

# 数据驱动的图像内容编辑

计算机系 张松海

# 主要内容

- 基于互联网图像的颜色空间降维技术
  - Data-Driven Color Manifolds. @SIGGRAPH 2015. Chuong H. Nguyen et. al.
- 基于互联网图像的辅助绘制技术
  - ShadowDraw: Real-Time User Guidance for Freehand Drawing @SIGGRAPH 2011. Yong Jae Lee et. al.
- 基于互联网图像的场景漫游
  - Photo tourism: Exploring photo collections in 3D. @SIGGRAPH 2006. Noah Snavely.et.al.
- 基于互联网数据的图像补全
  - Scene Completion Using Millions of Photographs. @SIGGRAPH 2007. J. Hays & A. A. Efros.
- 基于互联网的图像检索与融合
  - Sketch2Photo: Internet Image Montage. @SIGGRAPH ASIA 2009. Tao Chen et. al.

# 主要内容

- 基于互联网图像的颜色空间降维技术
  - Data-Driven Color Manifolds. @SIGGRAPH 2015. Chuong H. Nguyen et. al.
- 基于互联网图像的辅助绘制技术
  - ShadowDraw: Real-Time User Guidance for Freehand Drawing @SIGGRAPH 2011. Yong Jae Lee et. al.
- 基于互联网图像的场景漫游
  - Photo tourism: Exploring photo collections in 3D. @SIGGRAPH 2006. Noah Snavely.et.al.
- 基于互联网数据的图像补全
  - Scene Completion Using Millions of Photographs. @SIGGRAPH 2007. J. Hays & A. A. Efros.
- 基于互联网的图像检索与融合
  - Sketch2Photo: Internet Image Montage. @SIGGRAPH ASIA 2009. Tao Chen et. al.

# Overview

Supplemental video (with audio narration)

## Data-Driven Color Manifolds

ACM Transactions on Graphics 34(1)

<sup>1,2</sup>Chuong Nguyen, <sup>1,2</sup>Tobias Ritschel, <sup>2</sup>Hans-Peter Seidel

