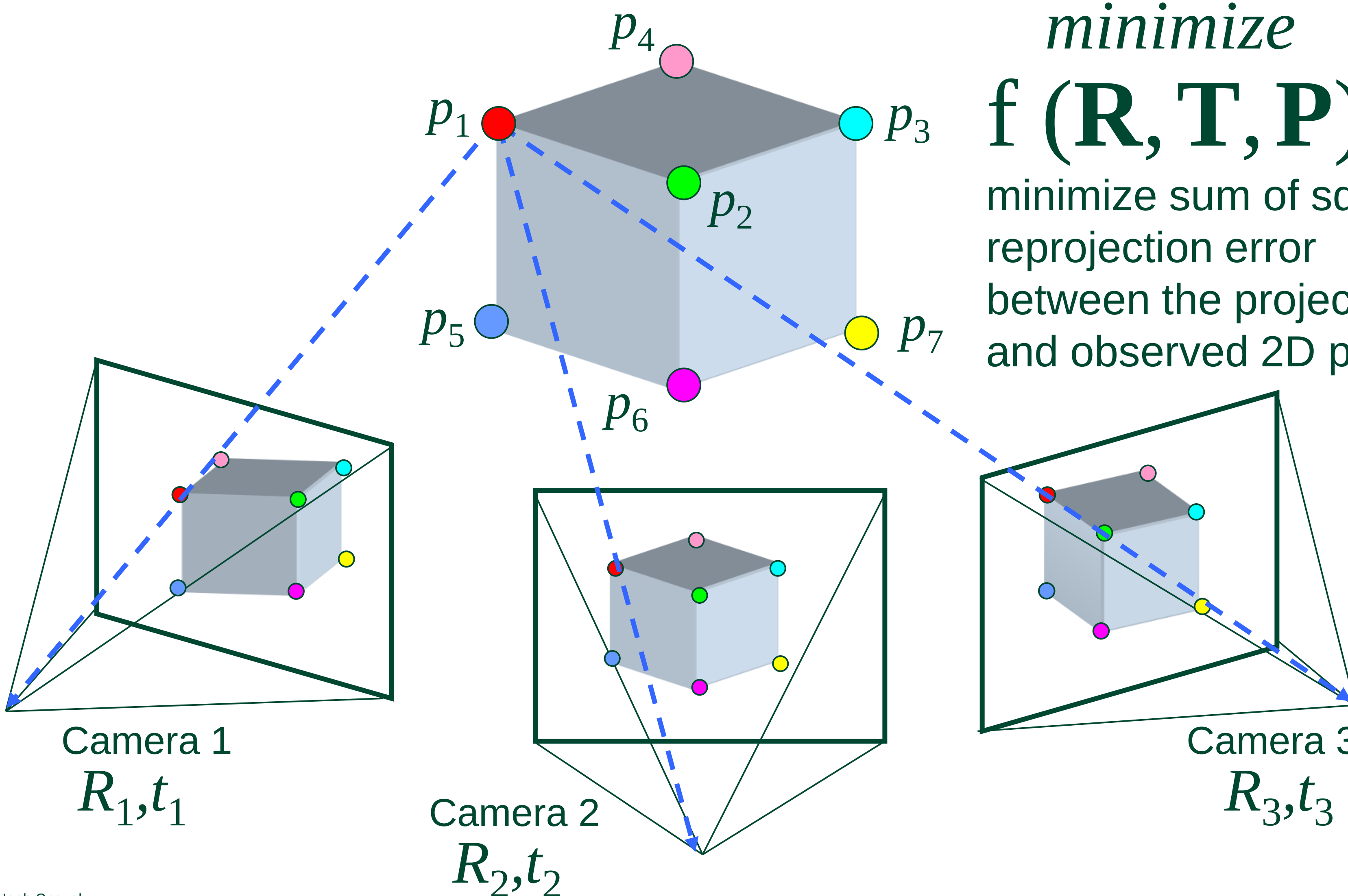


Feature matching

Refine matching using RANSAC [Fischler & Bolles 1987] to estimate fundamental matrices between pairs and then keep only matches consistent with that fundamental matrix.



Structure from motion



minimize

$$f\,(\mathbf{R},\mathbf{T},\mathbf{P})$$

minimize sum of squared reprojection error between the projected and observed 2D points

Structure from motion (Cont.)

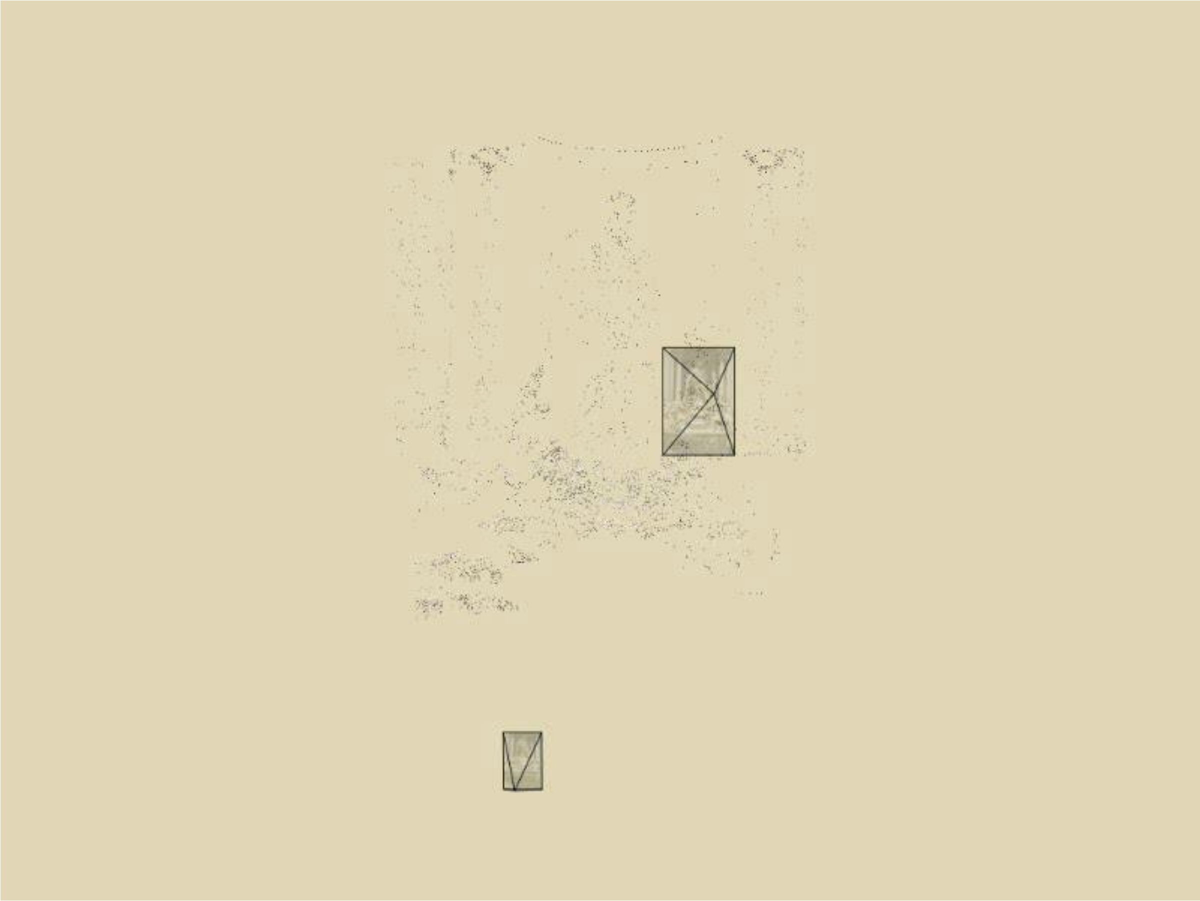
- This is a non-linear least squares problem and can be solved with algorithms such as Levenberg-Marquart.
- However, because the problem is non-linear, it can be sensitive to local minima.
- Therefore, it's important to initialize the parameters of the system carefully.
- In addition, we need to be able to deal with erroneous correspondences.

Incremental structure from motion

reconstruct the scene incrementally

Incremental structure from motion

Incremental structure from motion
Demo 1



Incremental structure from motion Demo 2

repeat until no more photos match any points in the scene

Reconstruction performance

- For photo sets from the Internet, 20% to 75% of the photos were registered
- Most unregistered photos belonged to different connected components (e.g. when searching for the Notre Dame cathedral, you get back photos of both the interior and exterior)
- Some failure cases (noisy,dark, too low resolution, too different angel than others)

- Running time: < 1 hour for 80 photos

  > 1 week for 2600 photo

Photo Tourism overview

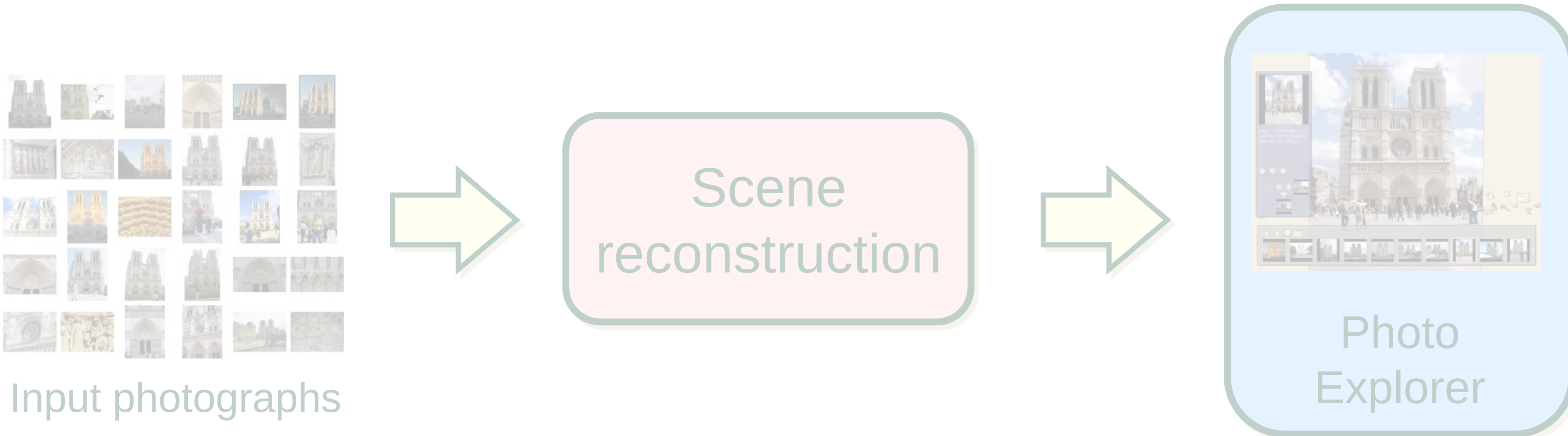
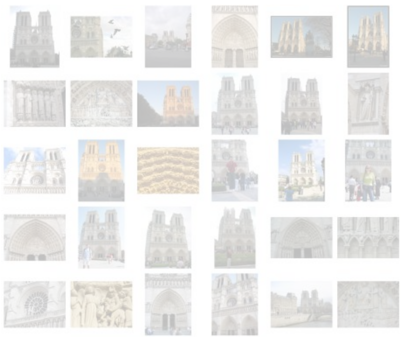


Photo Tourism overview



Input photographs

Scene reconstruction

Photo Explorer

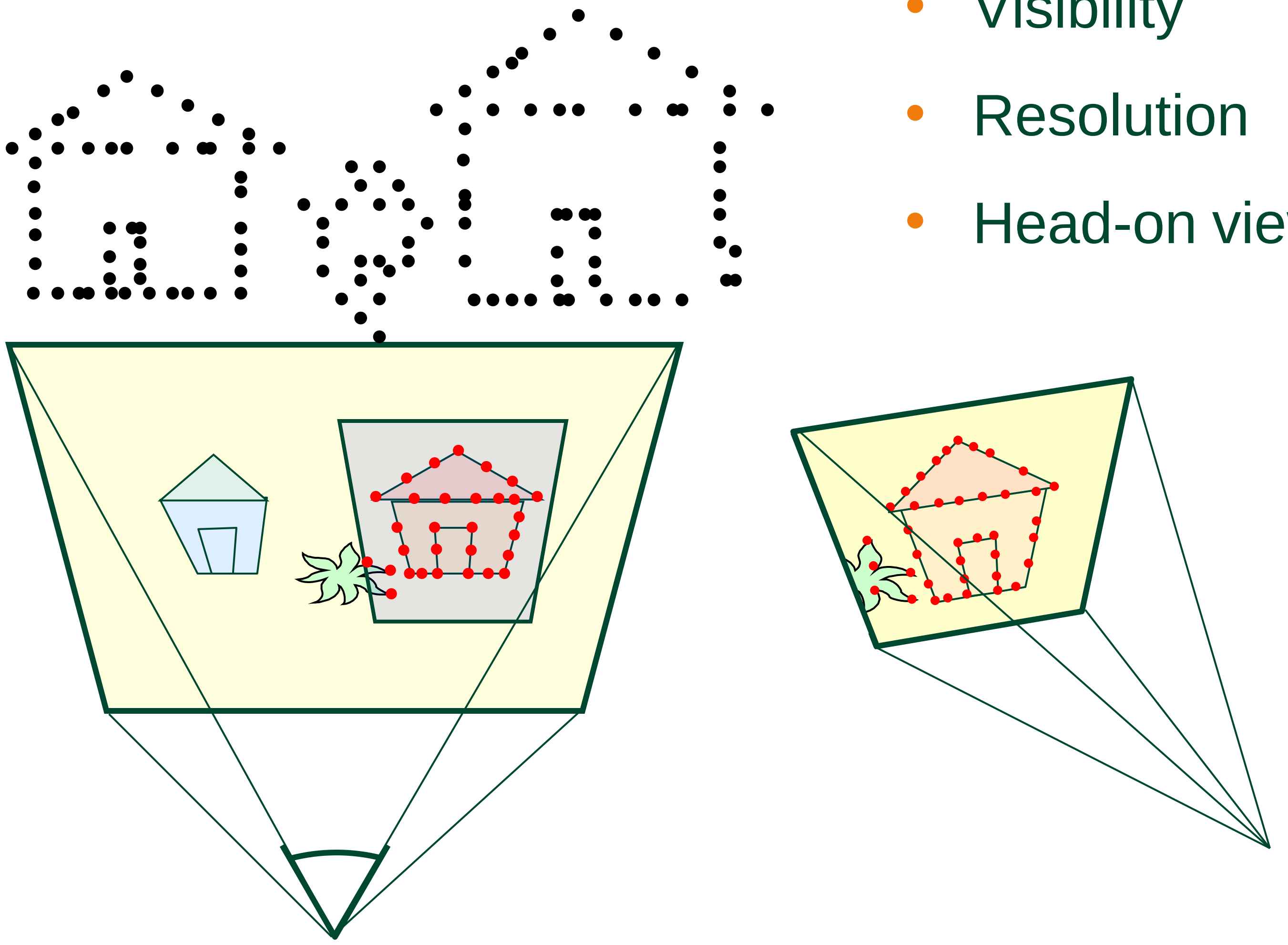
- Navigation
- Rendering
- Annotations

Navigation controls

- Free-flight navigation
- Object-based browsing
- Relation-based browsing
- Overhead map

Object-based browsing

Object-based browsing



- Visibility
- Resolution
- Head-on view or oblique

Relation-based browsing

Find all similar images

Find all details

Find all zoom outs

Zoom in

Move left

Move right

Zoom out

Relation-based browsing

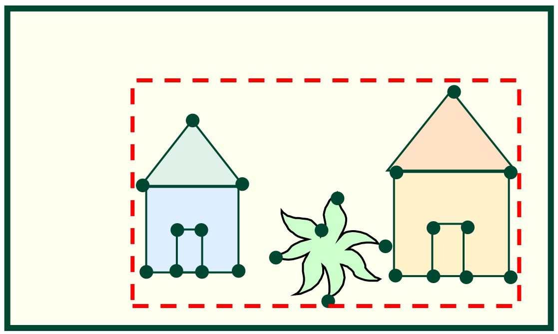


Image A

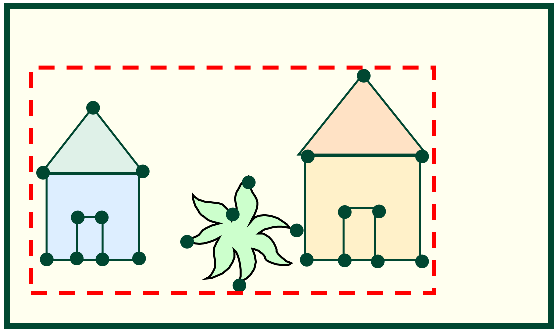


Image B

These relations are inferred based on the relative positions of corresponding feature points between photographs.

Relation-based browsing

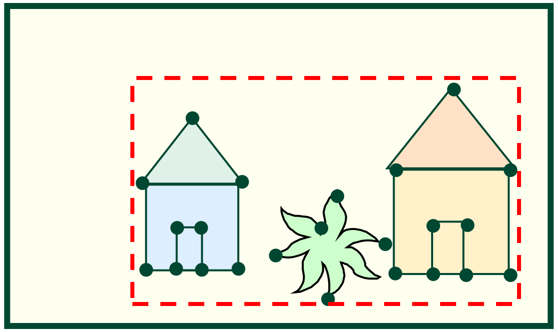


Image A

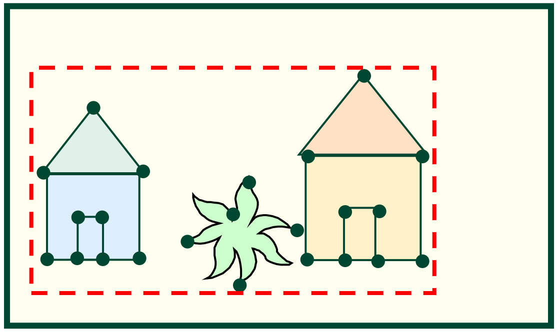
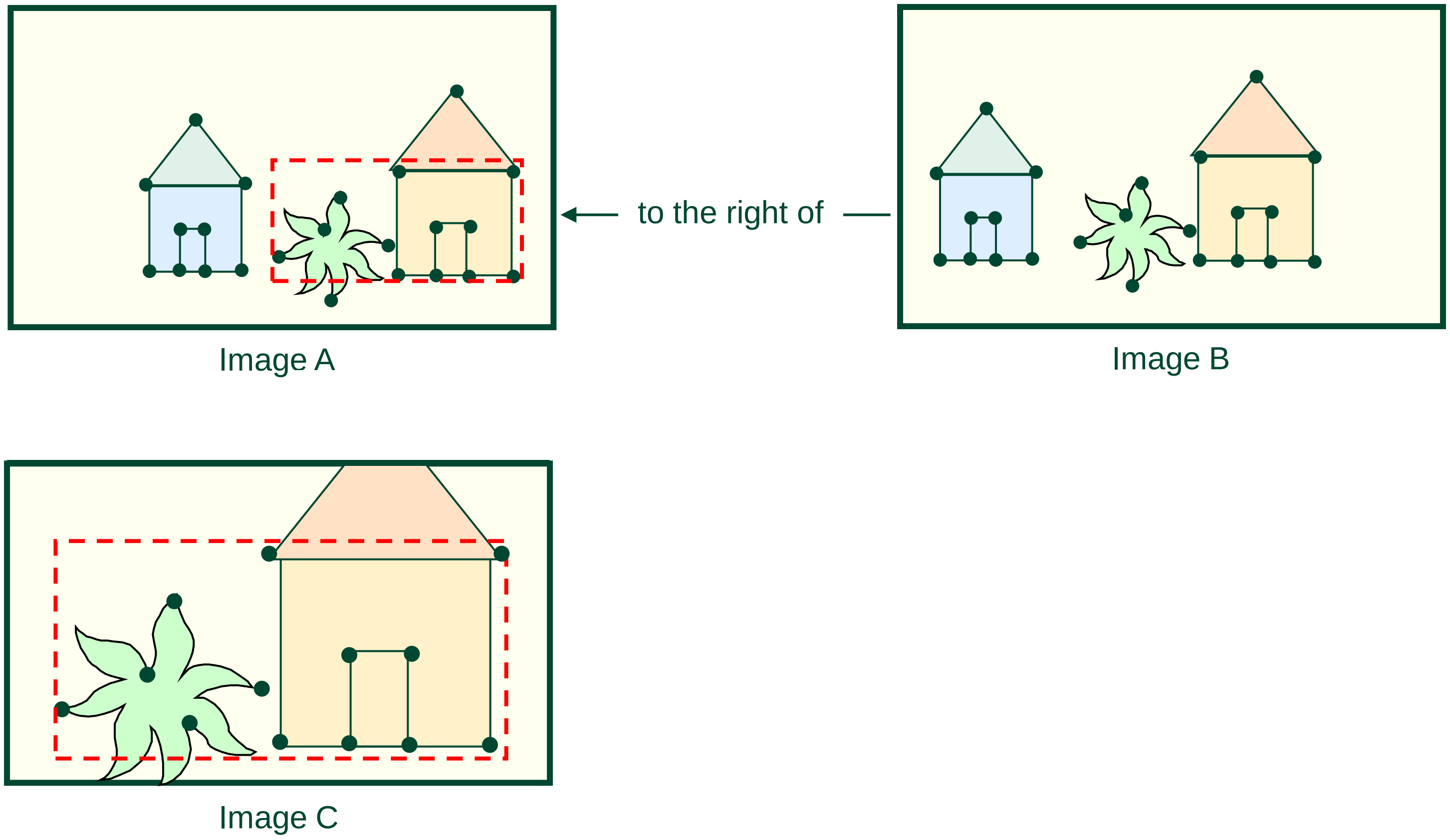
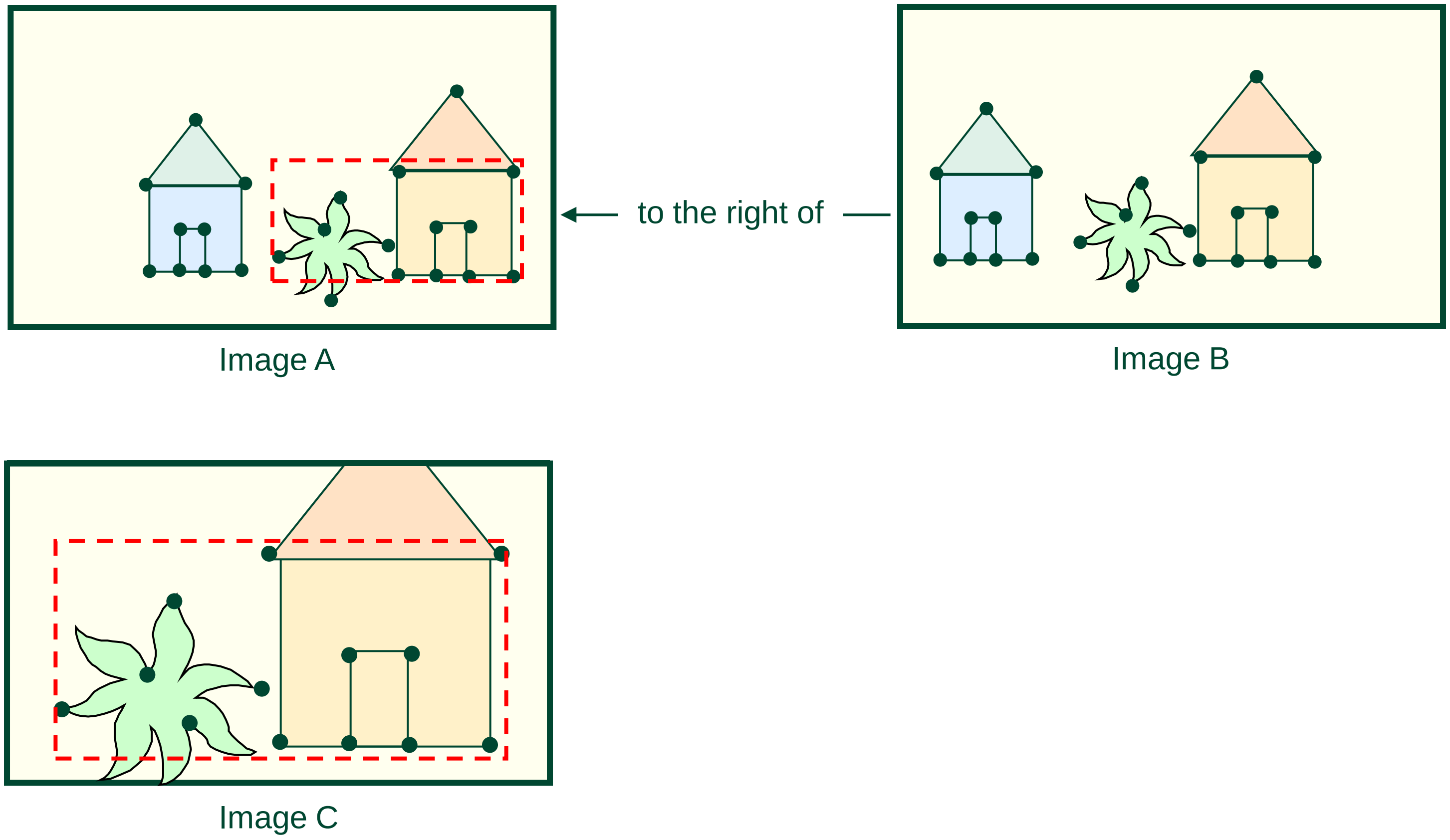