<sup>1</sup>MMCI / Saarland University

<sup>2</sup>MPI Informatik

# Introduction

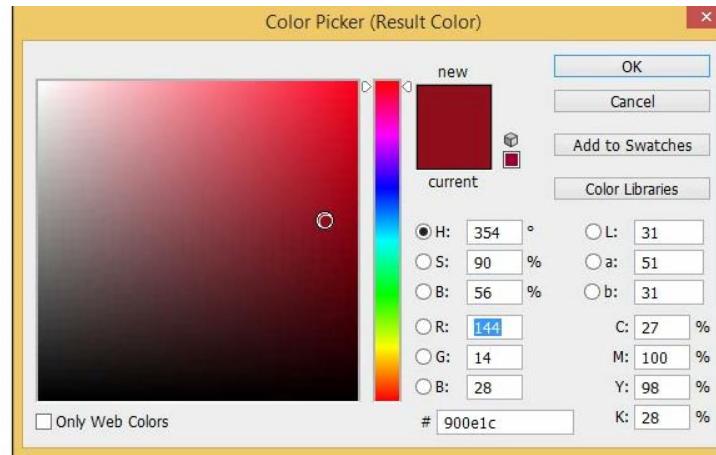


Source



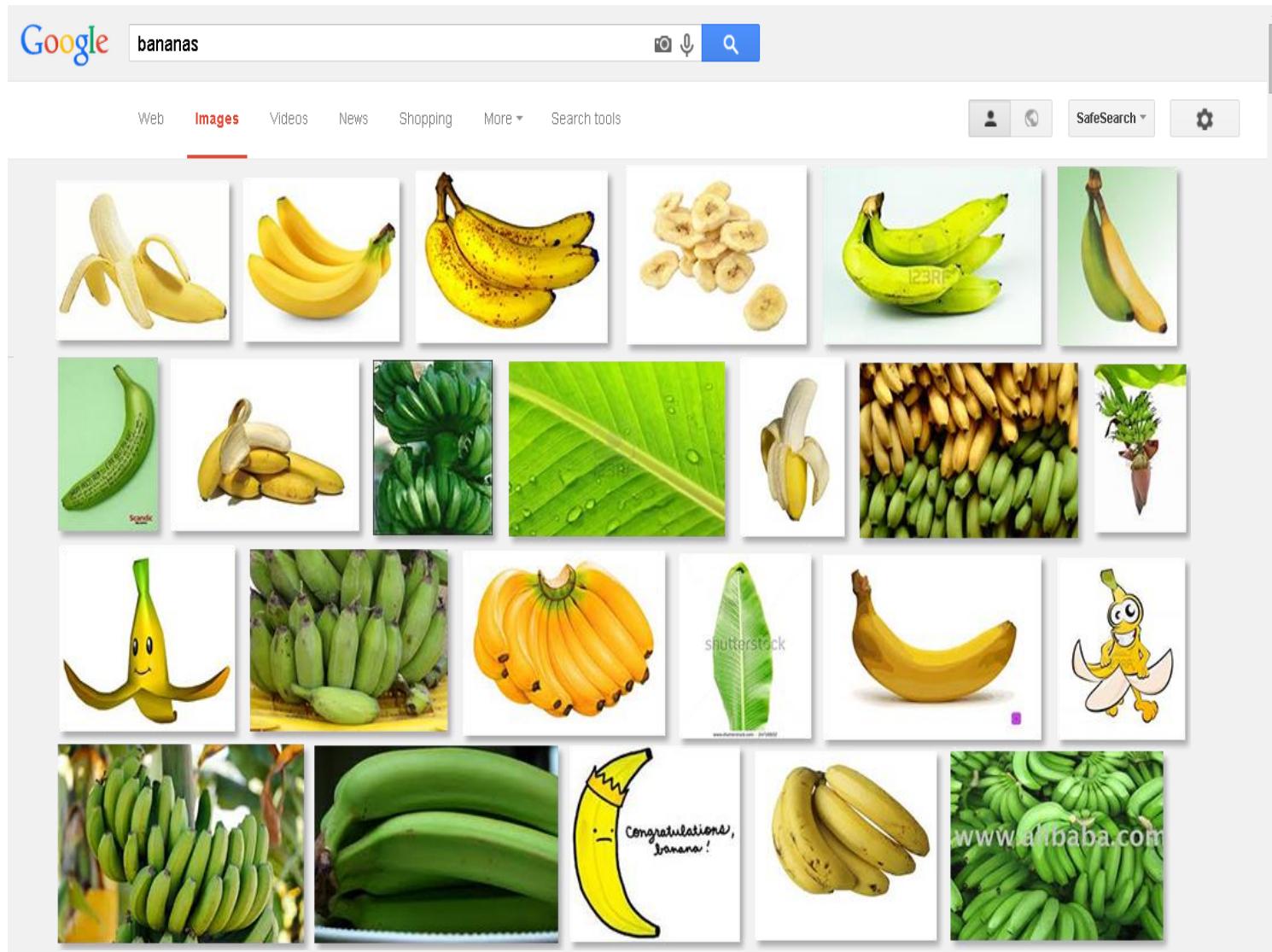