Image B

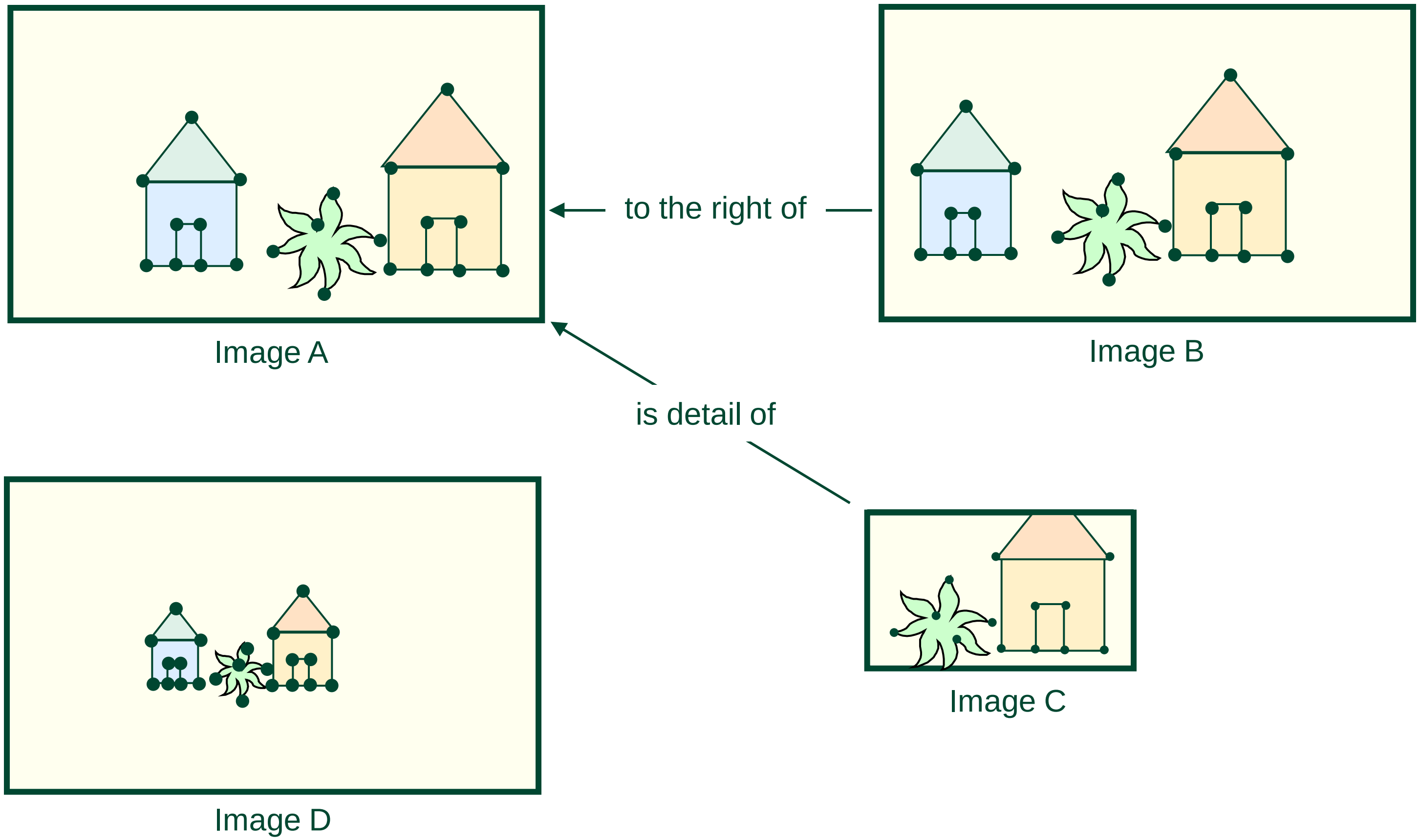
Relation-based browsing



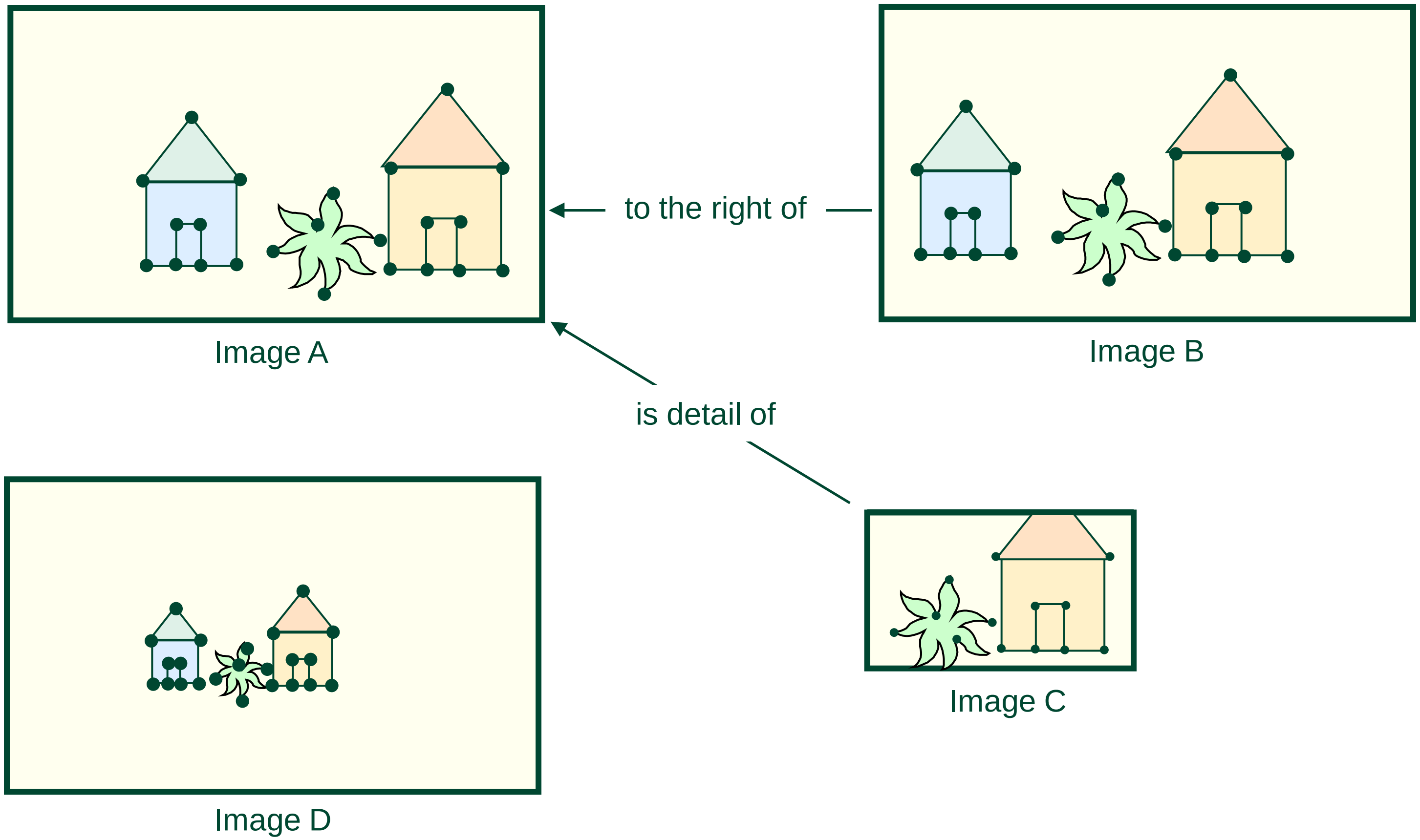
Relation-based browsing



Relation-based browsing



Relation-based browsing



Relation-based browsing

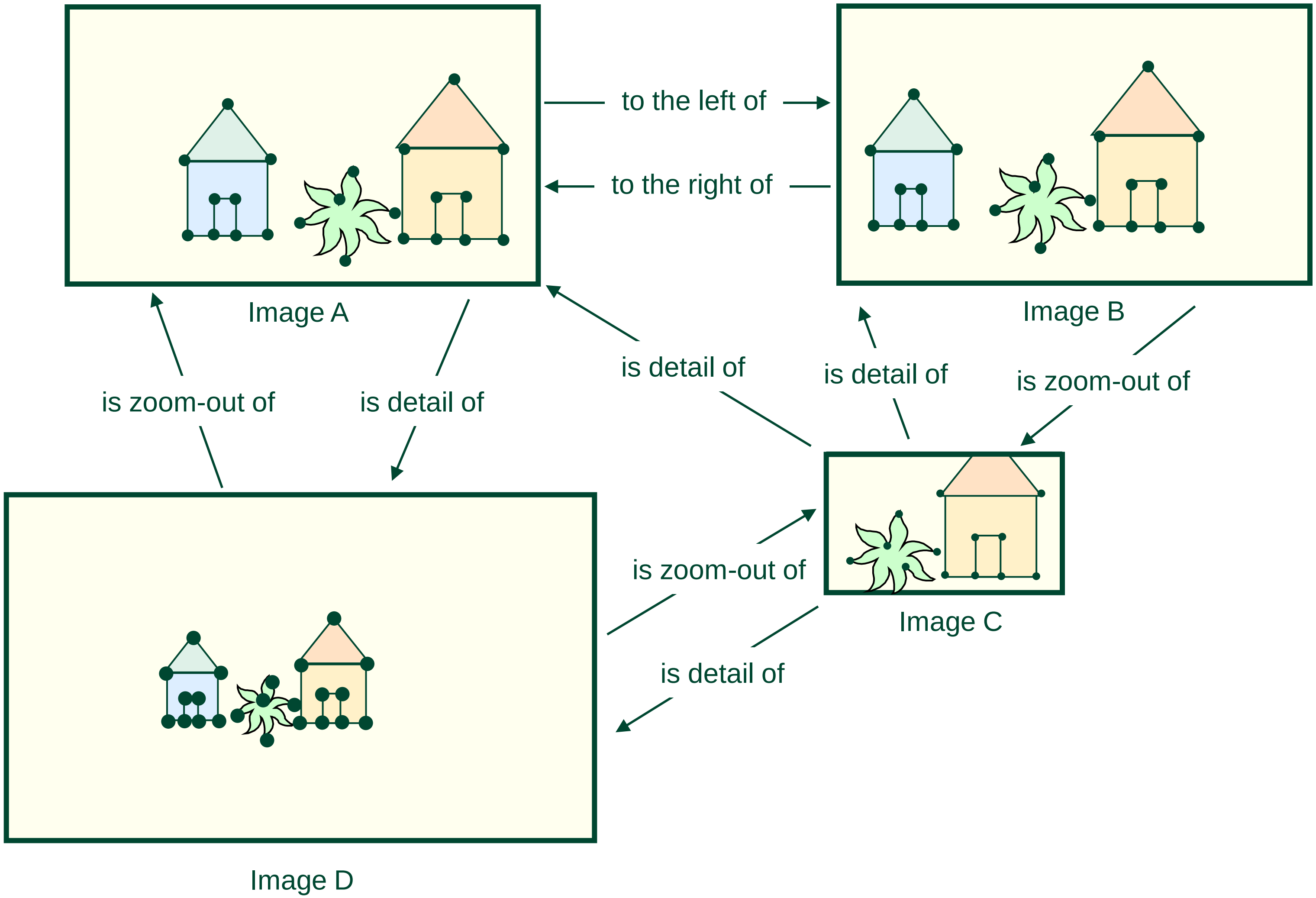
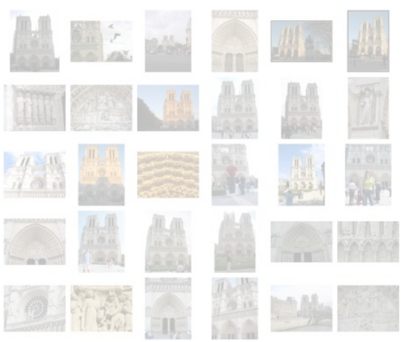


Photo Tourism overview



Input photographs

Scene reconstruction

Photo Explorer

- Navigation
- Rendering
- Annotations

Rendering

Rendering(Scene)

3D line segments extracted automatically from the photo collection and these washed-out looking colors

Rendering

Rendering transitions (photographs)

Image-based technique

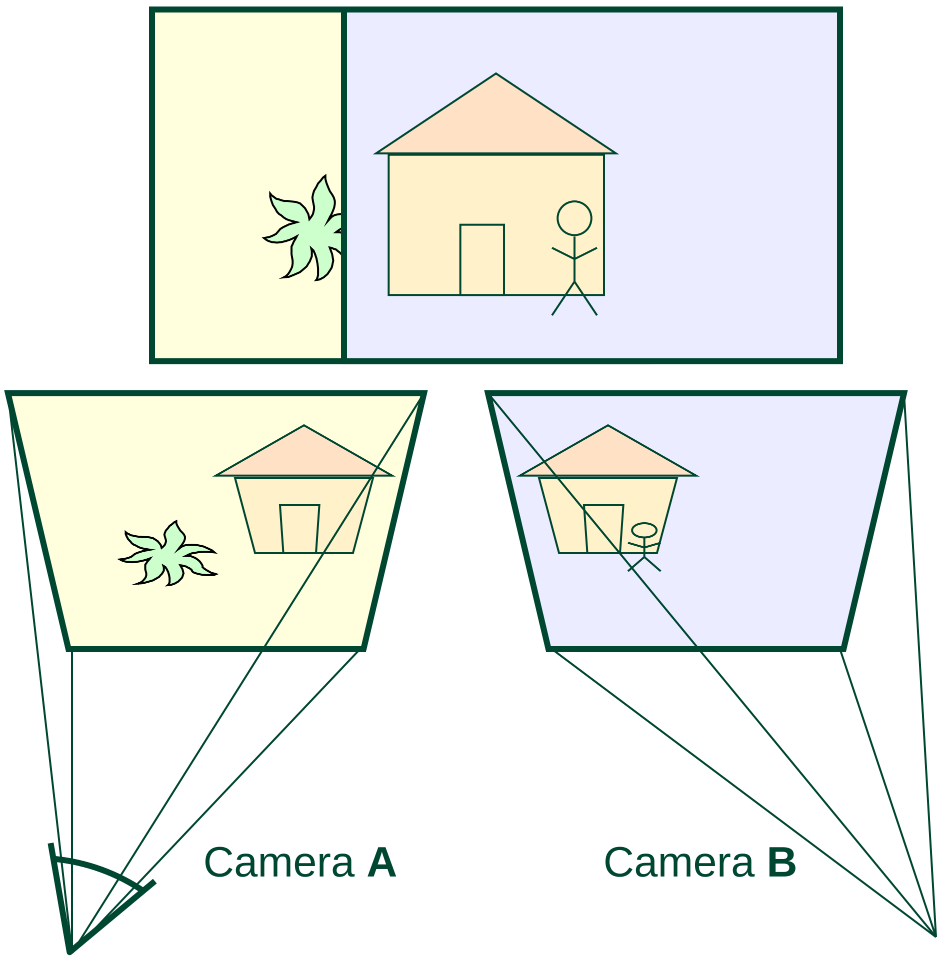
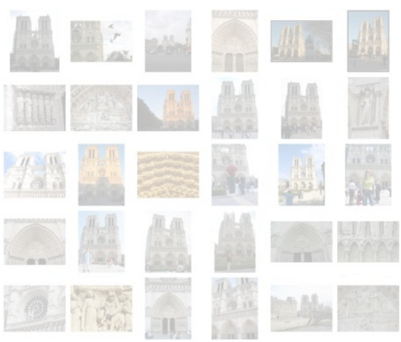


Photo Tourism overview



Input photographs

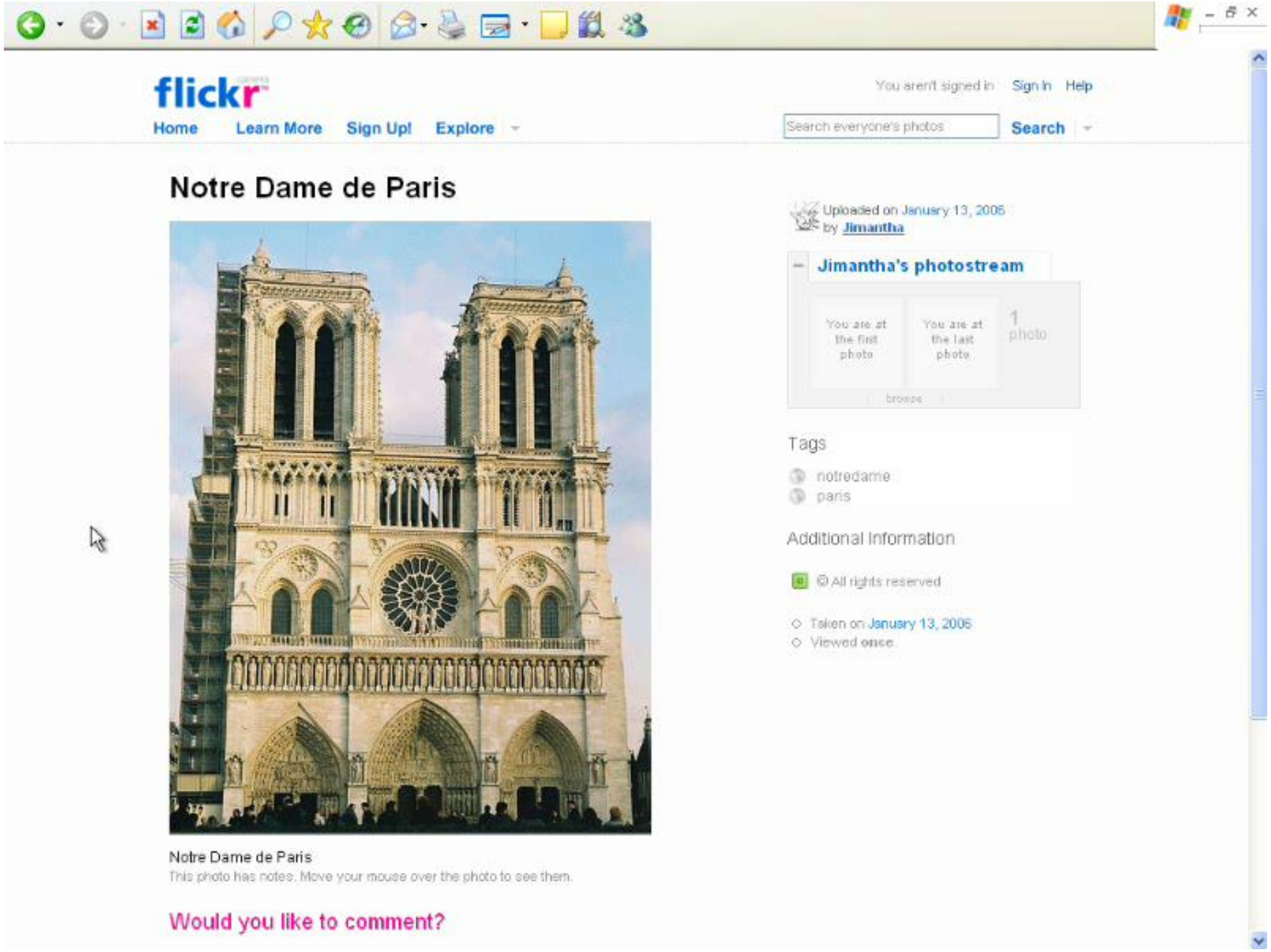
Scene reconstruction

Photo Explorer

- Navigation
- Rendering
- Annotations

Annotations

Annotations



Contributions

- Automated system for registering photo collections in 3D for interactive exploration
- Structure from motion algorithm demonstrated on hundreds of photos from the Internet
- Photo exploration system combining new image-based rendering and photo navigation techniques

Limitations / Future work

- Not all photos can be reliably matched








- Structure from motion scalability

  → More ns

- Plane-base

Conclusion

Indexing everyone's photos provides a new way to share and experience our world

To find out more:

- http://phototour.cs.washington.edu
- http://research.microsoft.com/IVM/PhotoTourism
- http://labs.live.com/photosynth
- Exhibition booth #2619

Saint Basil's Cathedral

Trafalgar Square

Rockefeller Center

Mount Rushmore

FaceTracer:

A Search Engine for Large Collections of Images with Faces

Neeraj Kumar,
Peter Belhumeur, Shree Nayar

Columbia University

How Can We Describe This Face?

Woman

Young

Asian

Brunette

Smiling

…

How Can We Describe The Image?

Indoors

Flash

In Focus

Frontal

Alone

…

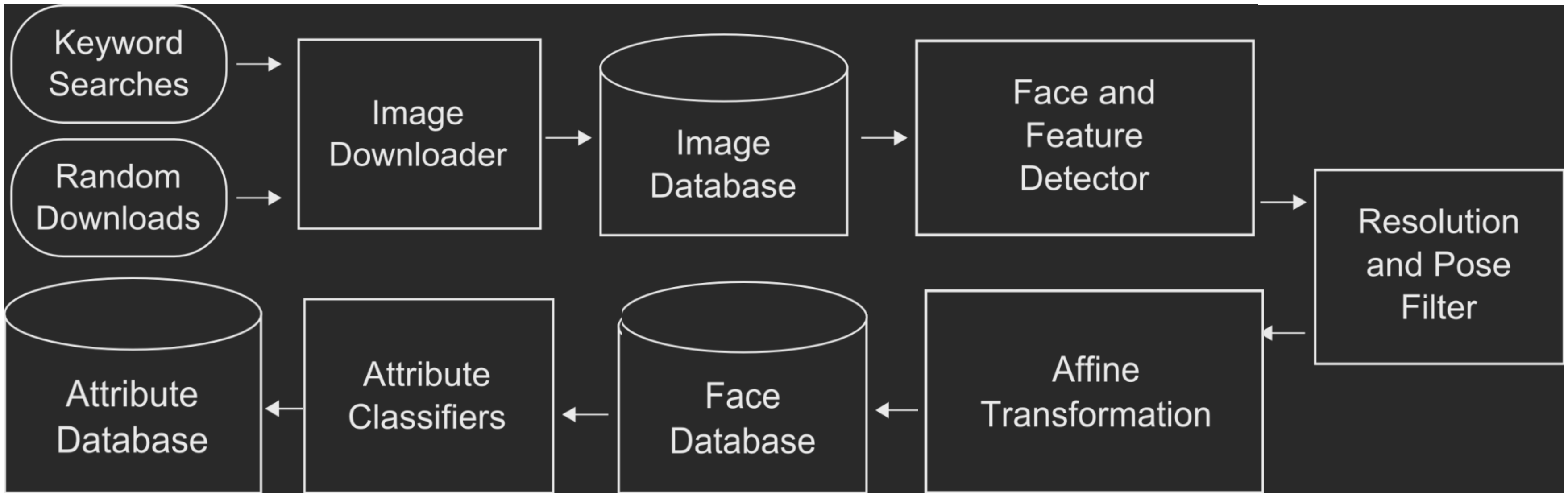
We Need a Search Engine Based on Facial and Image Appearance

Some Numbers

- Billions of Images
- Hundreds of Attributes
- Thousands of Manual Labels

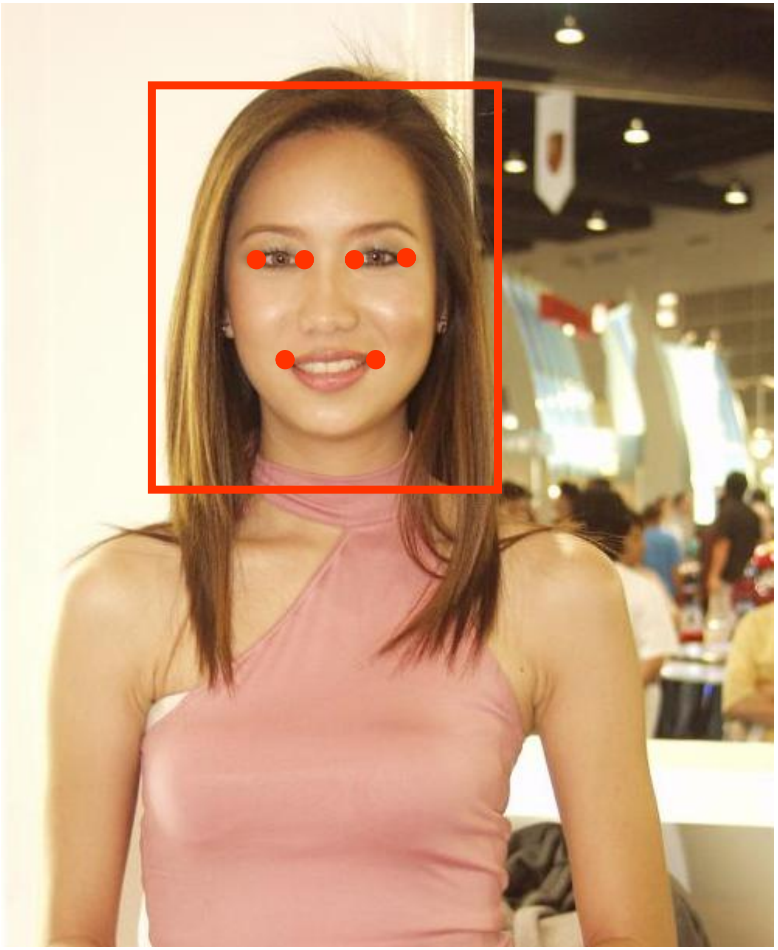
We need to do this automatically

Overview of Database Creation

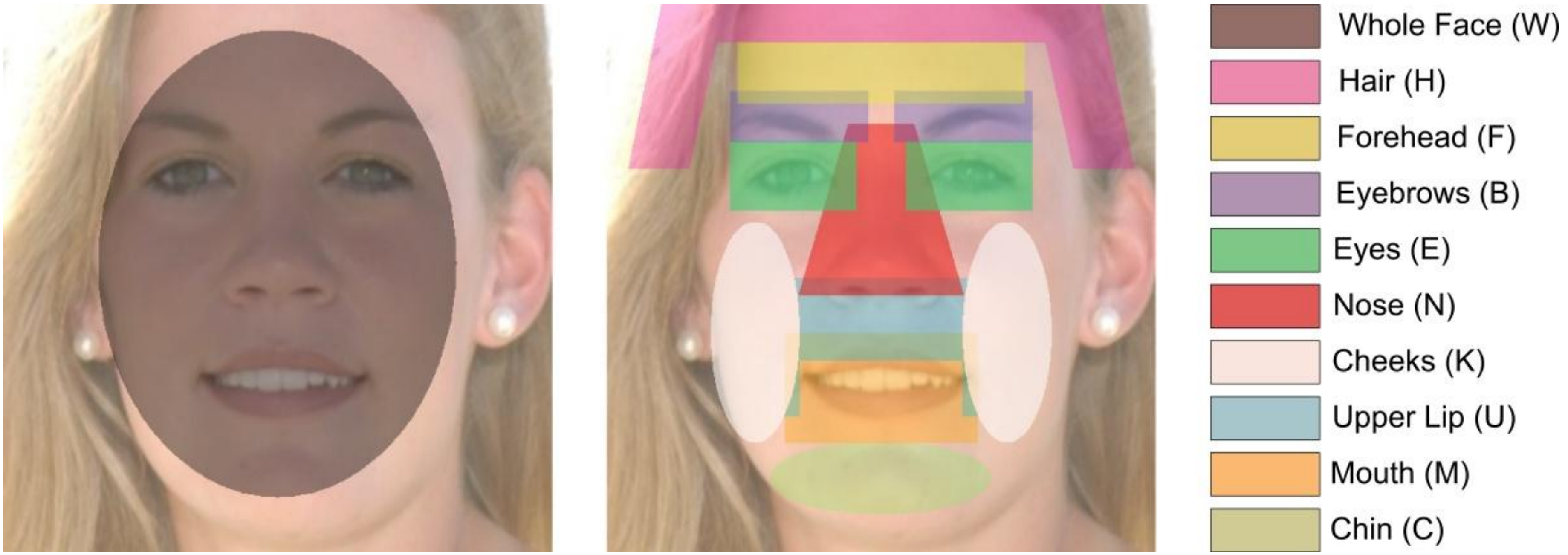


- •OKAO face detector
  - •Pose angles
  - • 6 Facial keypoints location
- • Alignment by linear least squares on detected facial points and corresponding points on a template face.

Detect and Align



1. Face Regions



Feature Types

2. Pixel Value Type	3. Normalizations	Aggregation
RGB (r)	None (n)	None (n)
HSV (h)	Mean-Norm (m)	Histogram (h)
Image Intensity (i)	Energy-Norm (e)	Statistics (s)
Edge Magnitude (m)		
Edge Orientation (o)		

Feature Types

Pixel Value Type	Normalizations	Aggregation
RGB (r)	None (n)	None (n)
HSV (h)	Mean-Norm (m)	Histogram (h)
Image Intensity (i)	Energy-Norm (e)	Statistics (s)
Edge Magnitude (m)		
Edge Orientation (o)		

"Region:pixel type.normalization.aggregation."

RGB, Mean Norm., No Aggreg. (r.m.n)

Feature Types

Pixel Value Type	Normalizations	Aggregation
RGB (r)	None (n)	None (n)
HSV (h)	Mean-Norm (m)	Histogram (h)
Image Intensity (i)	Energy-Norm (e)	Statistics (s)
Edge Magnitude (m)		
Edge Orientation (o)		

"Region:pixel type.normalization.aggregation."

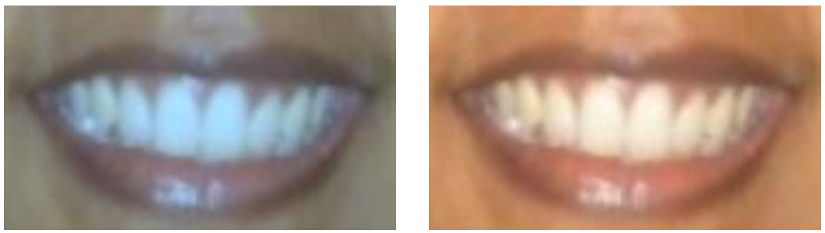
Edge Orientations, No Norm, Histogram (o.n.h)

Train Classifiers

Mouth
Raw RGB

Pool of Classifiers- one per region/feature type

Train Classifiers

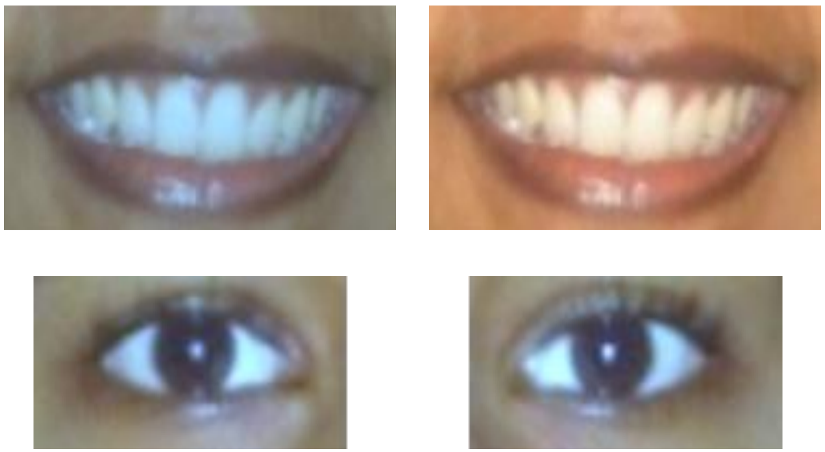


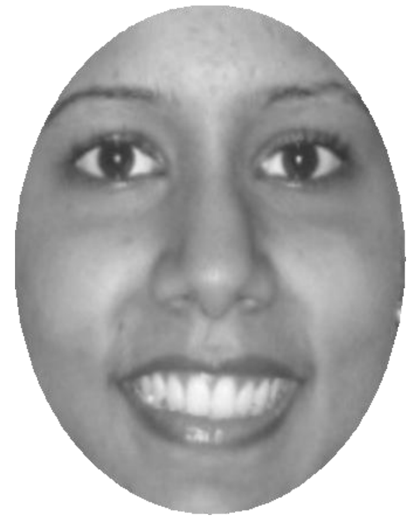
Eyes

Mean-Normalized RGB

Pool of Classifiers- one per region/feature type

Train Classifiers



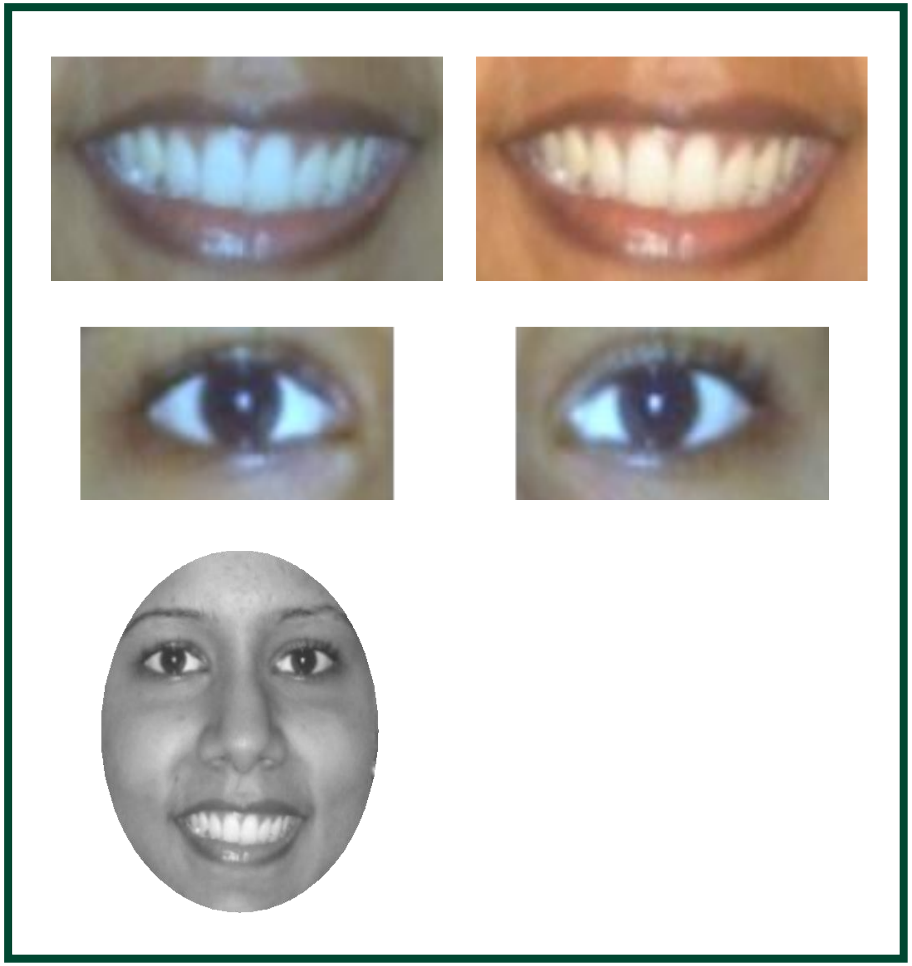


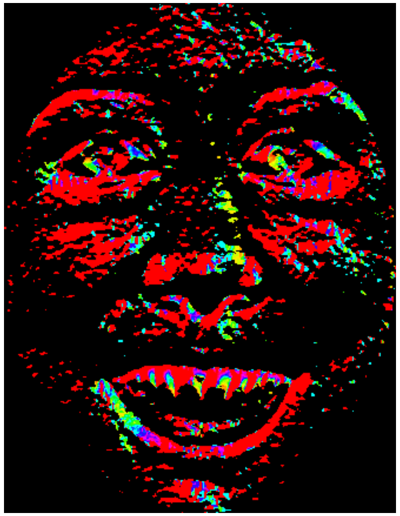
Whole Face

Raw Intensity

Pool of Classifiers- one per region/feature type

Train Classifiers



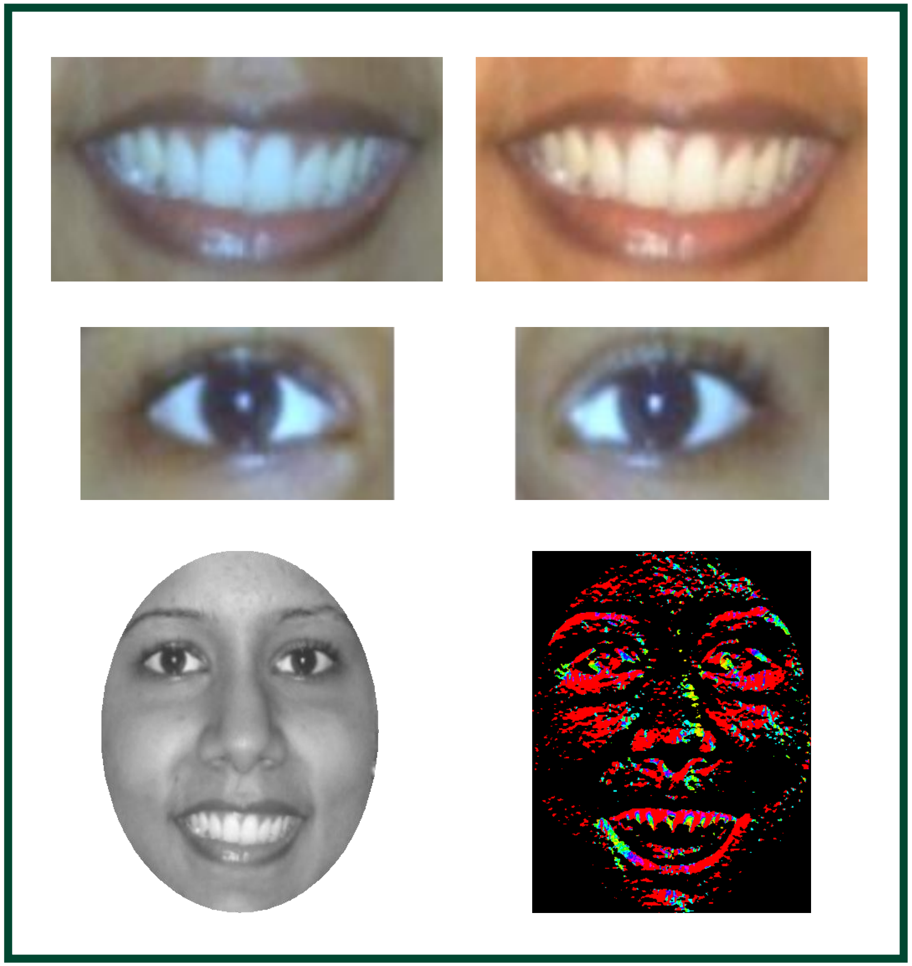


Whole Face

Gradient Directions

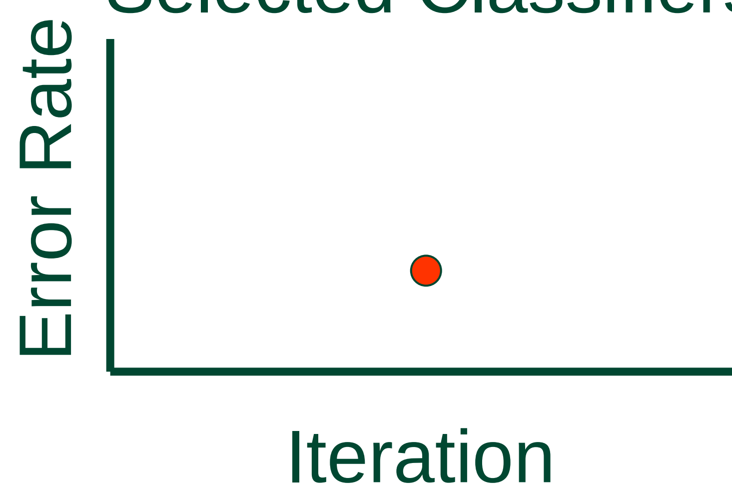
Pool of Classifiers- one per region/feature type

Select Classifiers



Pool of Classifiers

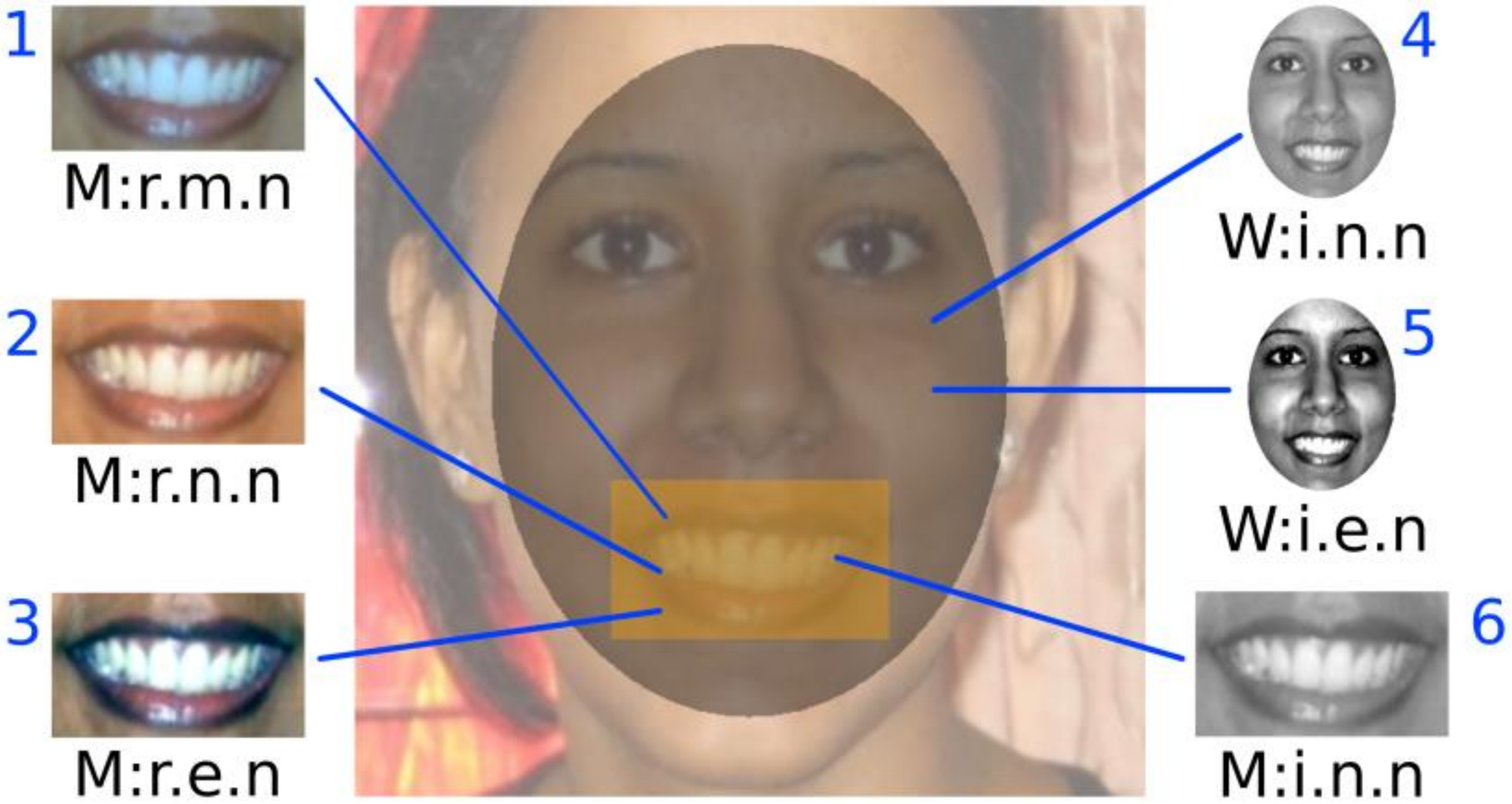
Selected Classifiers



Feature Selection: Smiling

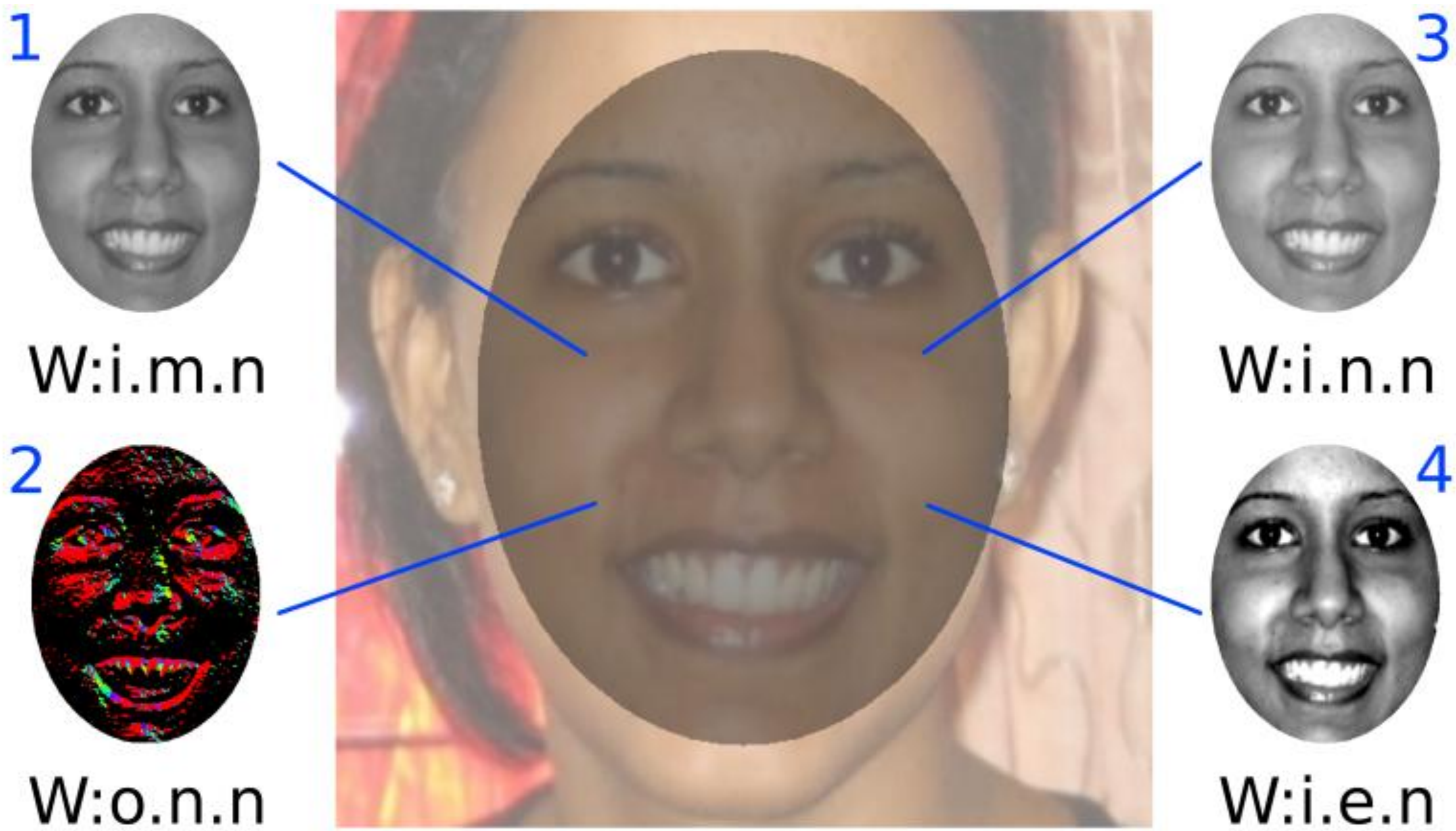
1. Mouth: RGB, Mean Norm., No Aggreg. (M:r.m.n)
2. Mouth: RGB, No Norm., No Aggreg. (M:r.n.n)
3. Mouth: RGB, Energy Norm., No Aggreg. (M:r.e.n)
4. Whole Face: Intensity, No Norm., No Aggreg. (W:i.n.n)
5. …

Selected Features



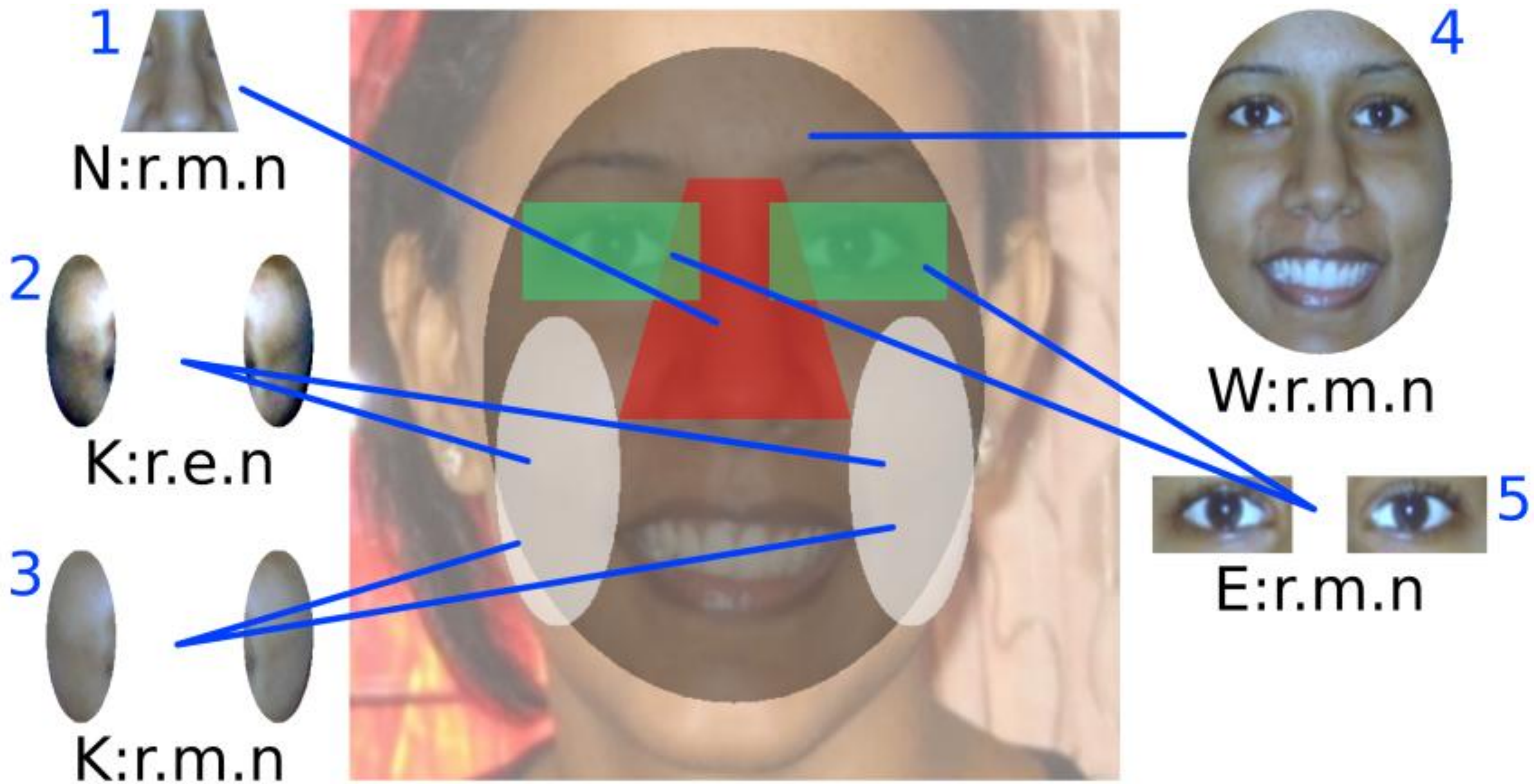
Smiling

Selected Features



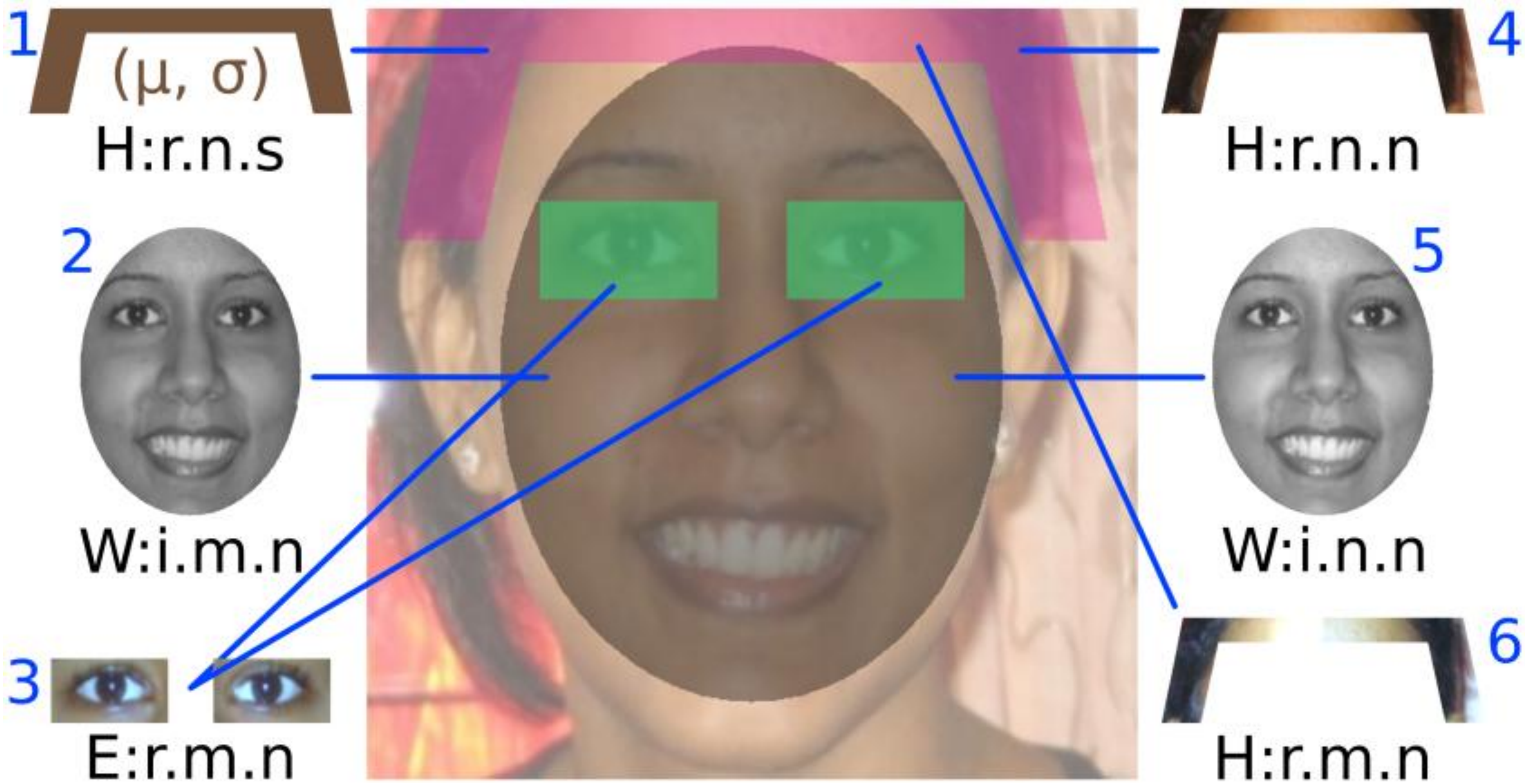
Gender

Selected Features



Indoor/Outdoor

Selected Features



Hair Color

Classification Accuracy

Attribute	Error Rate
Gender	8.62%
Age	16.65%
Race	6.49%
Hair Color	5.54%
Eye Wear	5.14%

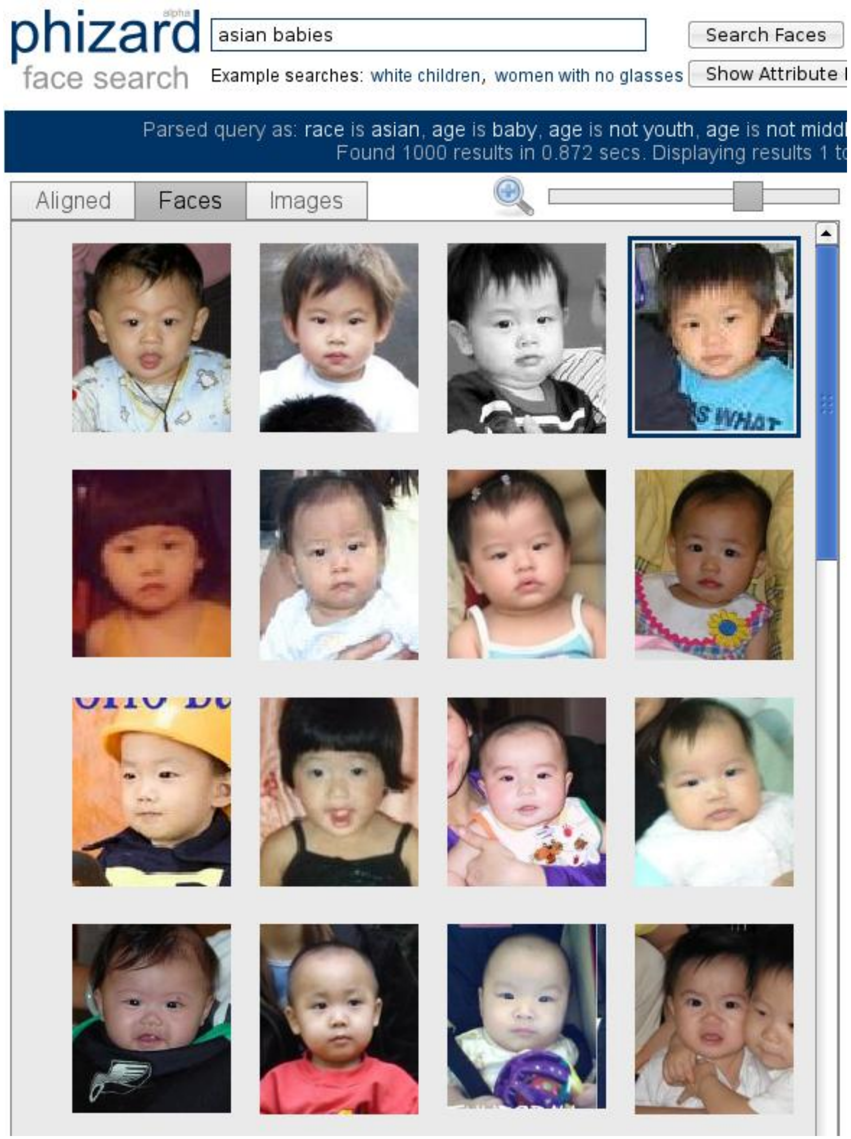
Attribute	Error Rate
Mustache	4.61%
Smiling	4.60%
Blurry	3.41%
Lighting	1.61%
Environment	12.15%

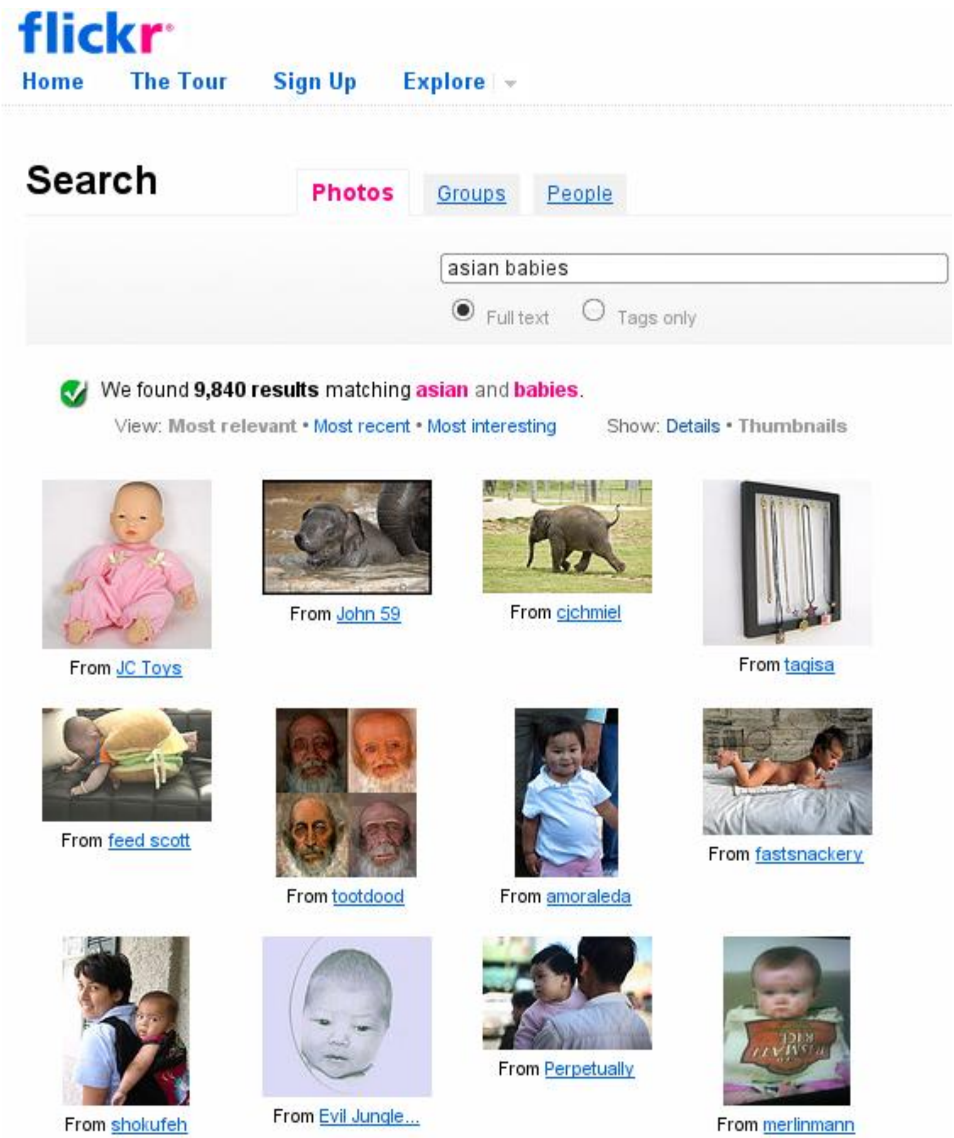
Comparison to State-of-the-Art

Method	Gender Error Rate	Smiling Error Rate
Proposed	8.62%	4.60%
Baluja & Rowley, IJCV 2007	13.13%	7.41%
Shakhnarovich et al., ICAFGR 2002	12.88%	6.40%
Moghaddam & Yang, TPAMI 2002	9.52%	13.54%

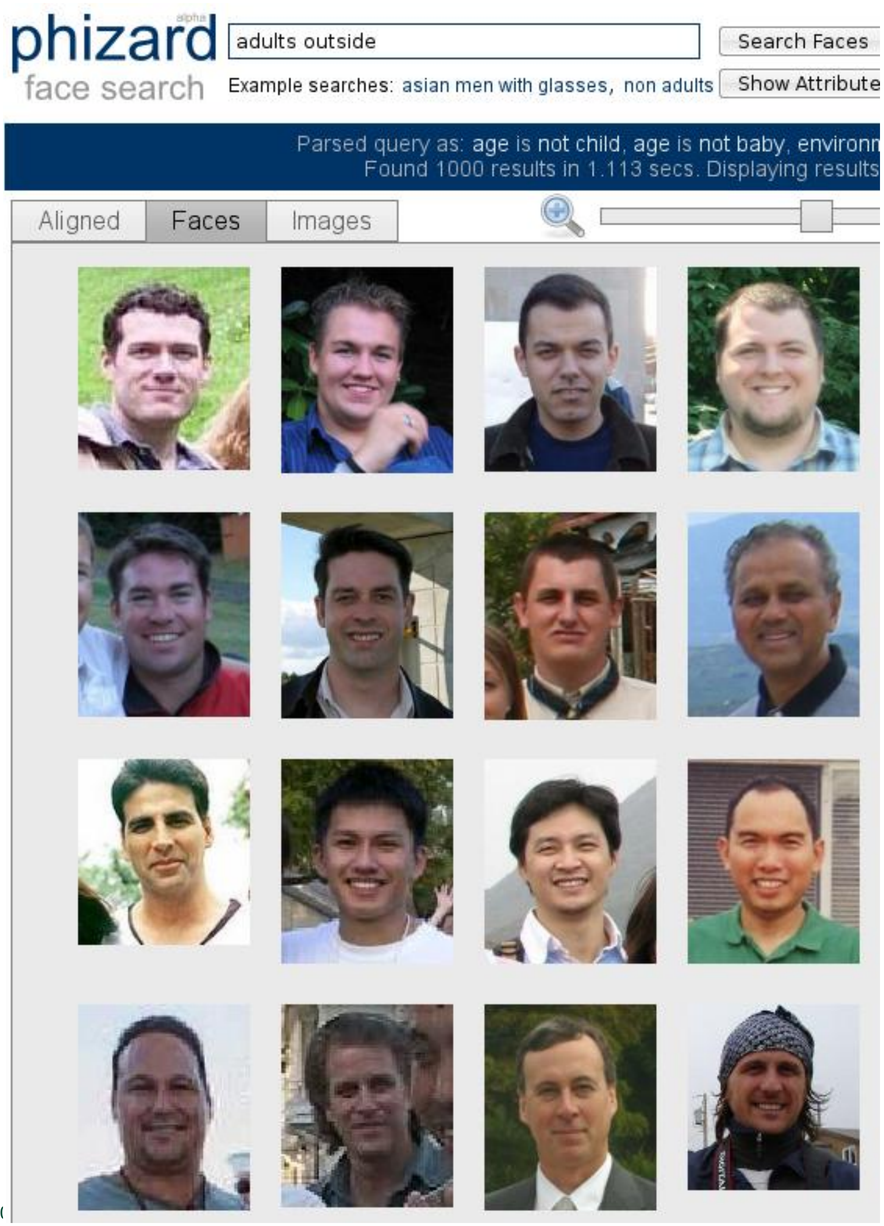
Results

"Asian Babies"

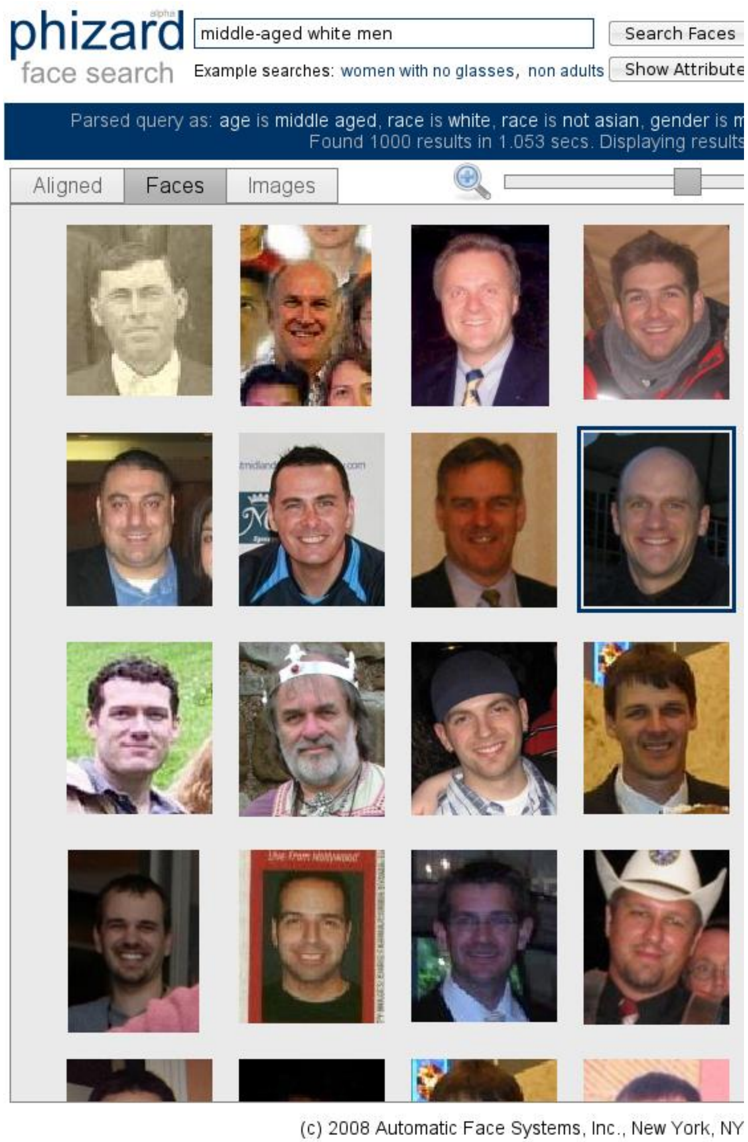


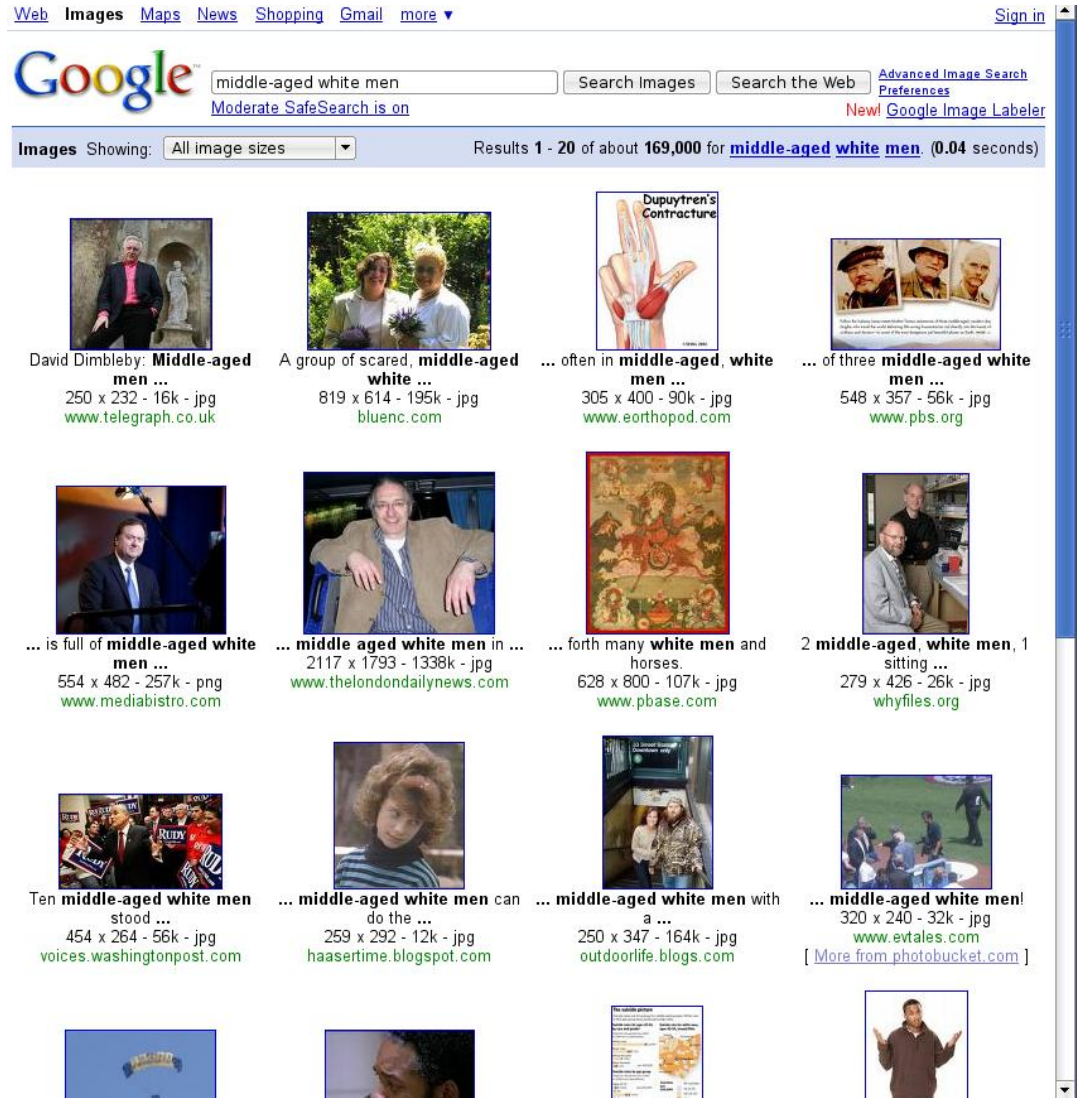


“Adults Outside”

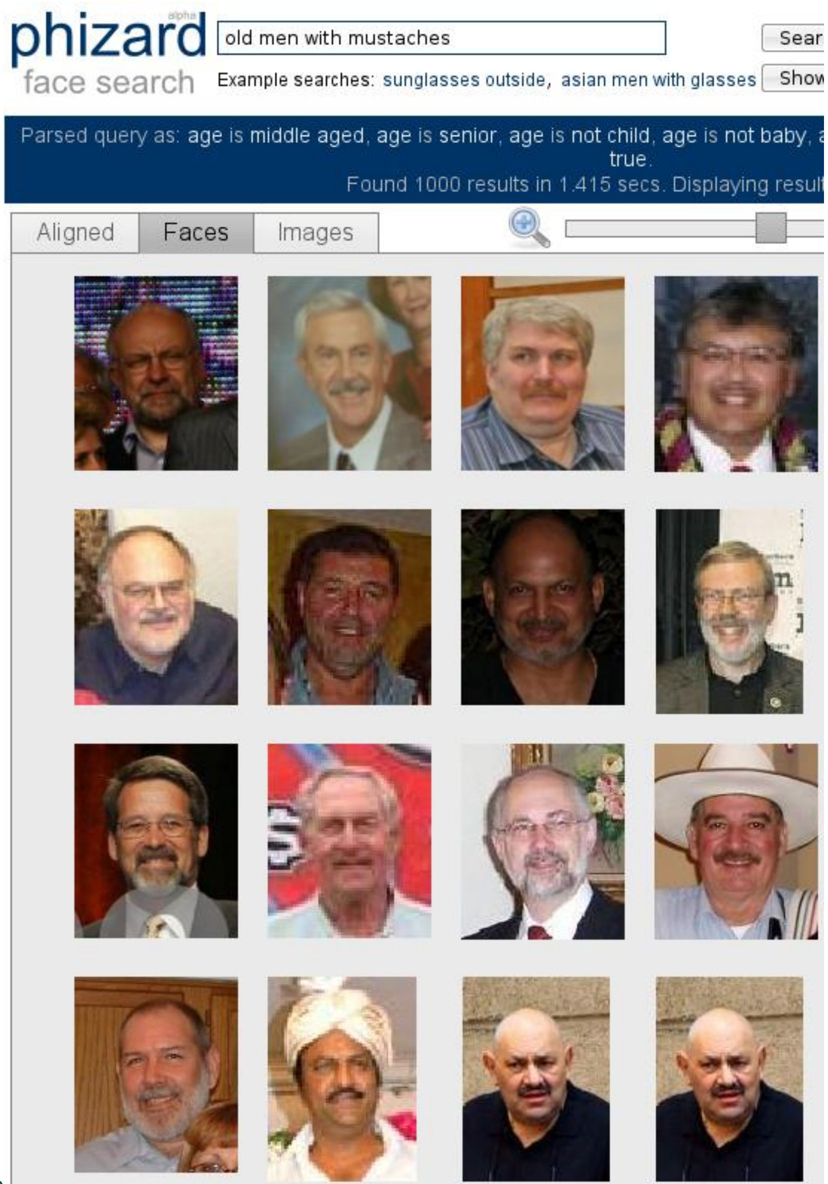


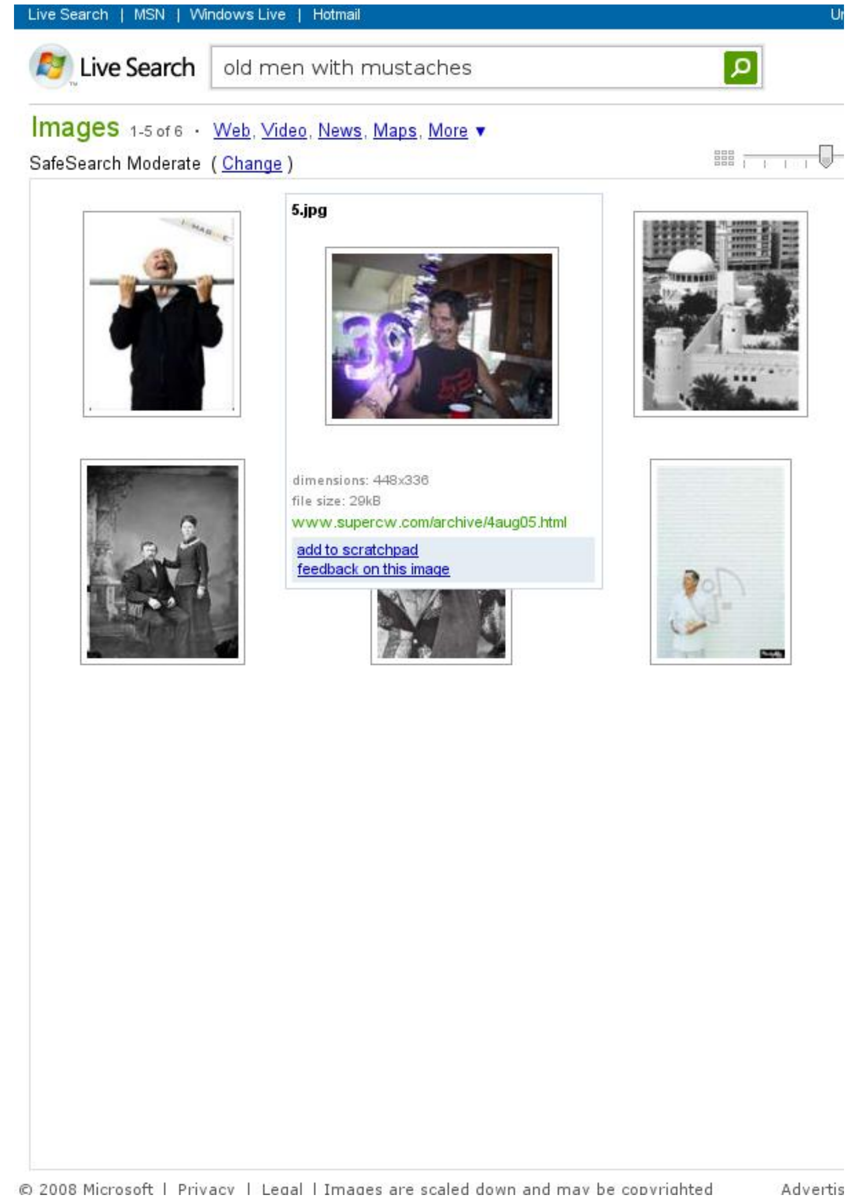
“Middle-Aged White Men”



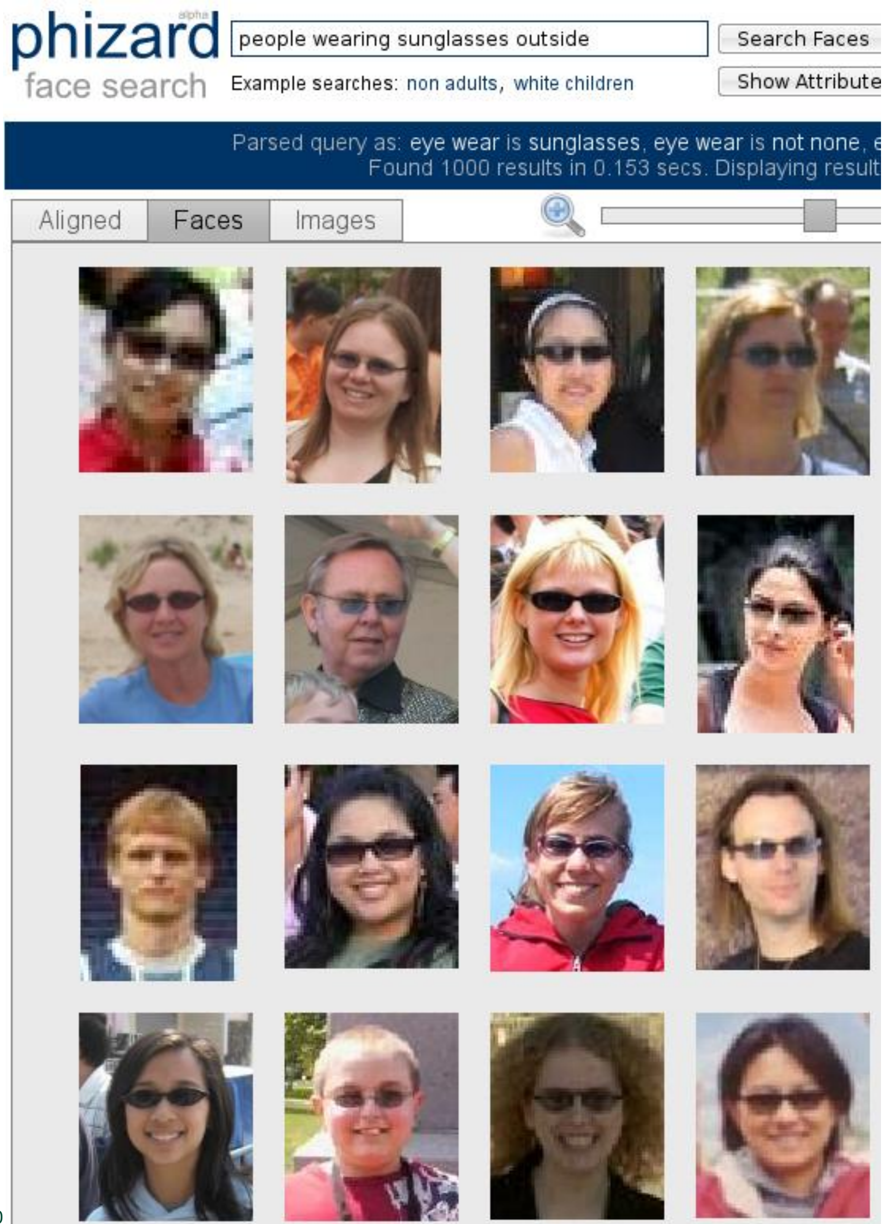


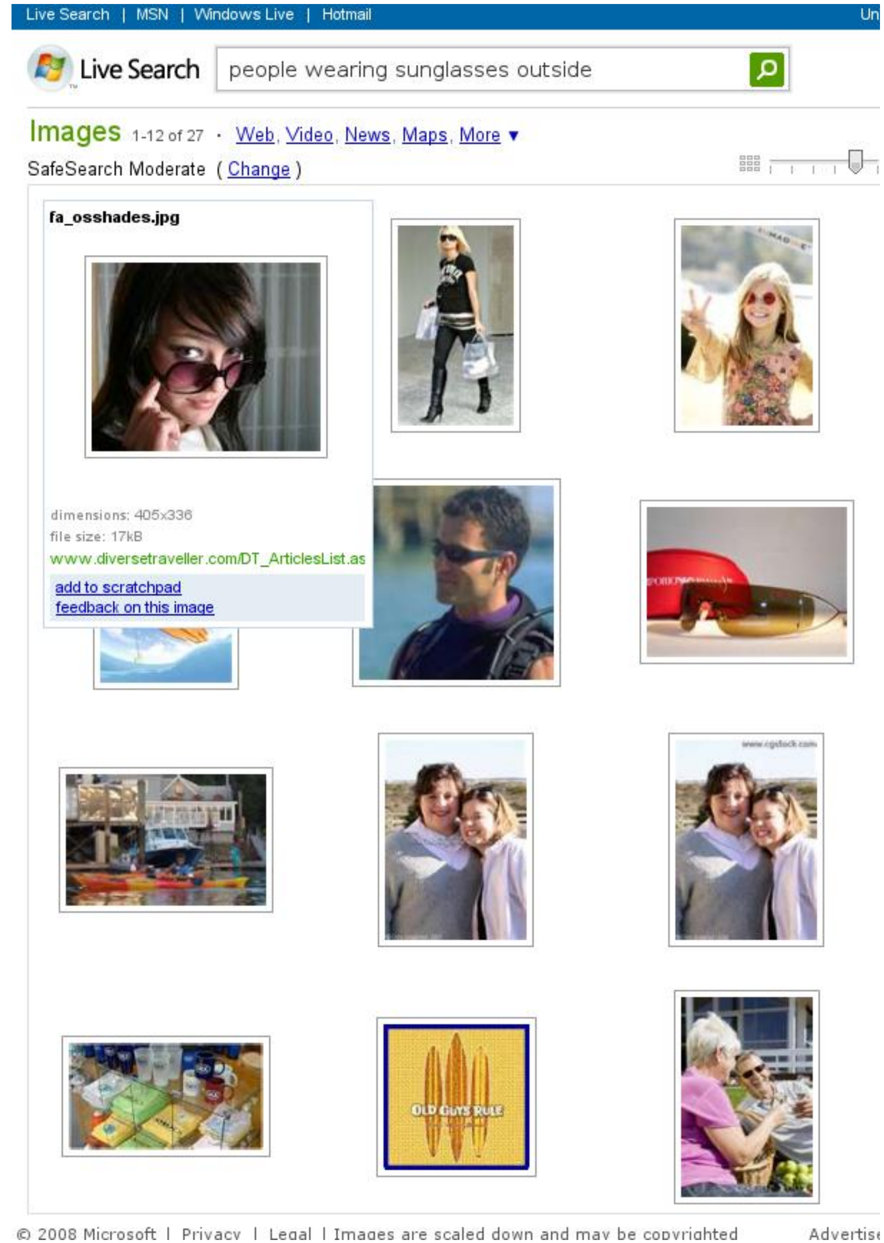
“Old Men With Mustaches”



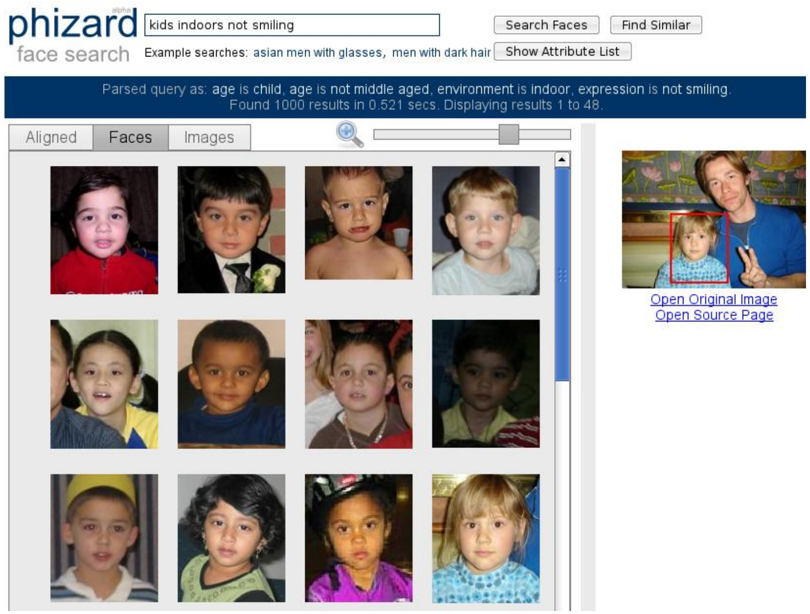


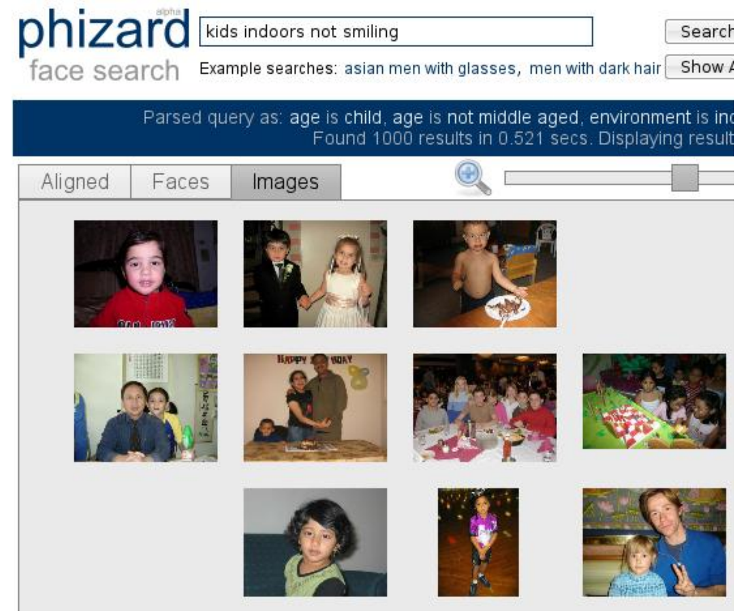
“People Wearing Sunglasses Outside”

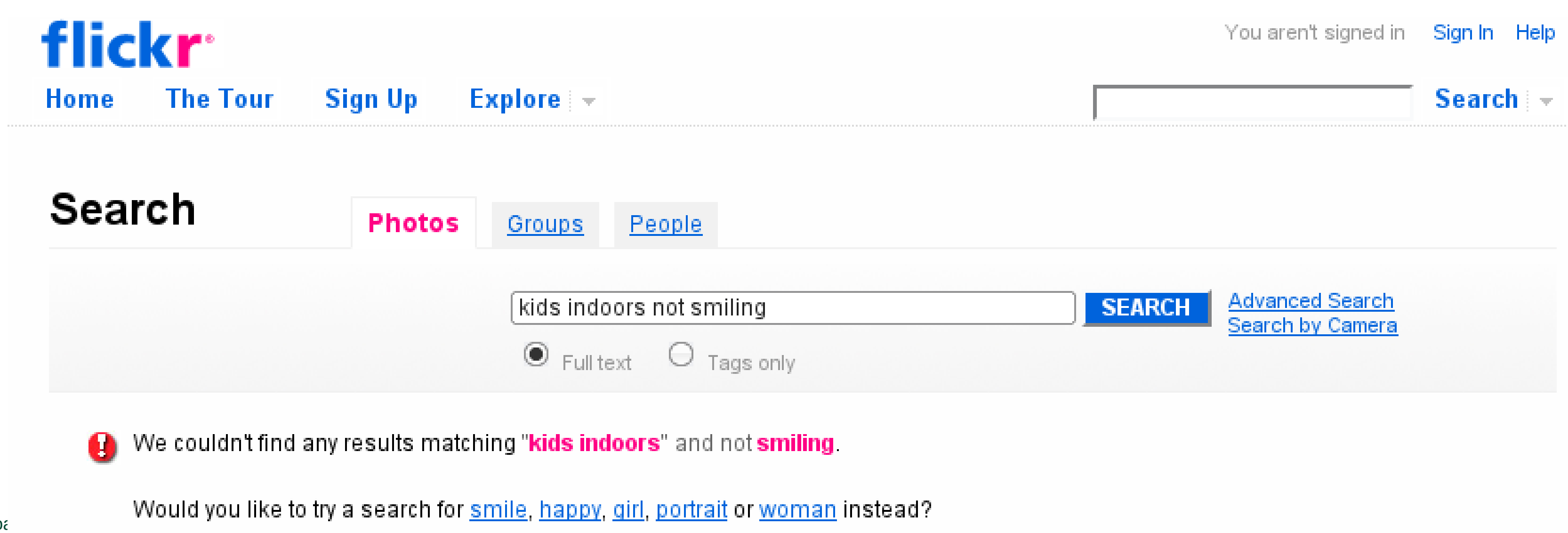




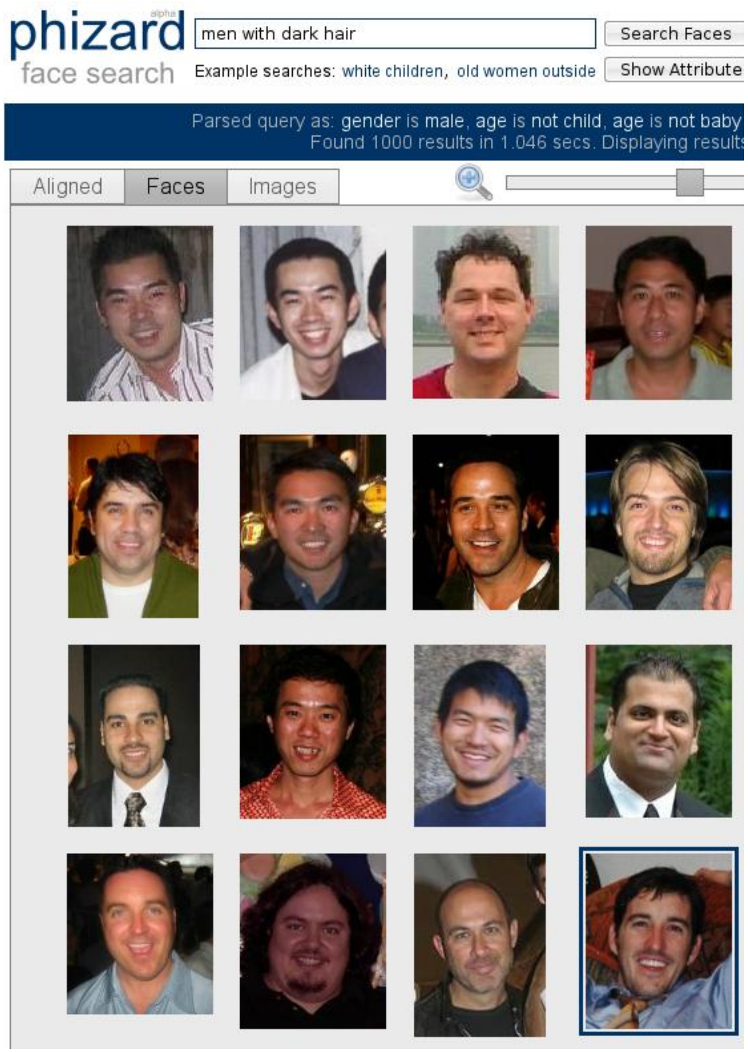
“Kids Indoors Not Smiling”

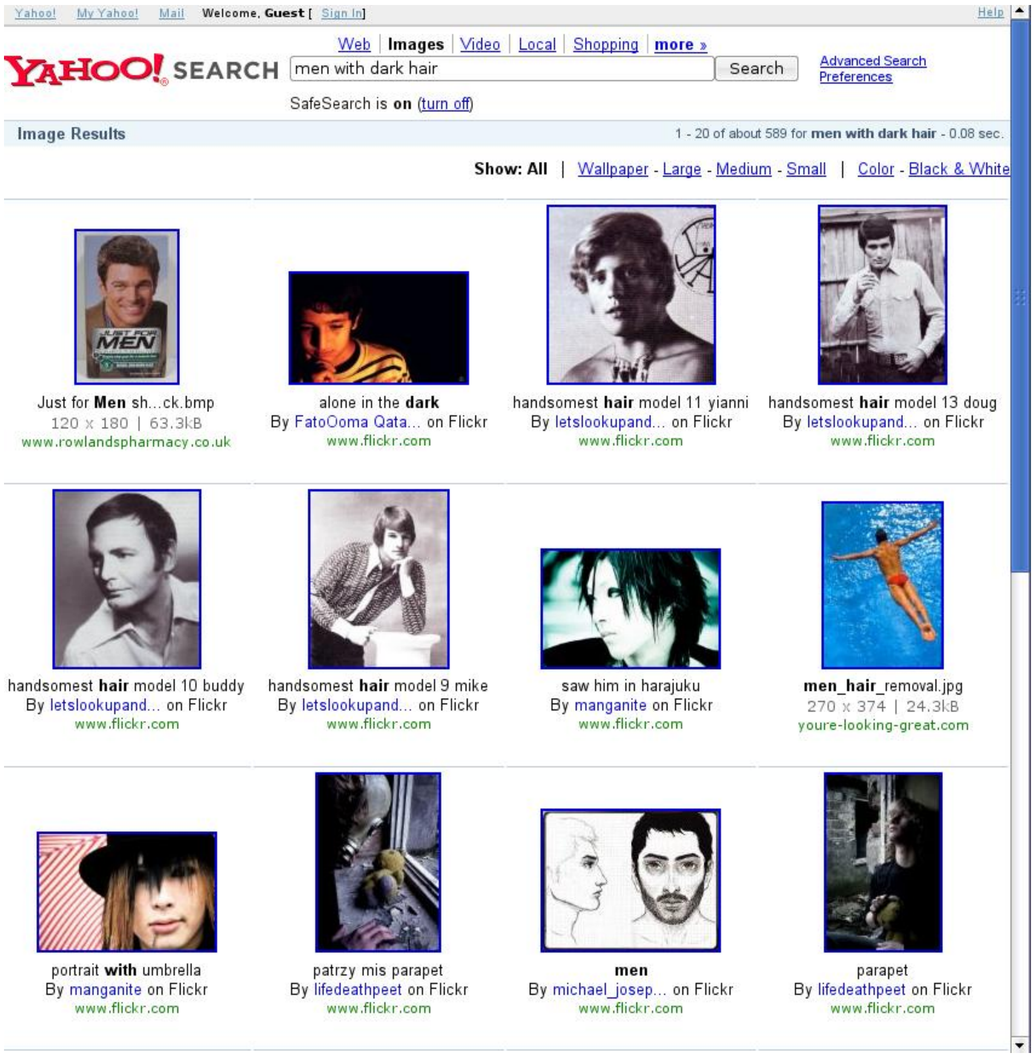




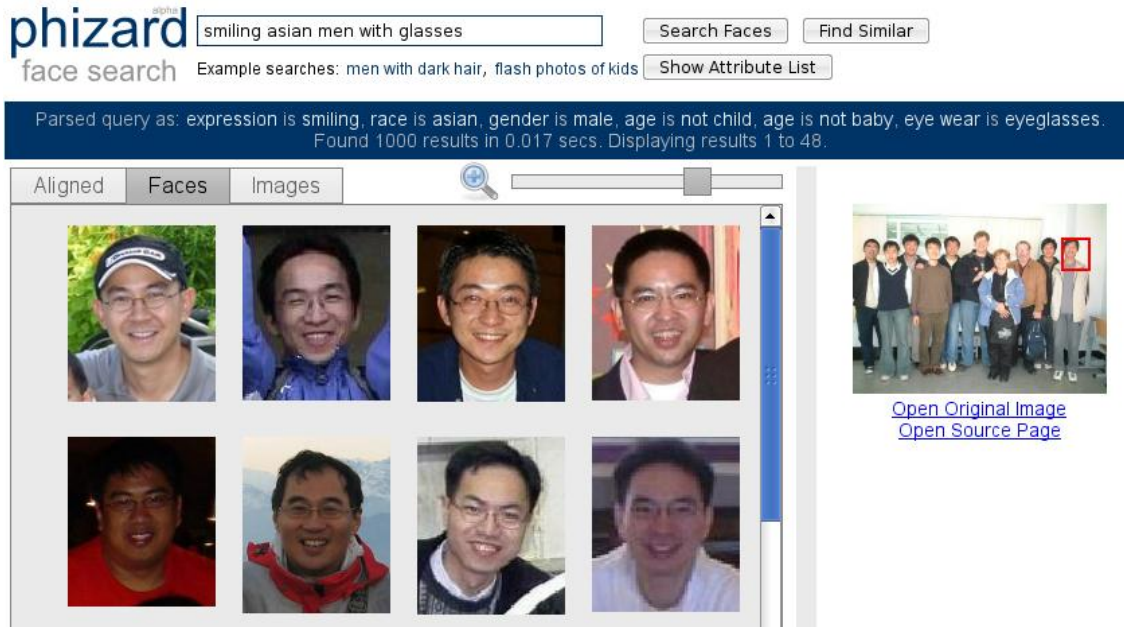


“Men With Dark Hair”

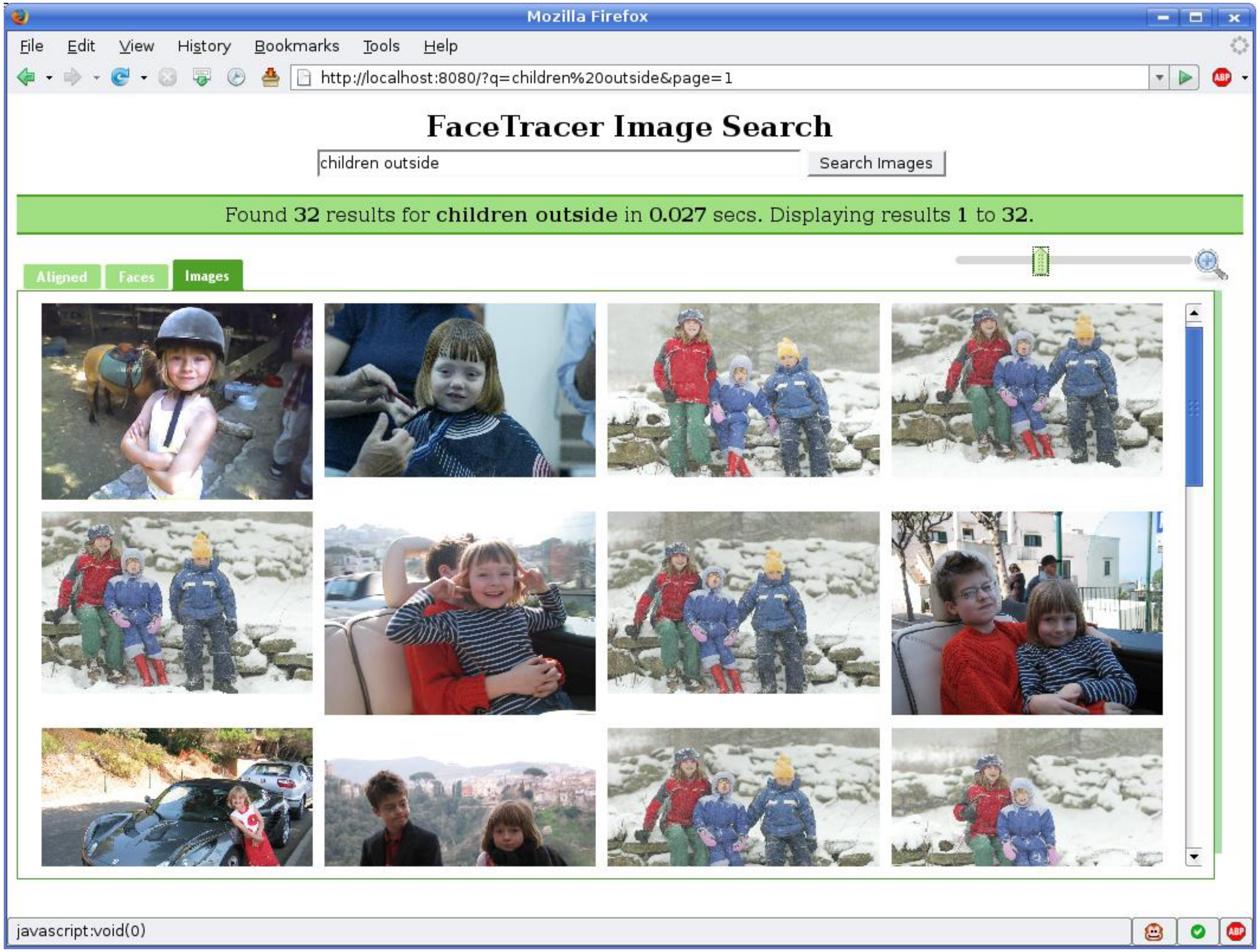




“Smiling Asian Men With Glasses”



Personal FaceTracer Search



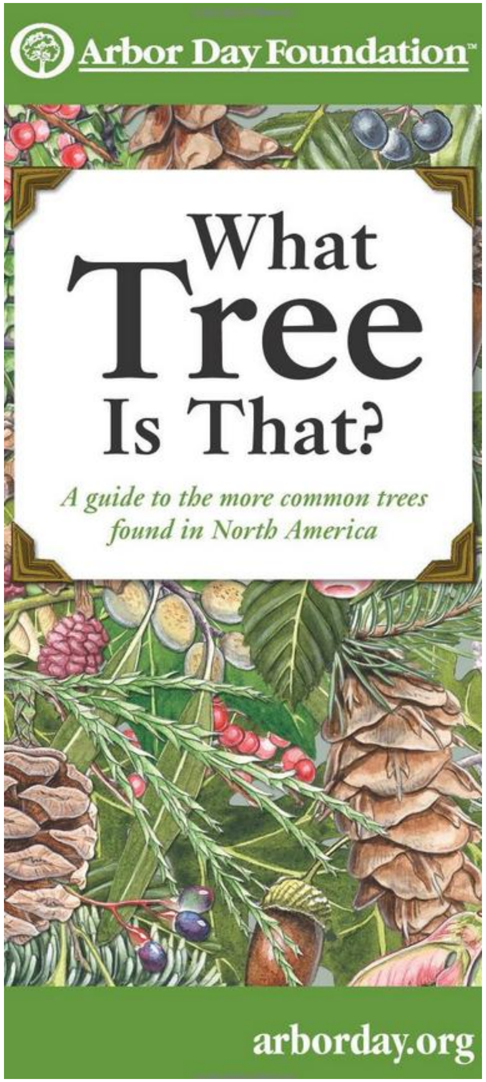
“Children outside”

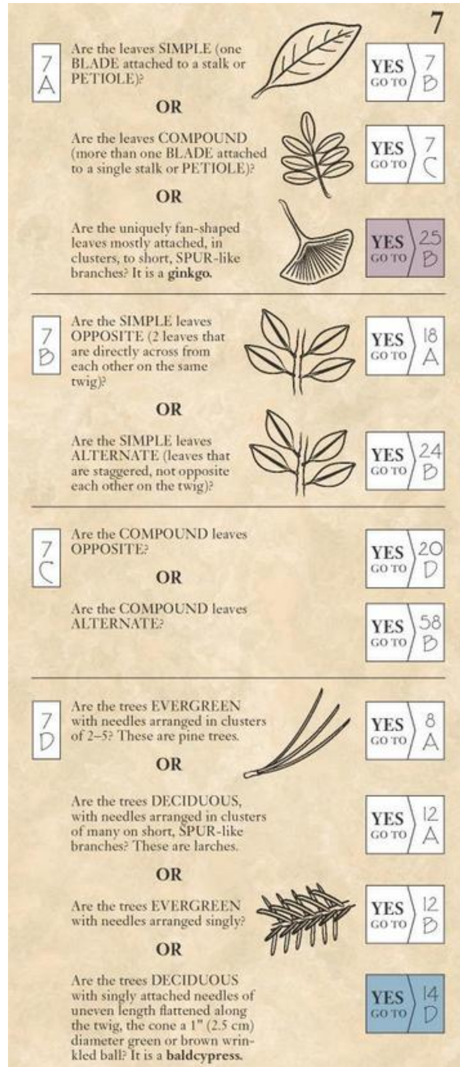
A Computer Vision System for Automatic Plant Species Identification

Neeraj Kumar — University of Washington
Peter N. Belhumeur — Columbia University
Arijit Biswas — University of Maryland
David W. Jacobs — University of Maryland
W. John Kress — Smithsonian Institution
Ida C. Lopez — Smithsonian Institution
João V.B. Soares — University of Maryland

What Plant Species is this?

Let’s Use a Field Guide





7

7A Are the leaves SIMPLE (one BLADE attached to a stalk or PETIOLE)? YES GO TO 7B

OR

Are the leaves COMPOUND (more than one BLADE attached to a single stalk or PETIOLE)? YES GO TO 7C

OR

Are the uniquely fan-shaped leaves mostly attached, in clusters, to short, SPUR-like branches? It is a **ginkgo.** YES GO TO 25B

7B Are the SIMPLE leaves OPPOSITE (2 leaves that are directly across from each other on the same twig)? YES GO TO 18A

OR

Are the SIMPLE leaves ALTERNATE (leaves that are staggered, not opposite each other on the twig)? YES GO TO 24B

7C Are the COMPOUND leaves OPPOSITE? YES GO TO 20D

OR

Are the COMPOUND leaves ALTERNATE? YES GO TO 58B

7D Are the trees EVERGREEN with needles arranged in clusters of 2–5? These are pine trees. YES GO TO 8A

OR

Are the trees DECIDUOUS, with needles arranged in clusters of many on short, SPUR-like branches? These are larches. YES GO TO 12A

OR

Are the trees EVERGREEN with needles arranged singly? YES GO TO 12B

OR

Are the trees DECIDUOUS with singly attached needles of uneven length flattened along the twig, the cone a 1" (2.5 cm) diameter green or brown wrinkled ball? It is a **baldcypress.** YES GO TO 14D

9

9A EASTERN WHITE PINE
Pinus strobus
Zones 3–8
The fabled tree eagerly sought by the first wave of loggers in America. The provincial tree of Ontario.

9B PITCH PINE
Pinus rigida
Zones 4–7
The fire resilient conifer of the East, even producing new branches and needles after fire kills the green foliage

9C LONGLEAF PINE
Pinus palustris
Zones 7–10
A tall, stately pine of the South long sought by loggers

9D LOBLOLLY PINE
Pinus taeda
Zones 6–9
Our most important and widely cultivated timber species in the southern United States

leafsnap

Like a normal field guide...

leafsnap

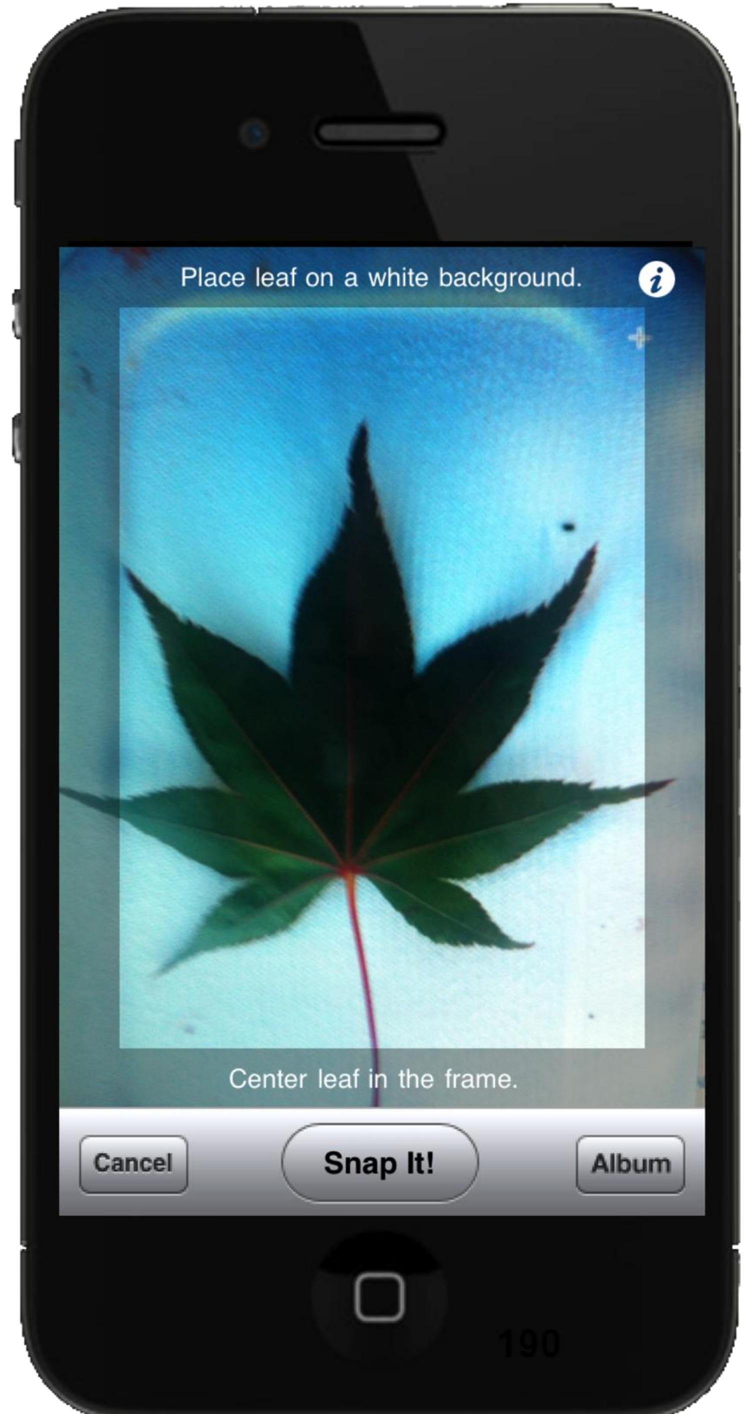
Like a normal field guide…

- that you can search and sort

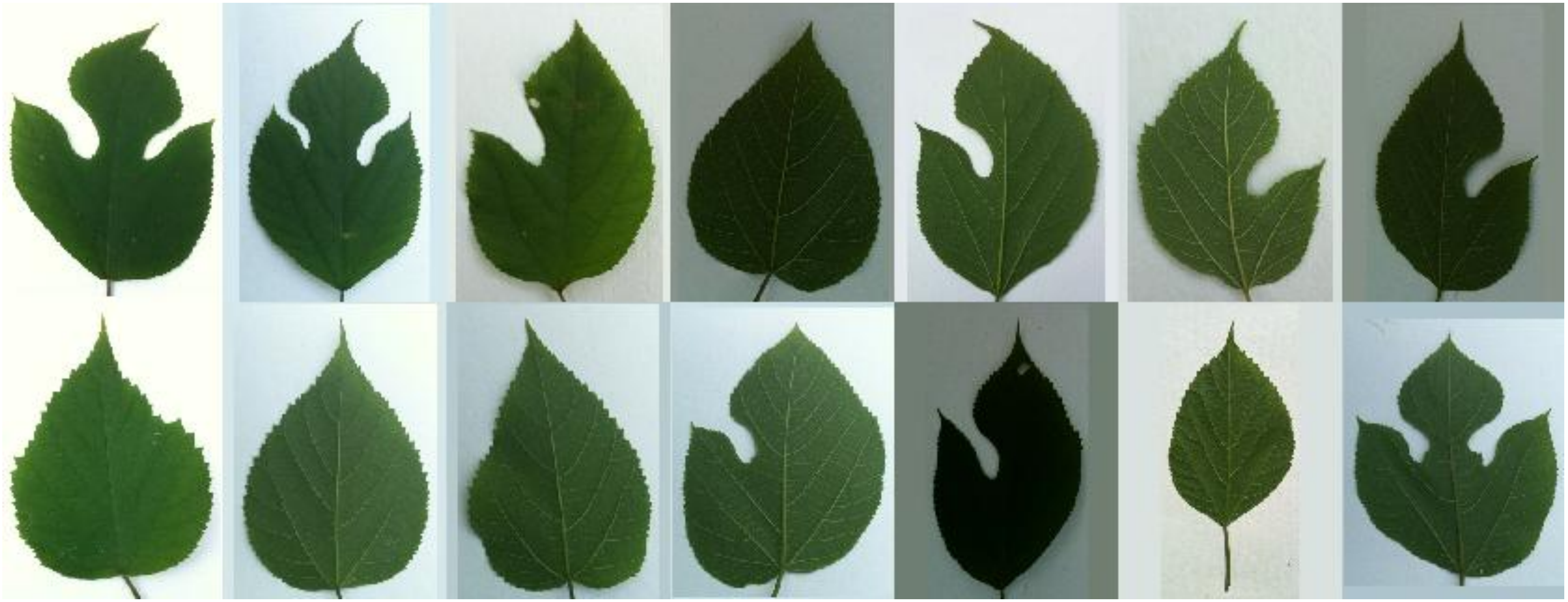
leafsnap

Like a normal field guide…

- that you can search and sort
- and with visual recognition



Large Intra-Class Variation



Images of Paper Mulberry (*Broussonettia papyrifera*)

What do iPhone Images Look Like?

Blurry, Not flat

What do iPhone Images Look Like?

Blurry, Not flat, Varying color, Shadows

What do iPhone Images Look Like?

Blurry, Not flat, Varying color, Shadows, No venation, Thin Structures

What do iPhone Images Look Like?

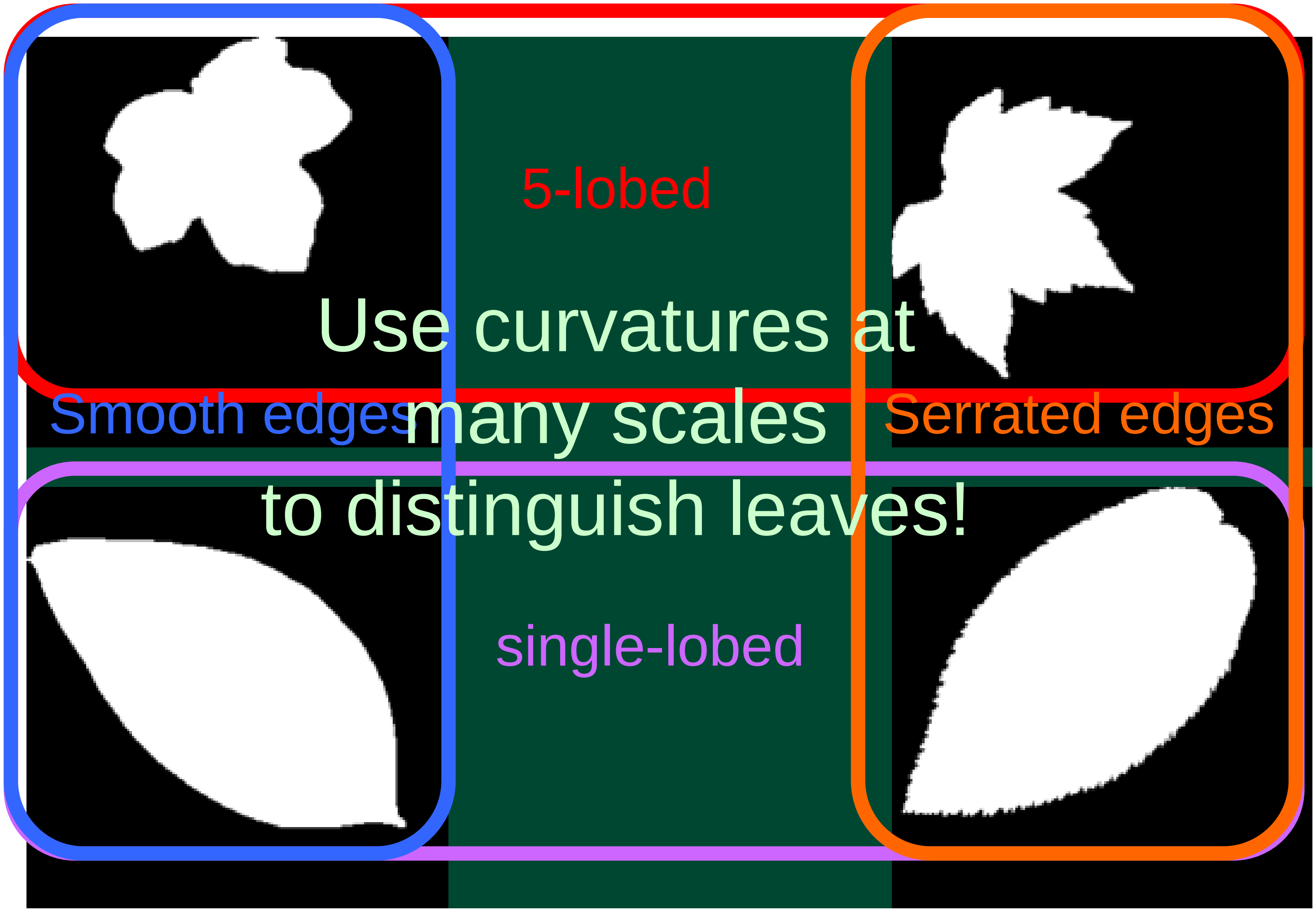
Blurry, Not flat, Varying color, Shadows, No venation, Thin Structures, White balance, Color splotches, …

Classification

- Classifying whether the image is of a valid leaf
  - Of a single leaf
  - Placed on light
  - Un-textured background with no other clutter
- They employ a binary RBF SVM classier applied to gist features (by LEAR).

Segmentation

Leaf Shape is Distinctive



Use curvatures at many scales to distinguish leaves!

Segmentation in HSV Colorspace

In Saturation-Value Space

Original Image
(On white background)

We do this by estimating foreground and background color distributions
in the saturation-value space of the SV colorspace.

Segmentation in HSV Colorspace

In Saturation-Value Space

Pixel-wise sum of two Gaussian Image Model:

$$\boxed{p(x|\Theta)} = \boxed{\sum_{k=1}^{2}} \frac{1}{2} \boxed{p(x|\mu_k}, \boxed{\Sigma})$$

Pixel label, given parameters

{foreground, background}

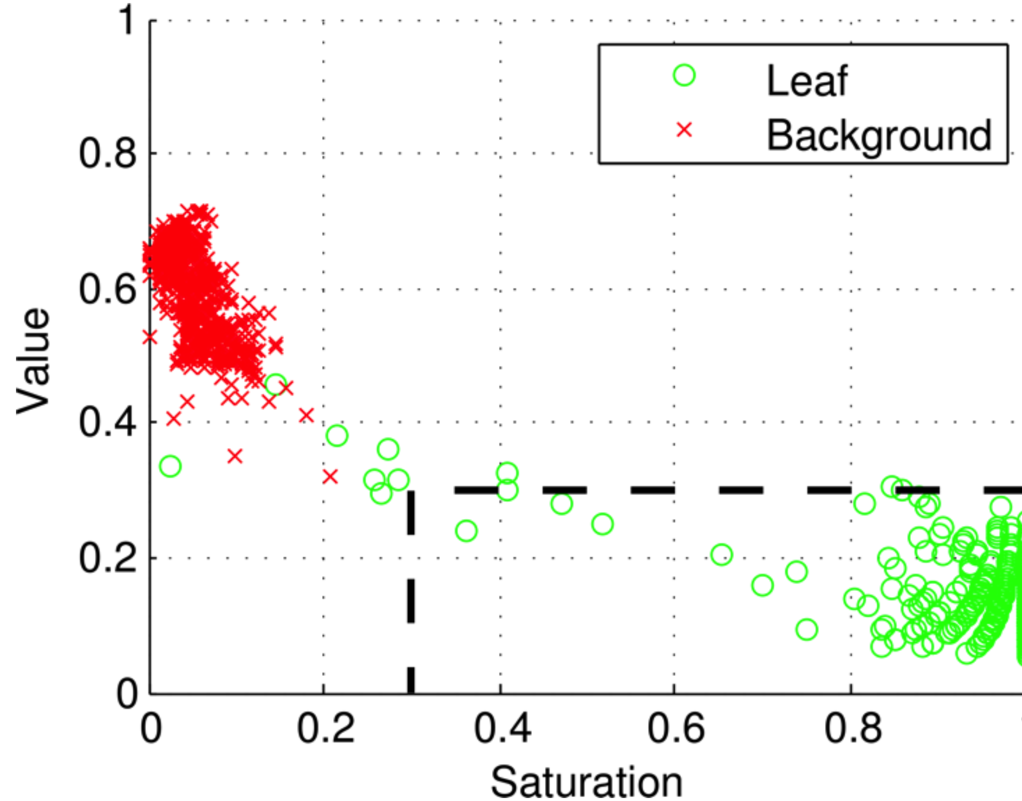
Mean color of class

Shared covariance

Hue is not useful because the background often has a greenish tinge due to reflections from the leaf or surrounding foliage

Segmentation in HSV Colorspace

Some difficulty with pine leaves→ they employed pixel weighting



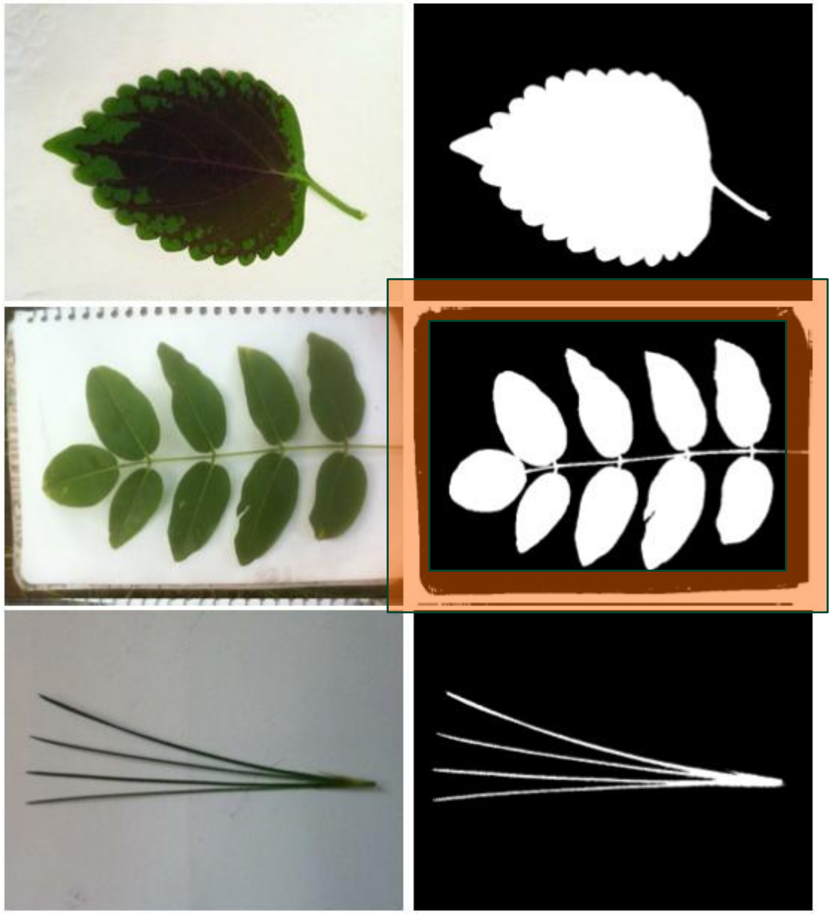
Segmented Image

In Saturation-Value Space

Expectation-Maximization (6-7 iterations)

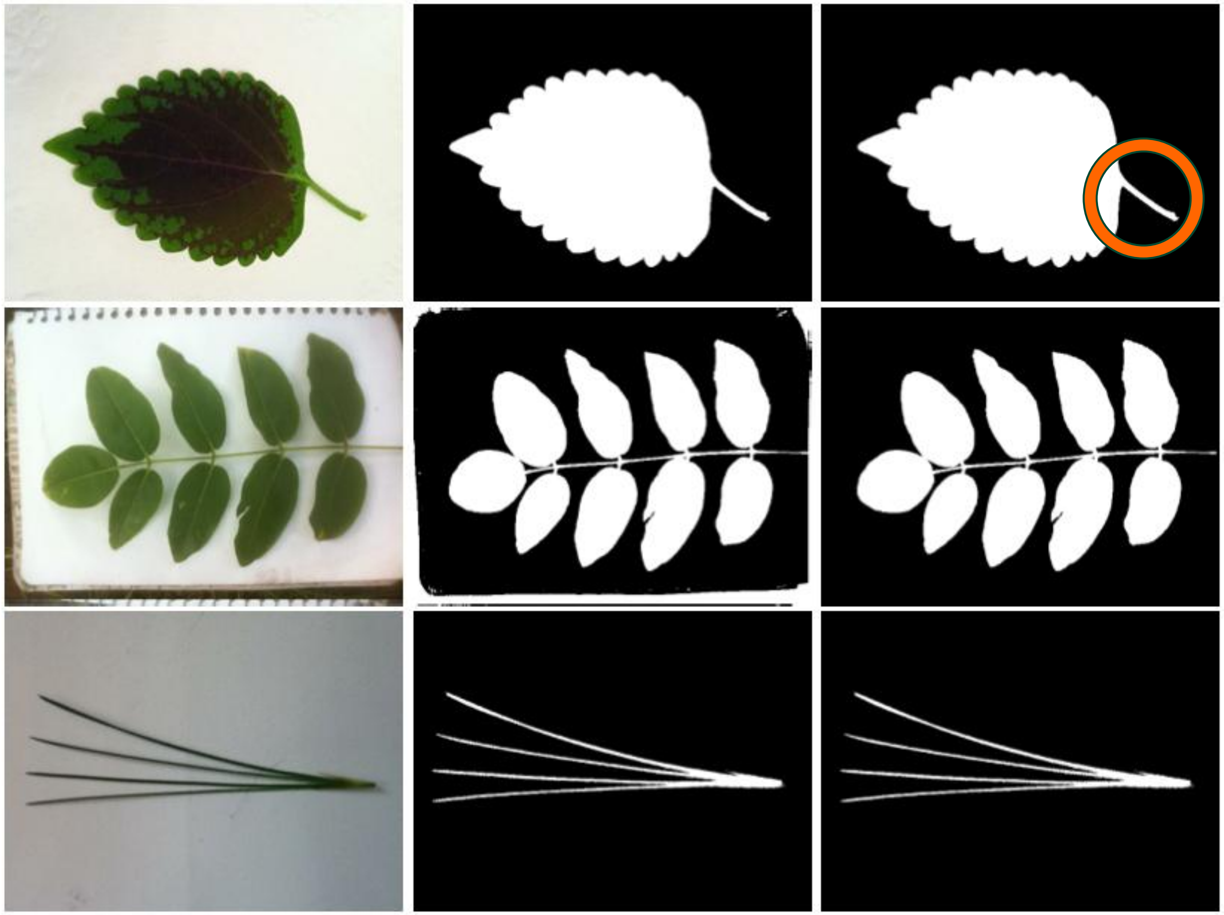
Use downsampeled version during EM

Segmentation Results



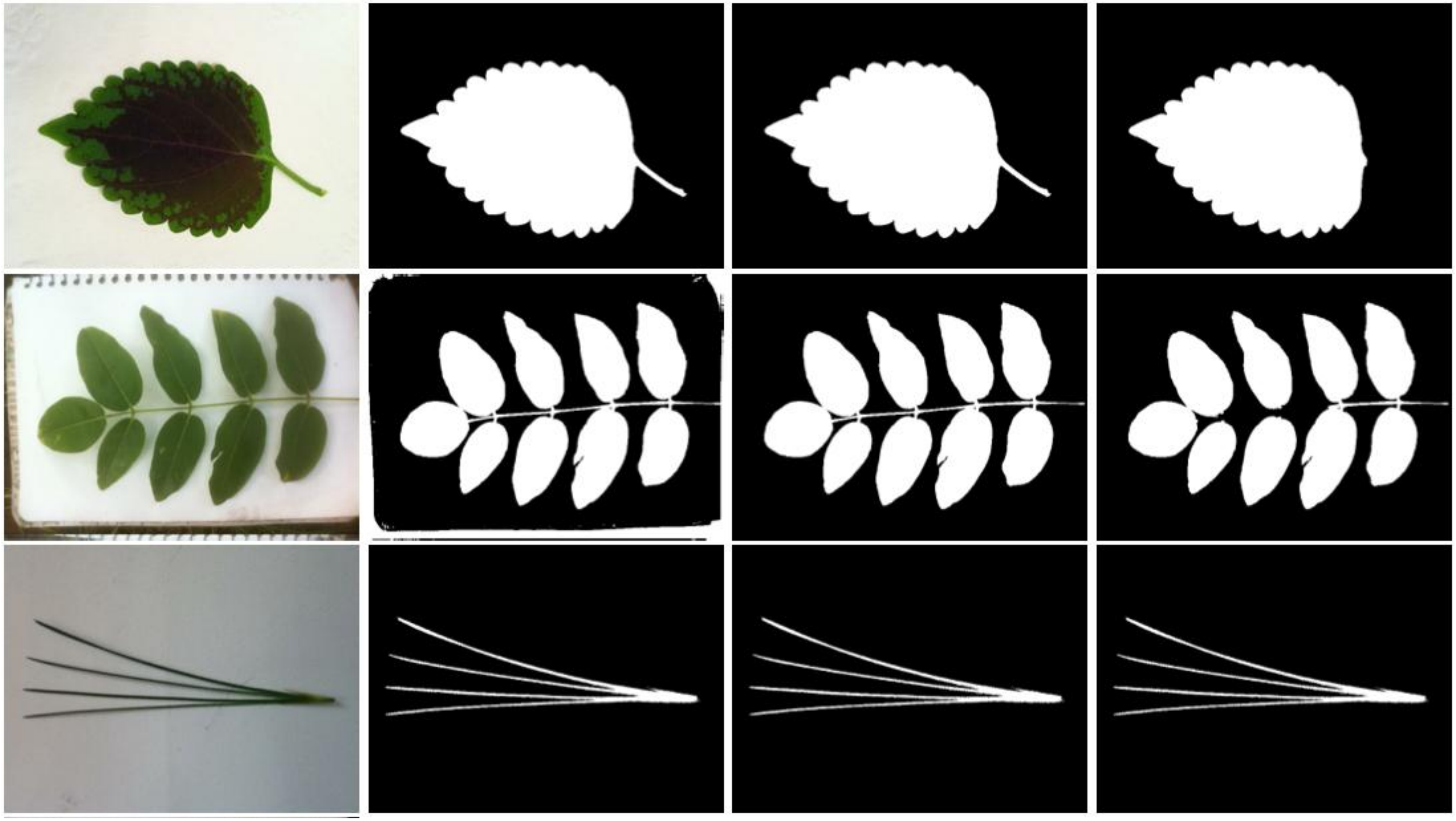
Original Initial Result (60 ms)

Segmentation Results



Original | Initial Result (60 ms) | Rem. False Positives (6 ms)

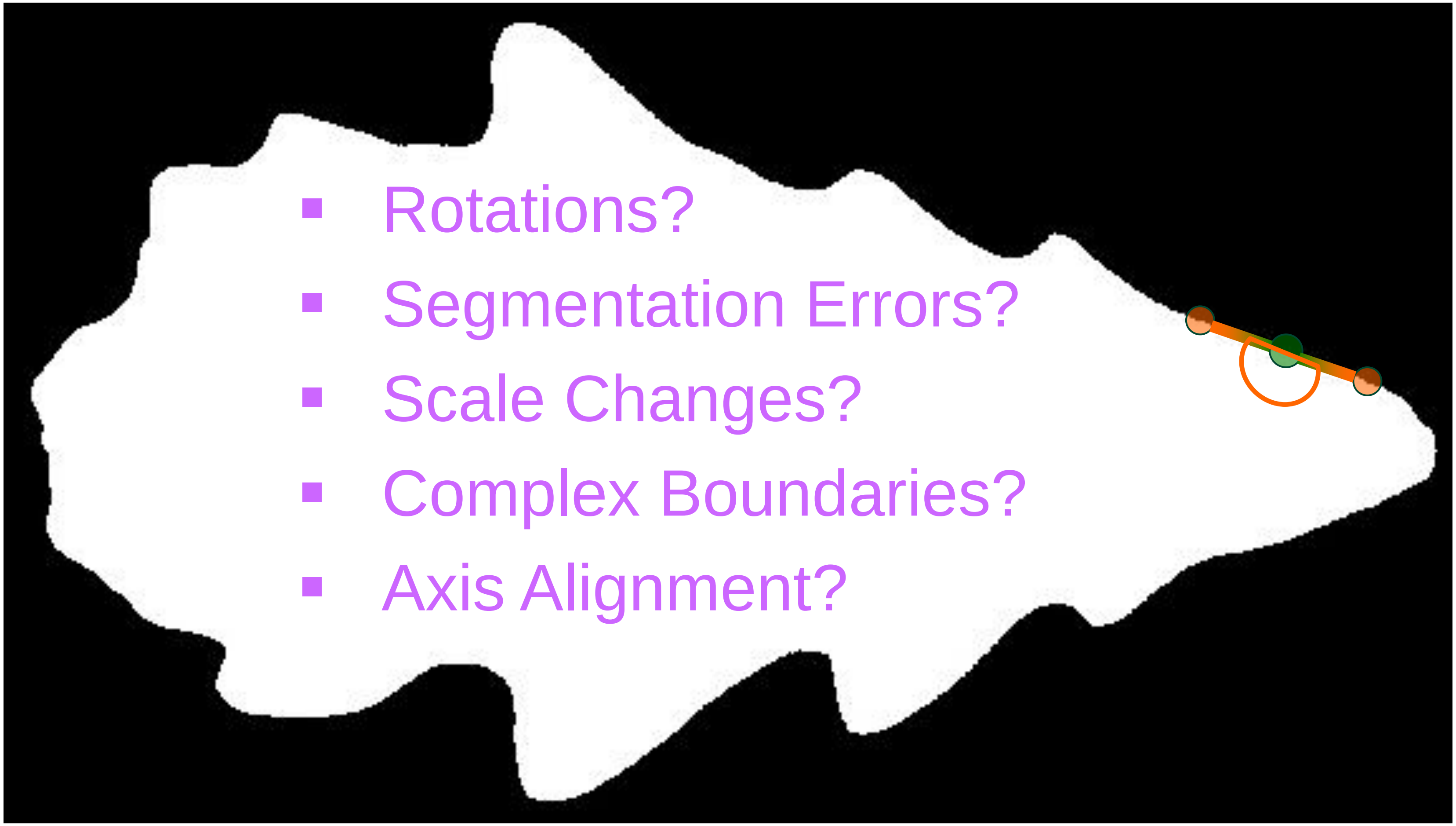
Segmentation Results



Original | Initial Result (60 ms) | Rem. False Positives (6 ms) | Remove Stem (36 ms)

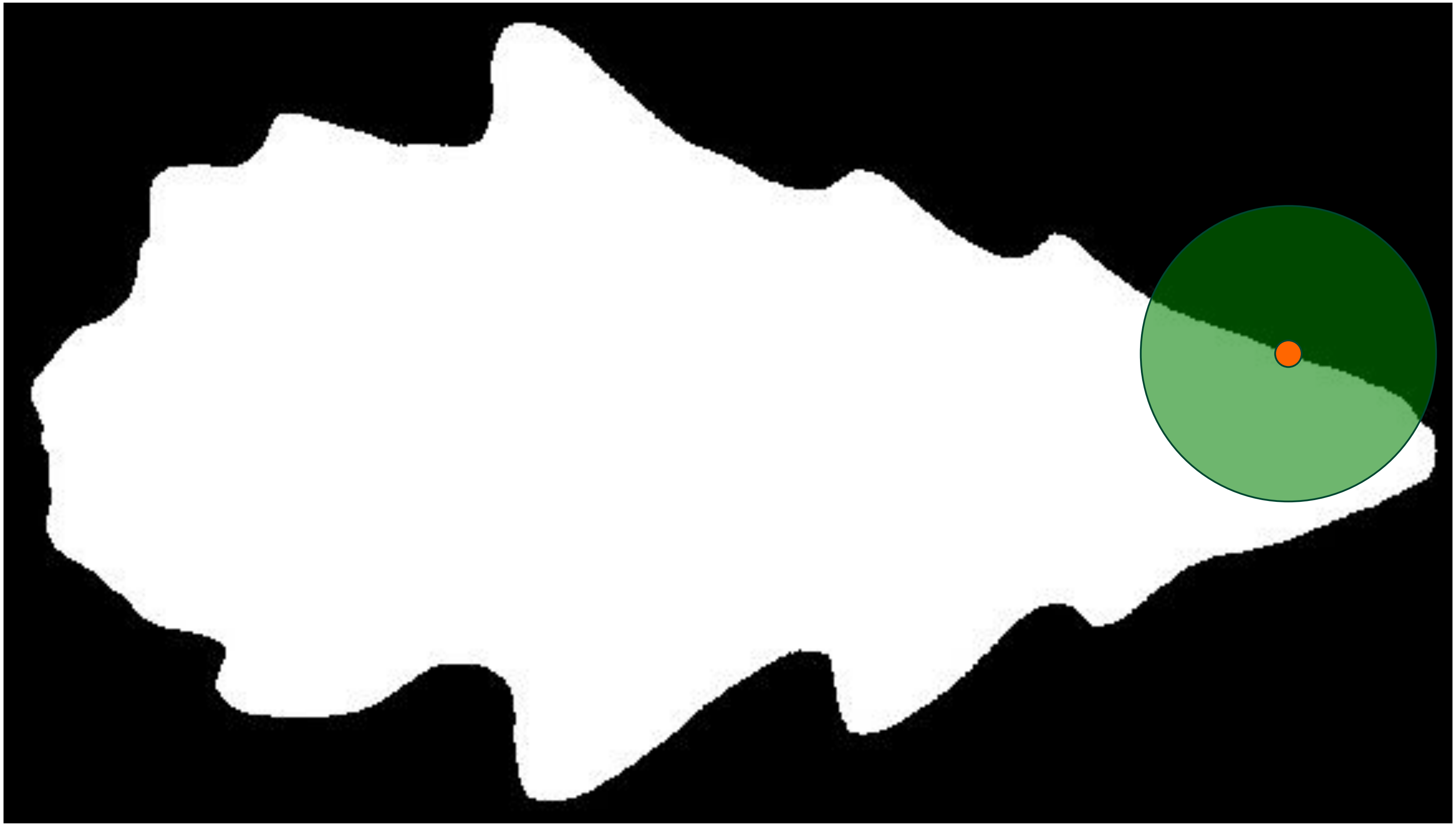
Morphological operations

Computing Curvature



Differential measures: not robust on discrete grids
And amplify noise

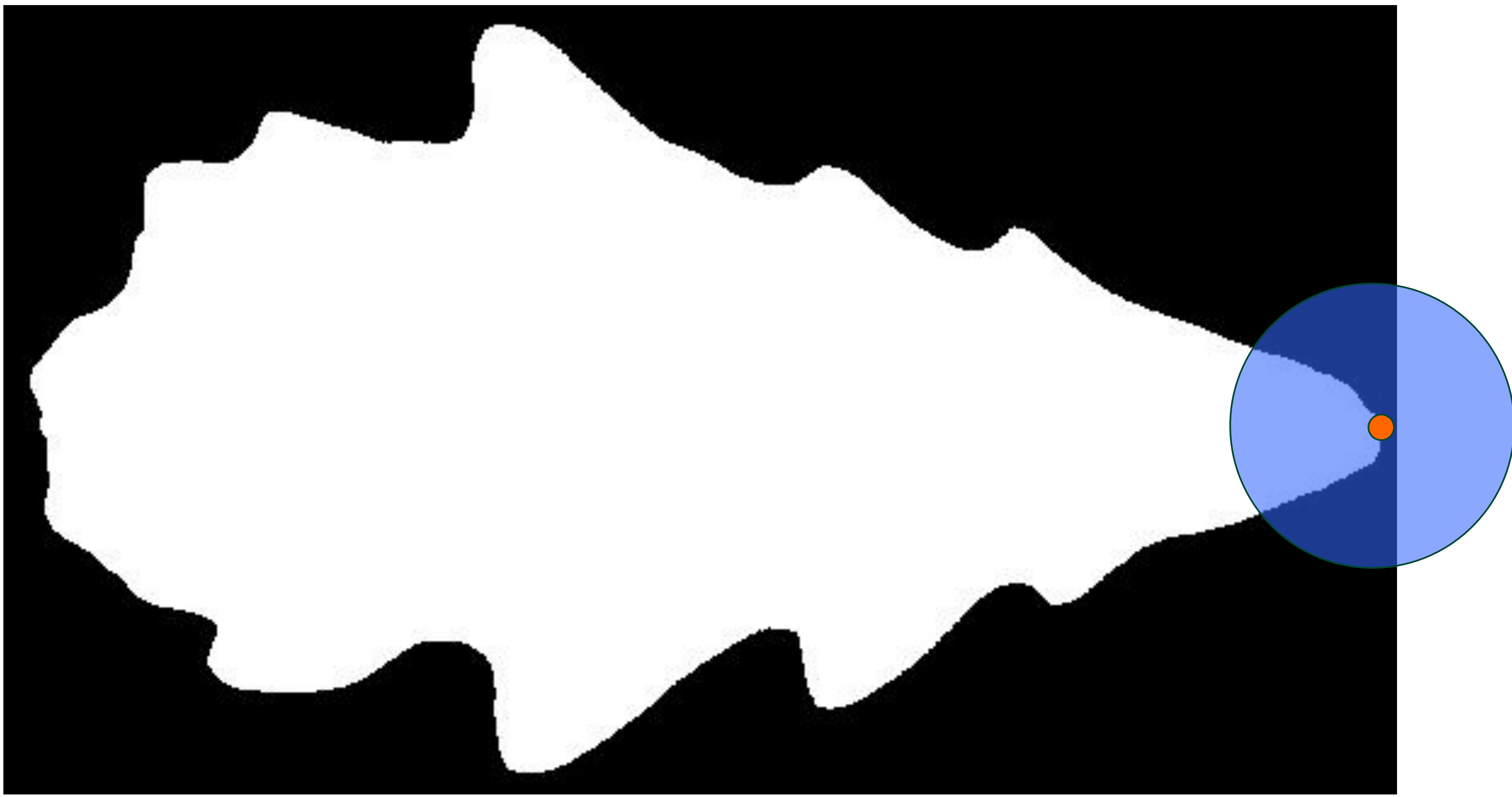
Curvature Using Integral Measures



Curvature = (white pixels in circle)/area = 0.5 (straight)

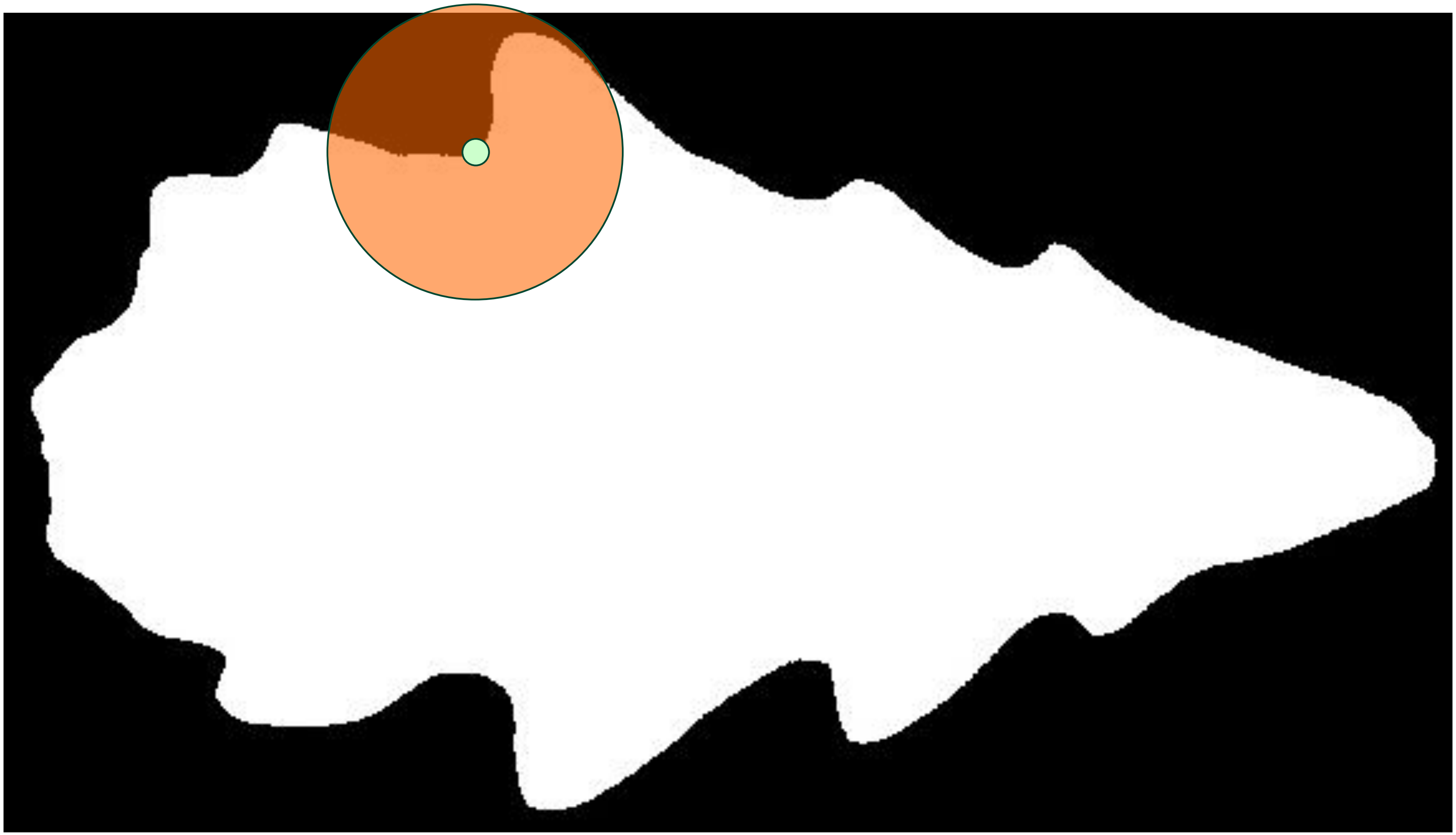
[Connolly 1986] [Manay et al. 2006] [Pottmann et al. 2009]

Curvature Using Integral Measures



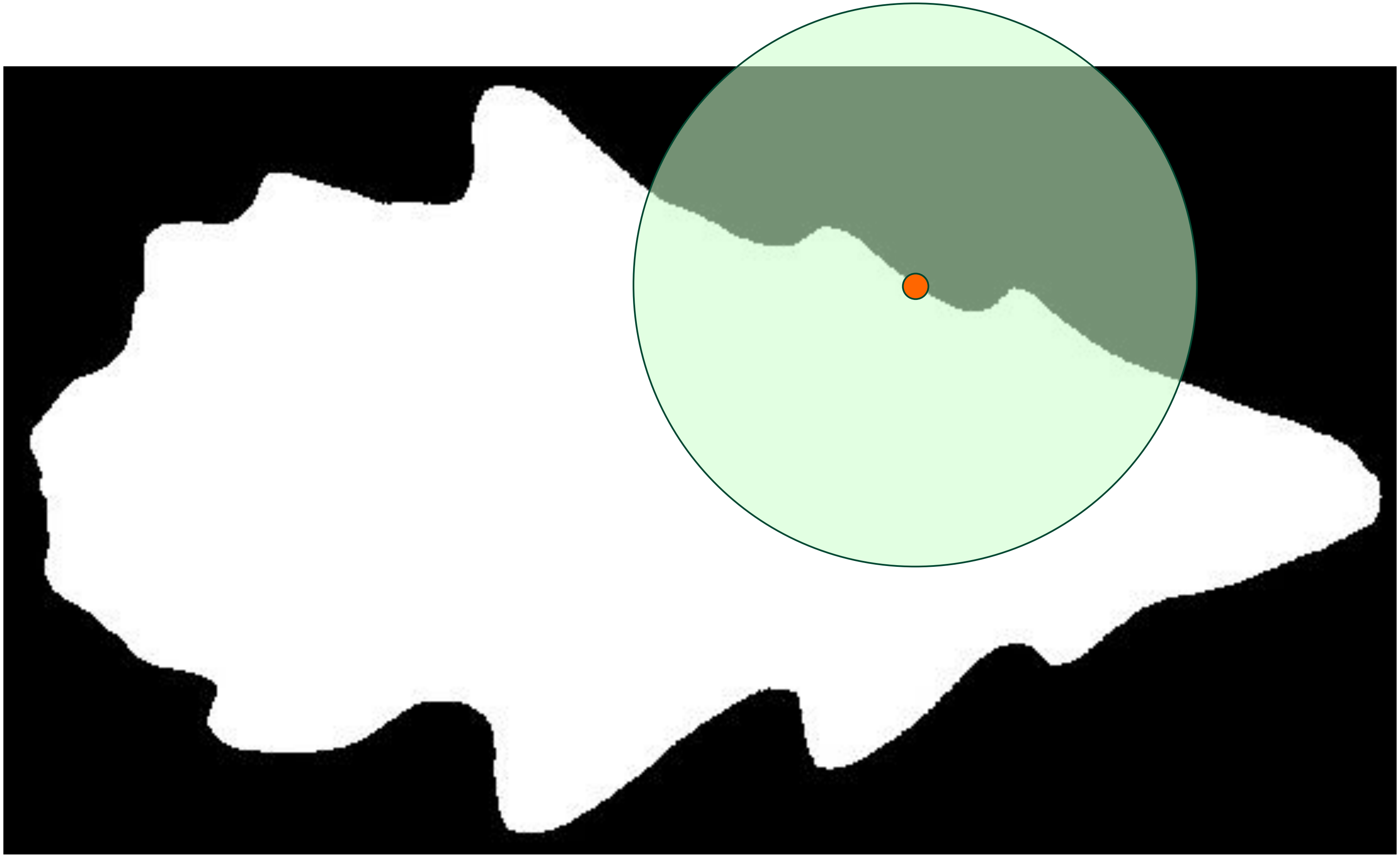
Curvature = (white pixels in circle)/area = 0.2 (convex)

Curvature Using Integral Measures



Curvature = (white pixels in circle)/area = 0.8 (concave)

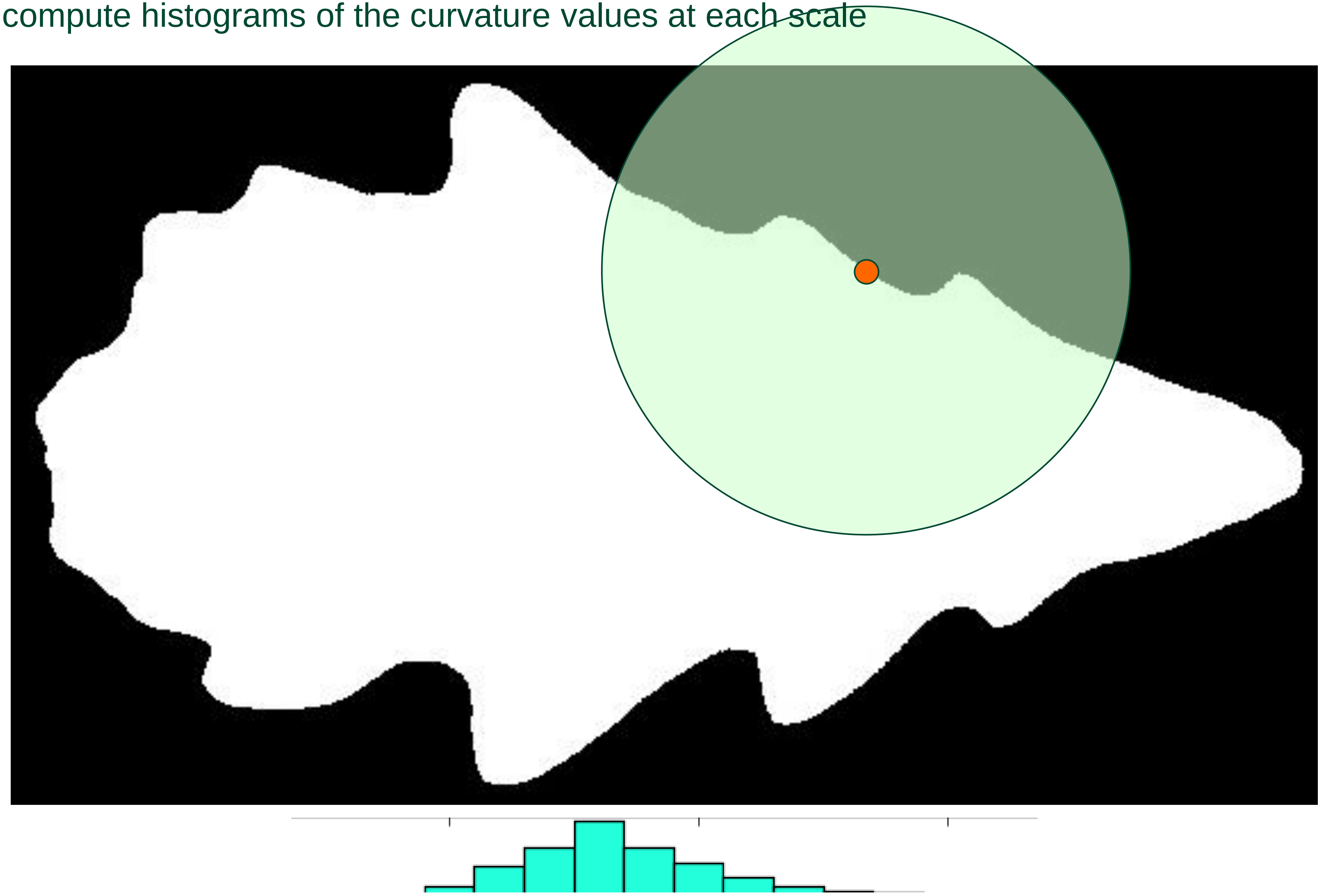
Curvature Using Integral Measures



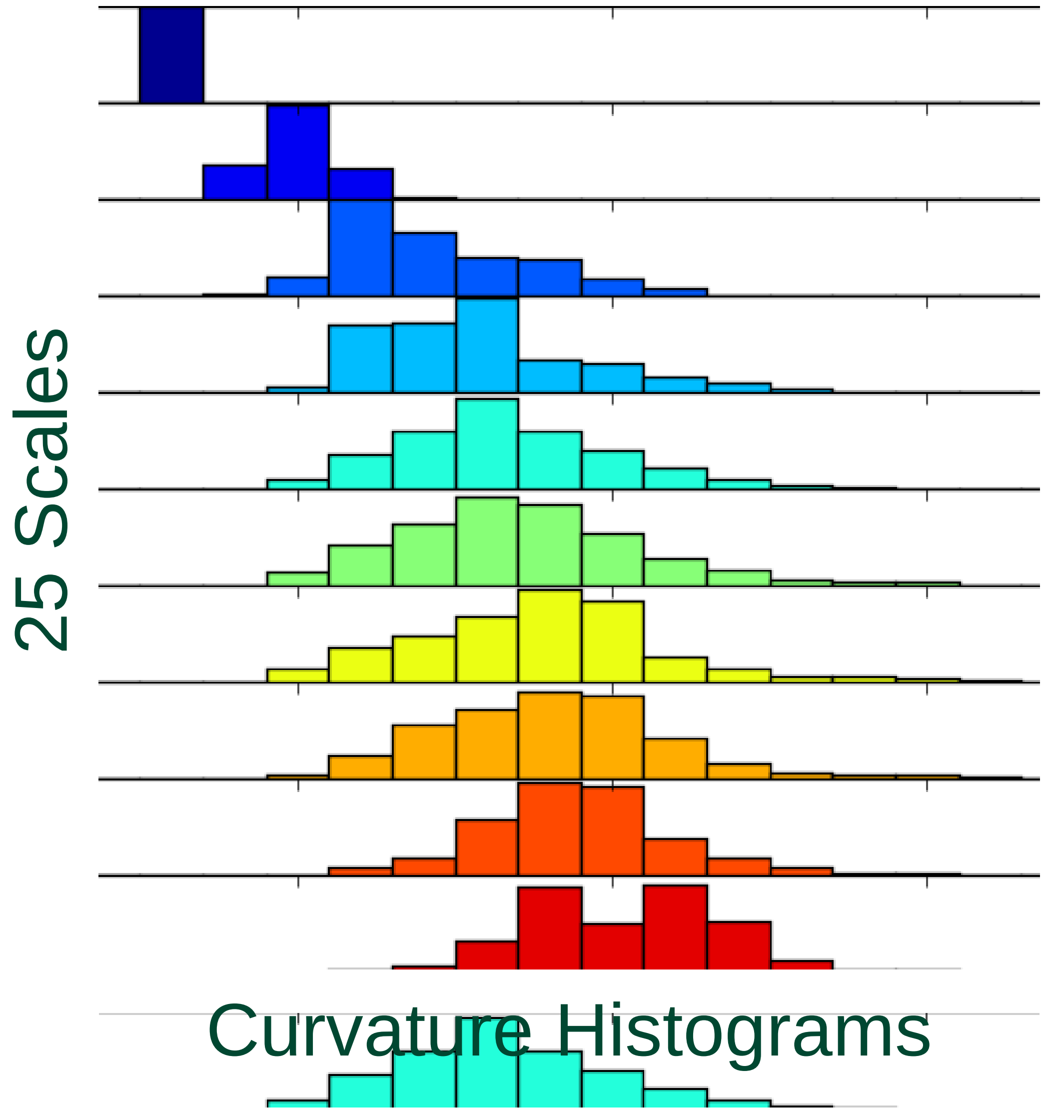
Curvature = (white pixels in circle)/area = 0.5

Build Histogram of Curvatures (HoCS)

compute histograms of the curvature values at each scale



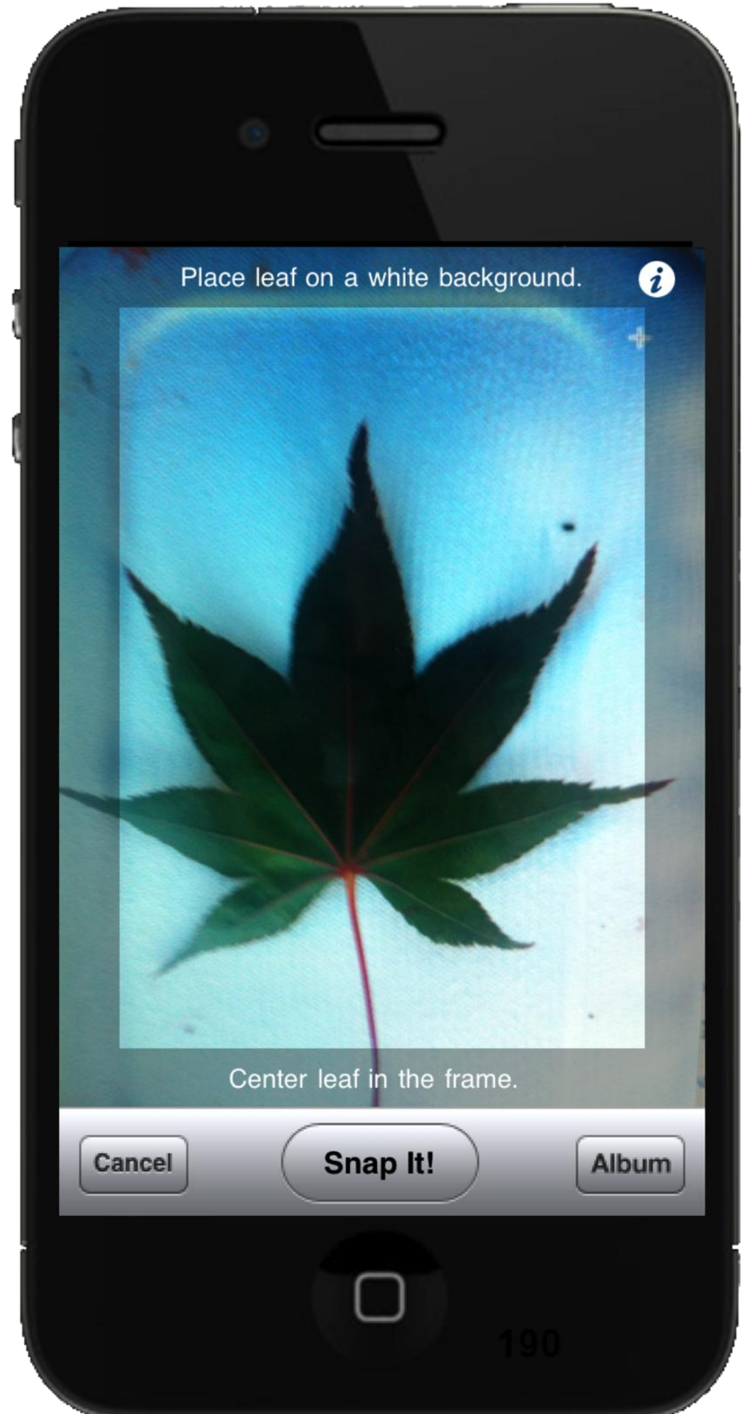
Histograms of Curvature over Scale



Total time: 0.11s

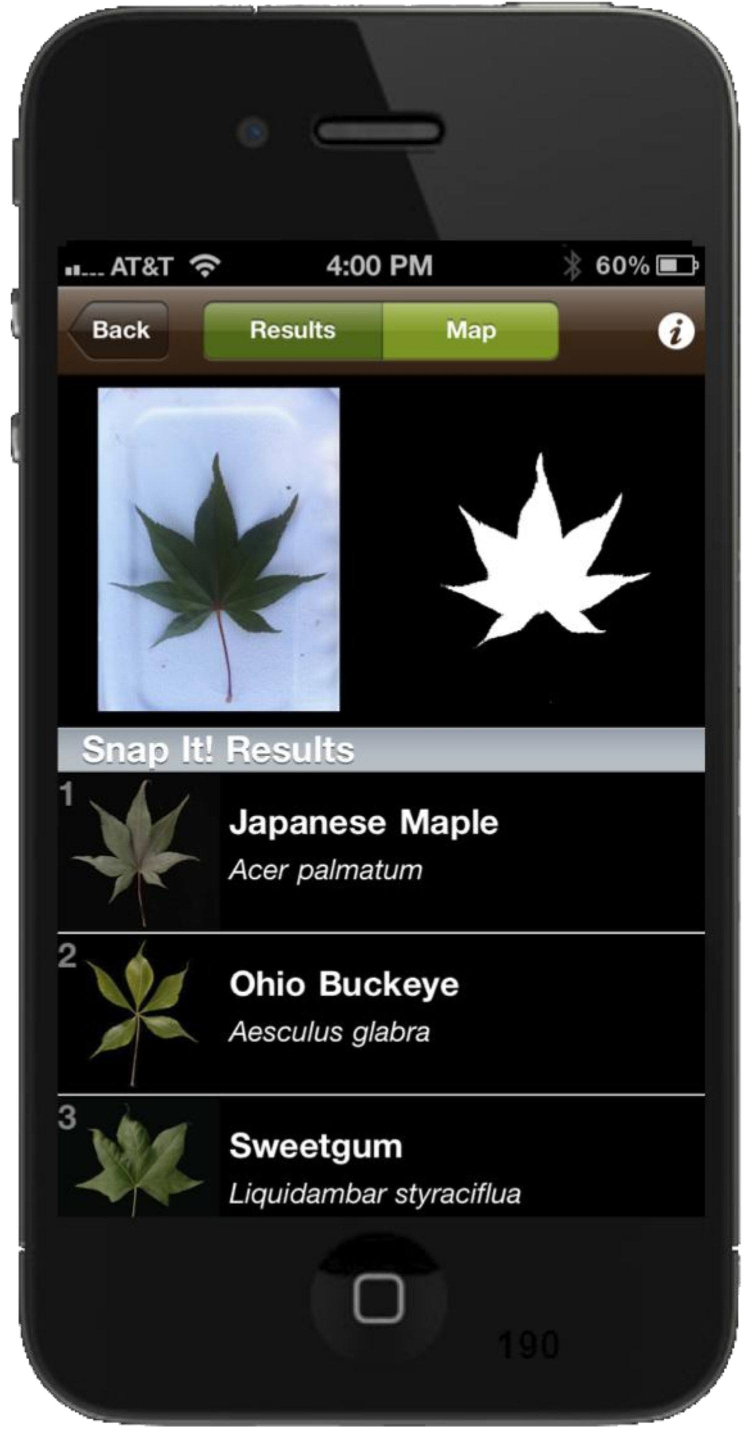
leafsnap

- Upload image
- Perform recognition (5.4s)
  - Nearest neighbor classifier
- Get ranked results

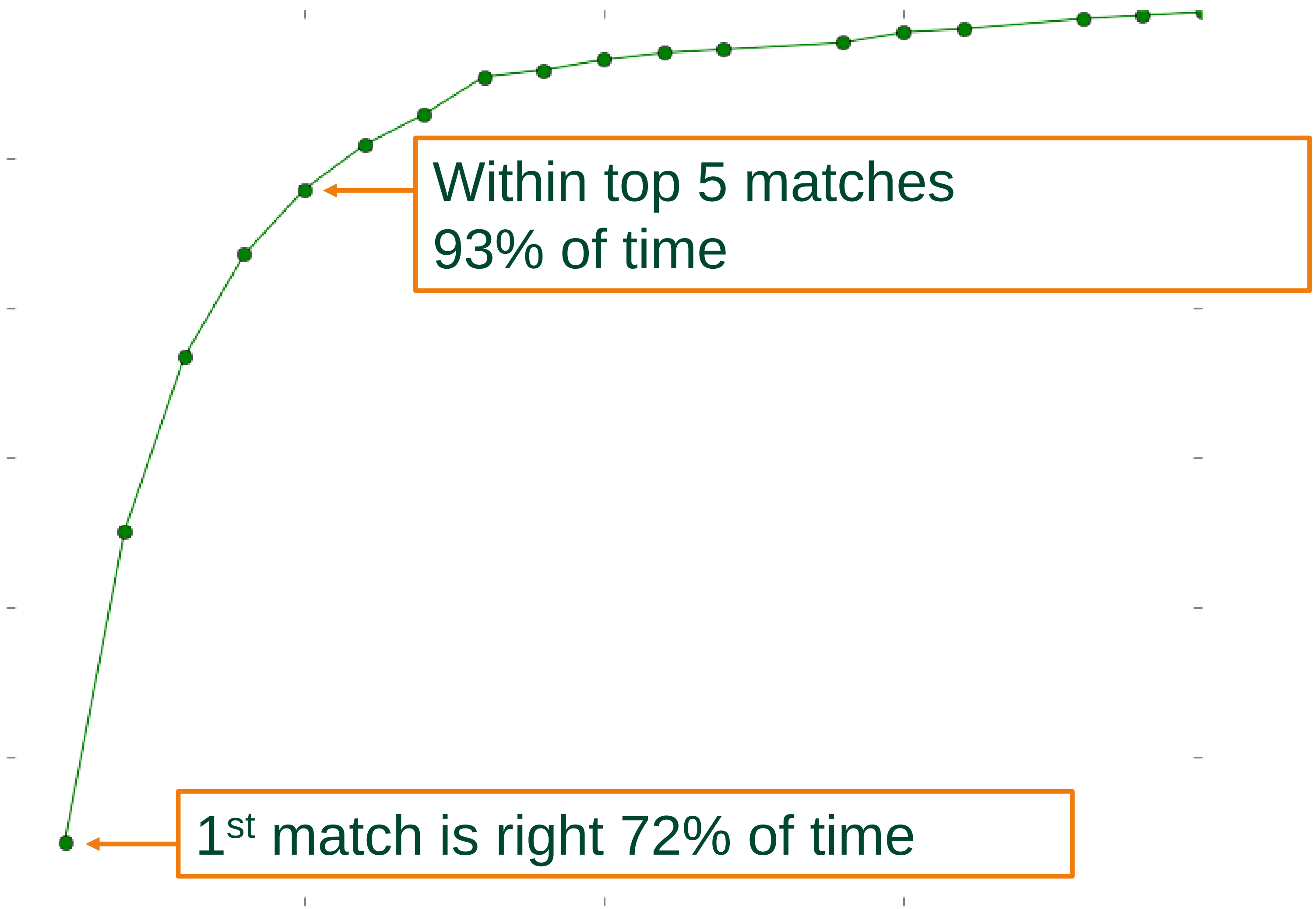


leafsnap

Which species is it?

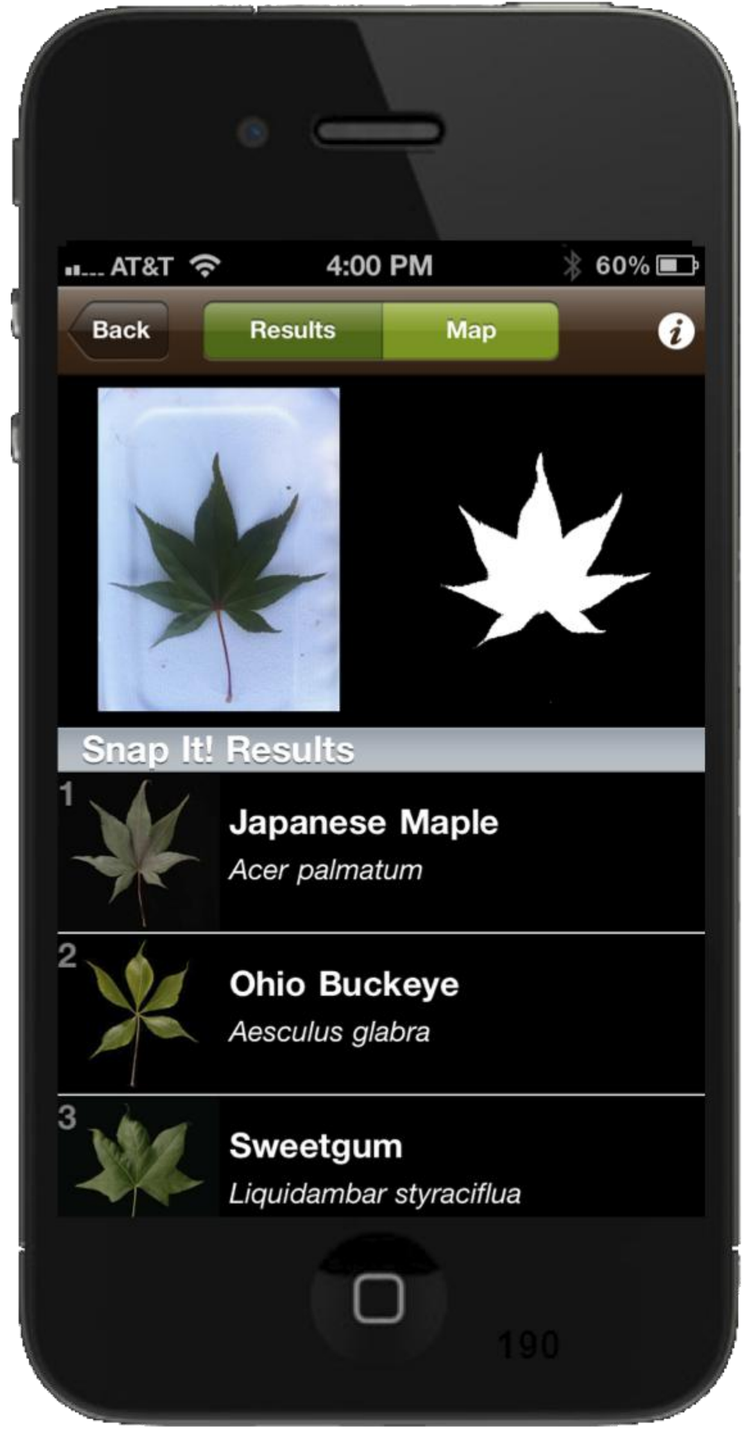


Accuracy on the Trees of the Northeast



leafsnap

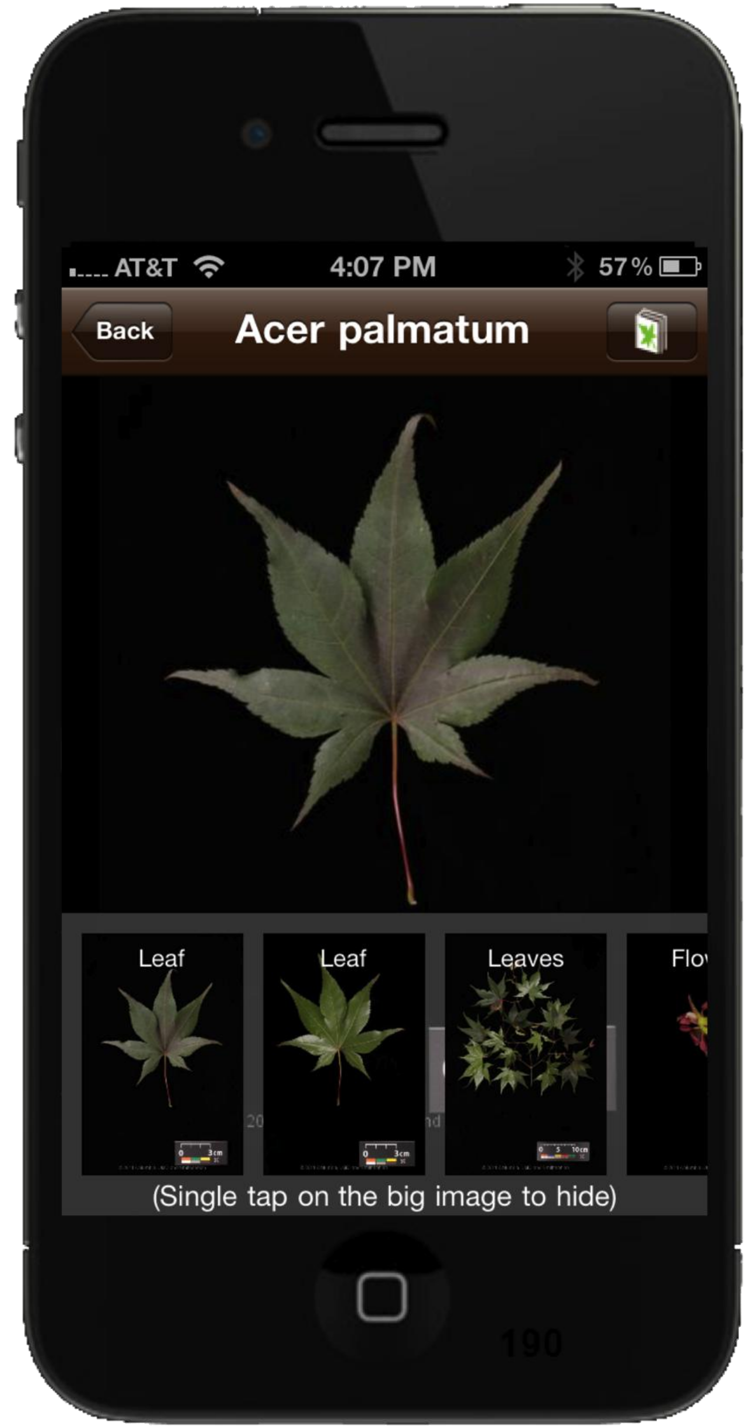
Which species is it?



leafsnap

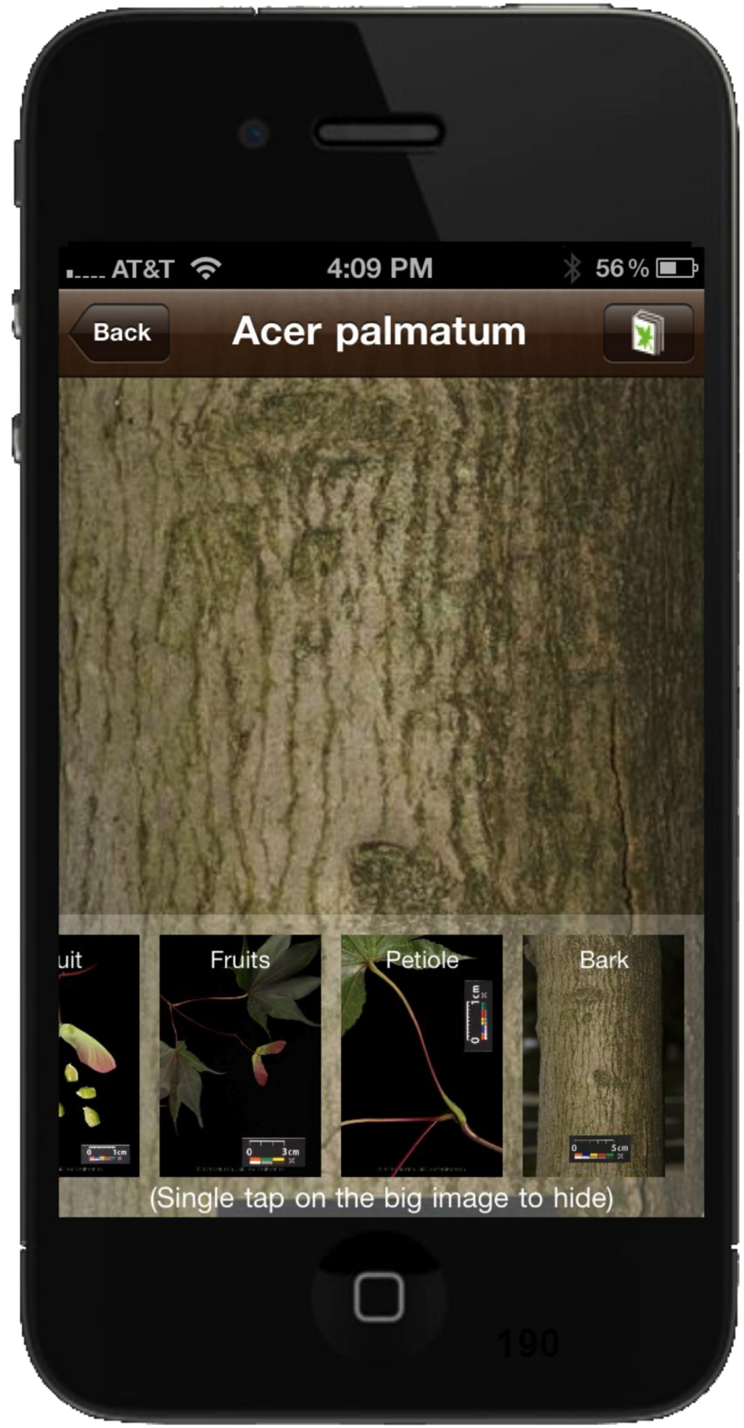
Which species is it?

- Explore the images in the app



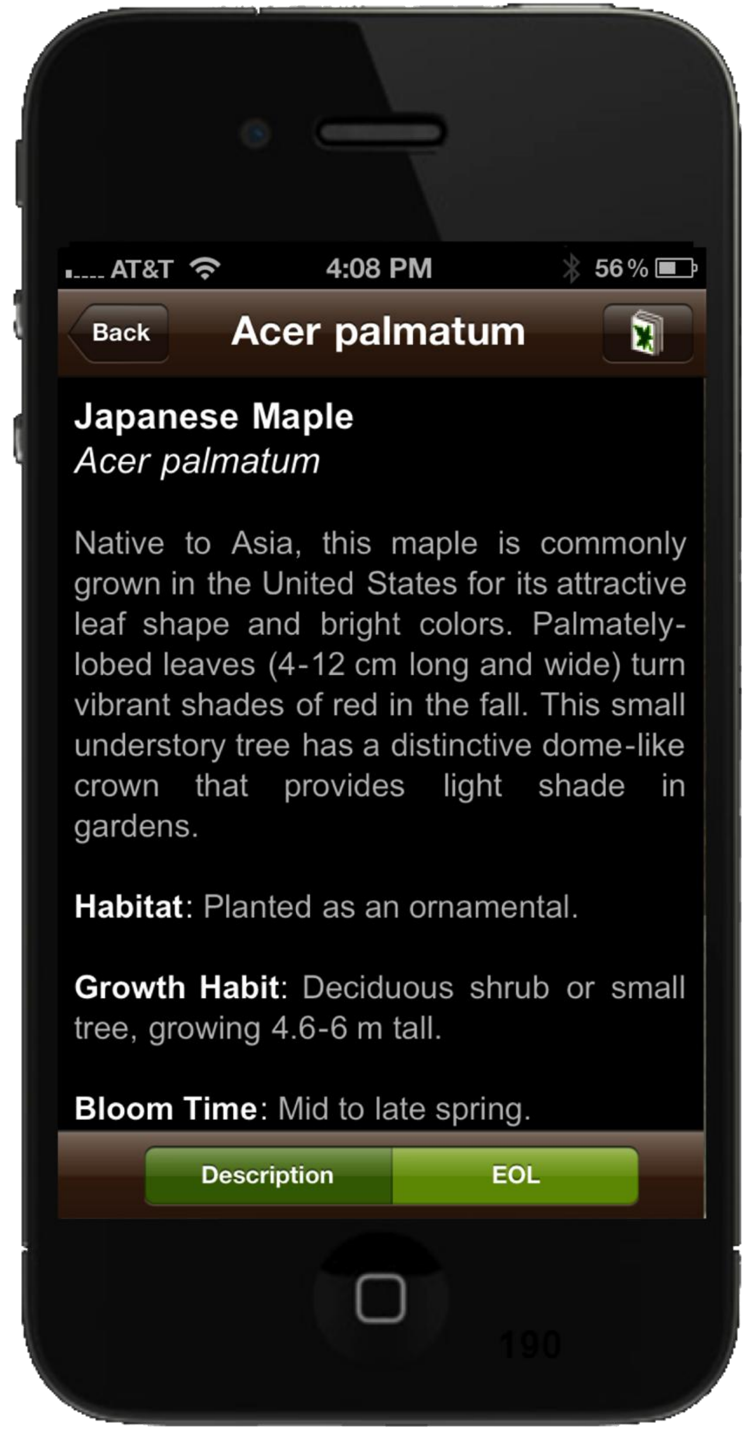
Which species is it?

- Explore the images in the app
  - Including the bark and stem



Which species is it?

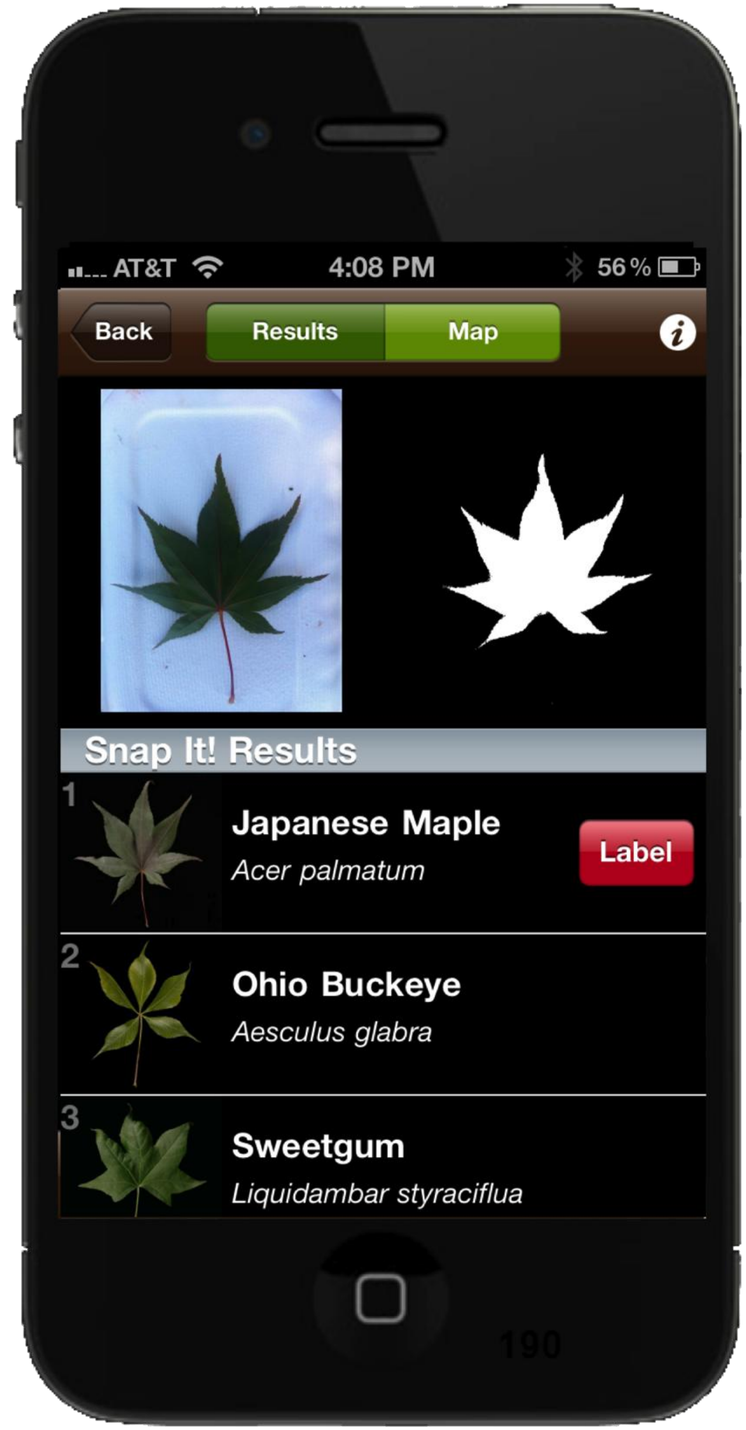
- Explore the images in the app
  - Including the bark and stem
- Read the text descriptions



leafsnap

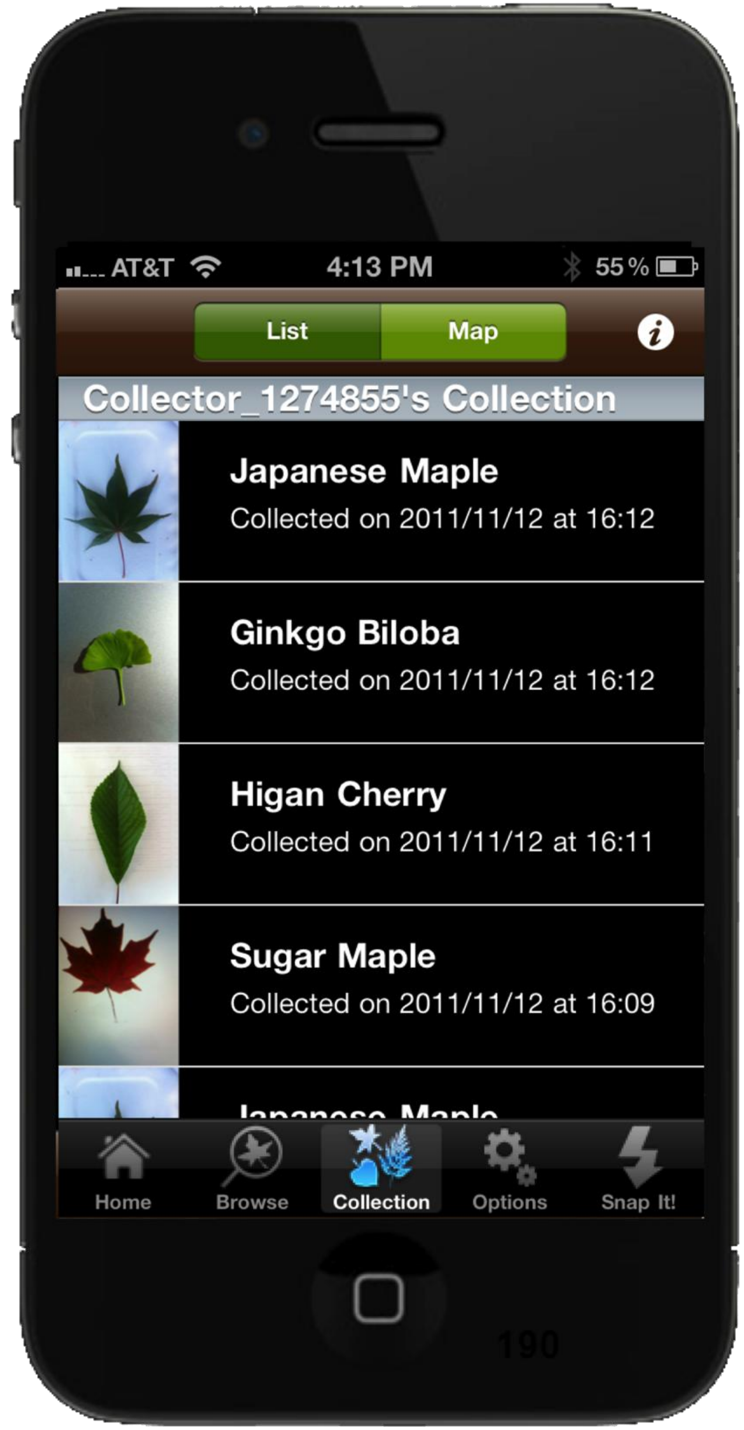
Which species is it?

- Explore the images in the app
  - Including the bark and stem
- Read the text descriptions
- Once confident, label it!



leafsnap

Look back through your collection…



leafsnap

Look back through your collection…

- Including the location!

leafsnap

Available on the App Store

Available on the iPad App Store

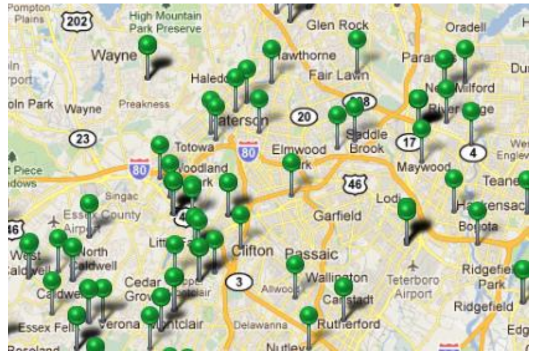
- Nearly 1 million downloads
  - 40k new users per month
  - 100k active users
- 1.7 million images taken
  - 100k new images/month
  - 100k users with > 5 images
- Users from all over the world
- Botanists, educators, kids, hobbyists, photographers, …

- very fast, suitable for use in an interactive application
- adapt to major sources of color variability such as lighting changes and natural variations in leaf color

- Nearly 1 million downloads
  - 40k new users per month
  - 100k active users
- 1.7 million images taken
  - 100k new images/month
  - 100k users with > 5 images
- Users from all over the world
- Botanists, educators, kids, hobbyists, photographers, …

Lots of Future Directions

Education
and Outreach



Tracking
Biodiversity

Discovering
New Species




leafsnap

- Apps Dataset Code available at leafsnap.com
- Plant images in apps taken by FindingSpecies.org

volunteers

Thank You