Target

# Color editing

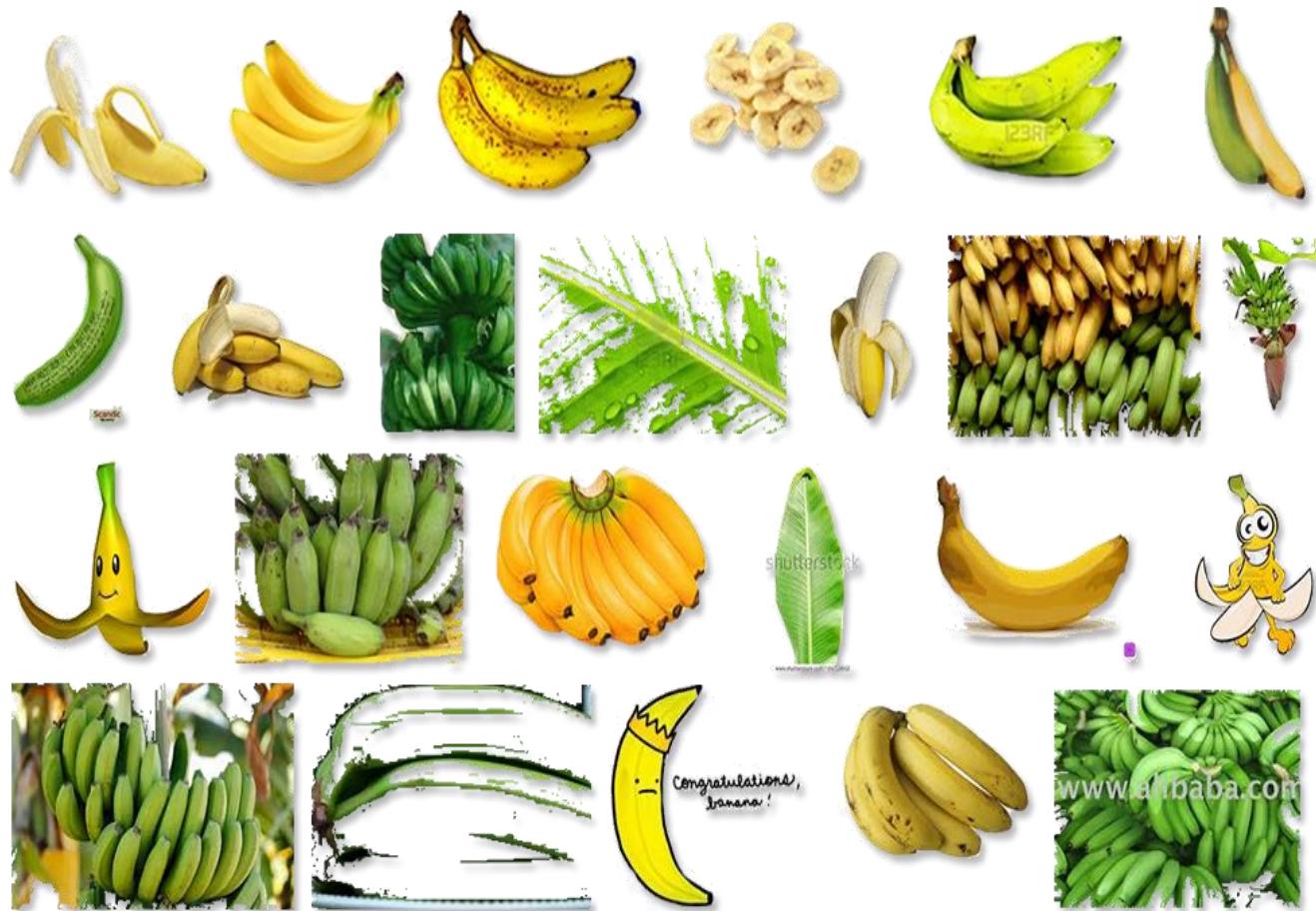


Navigation in a 3D space using a 2D+1D interface

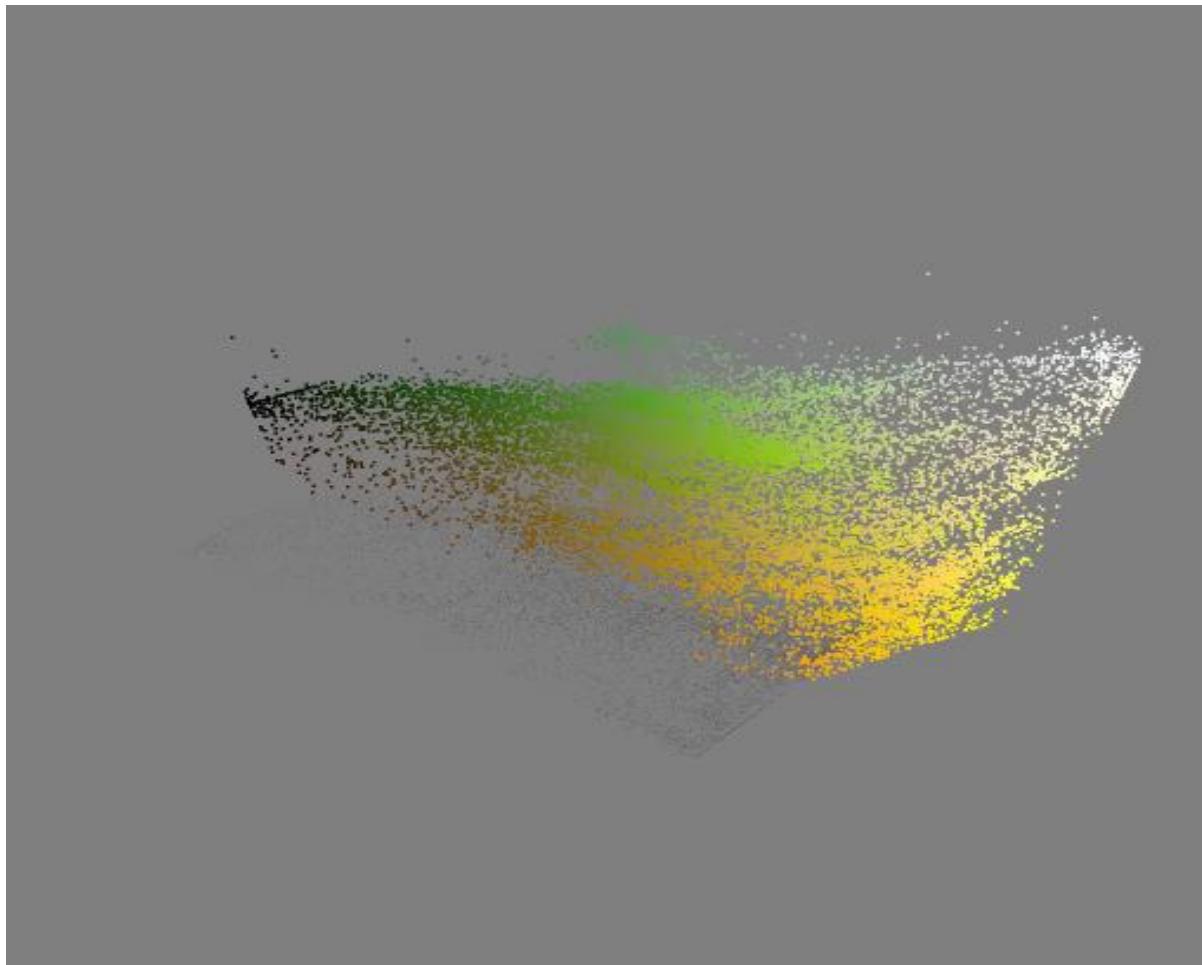
# Data Acquisition



# Background Removal



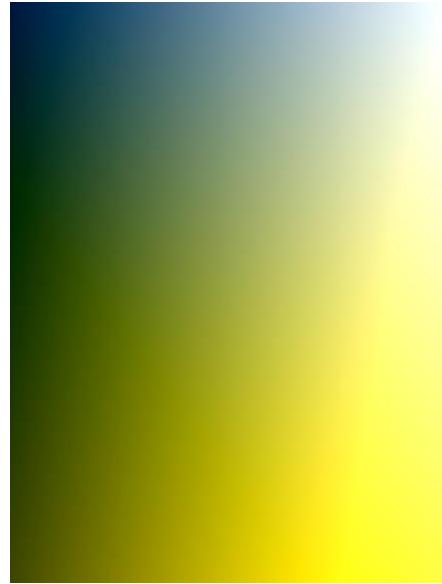
# “Banana” – Point Cloud



# Dimensionality Reduction



1D

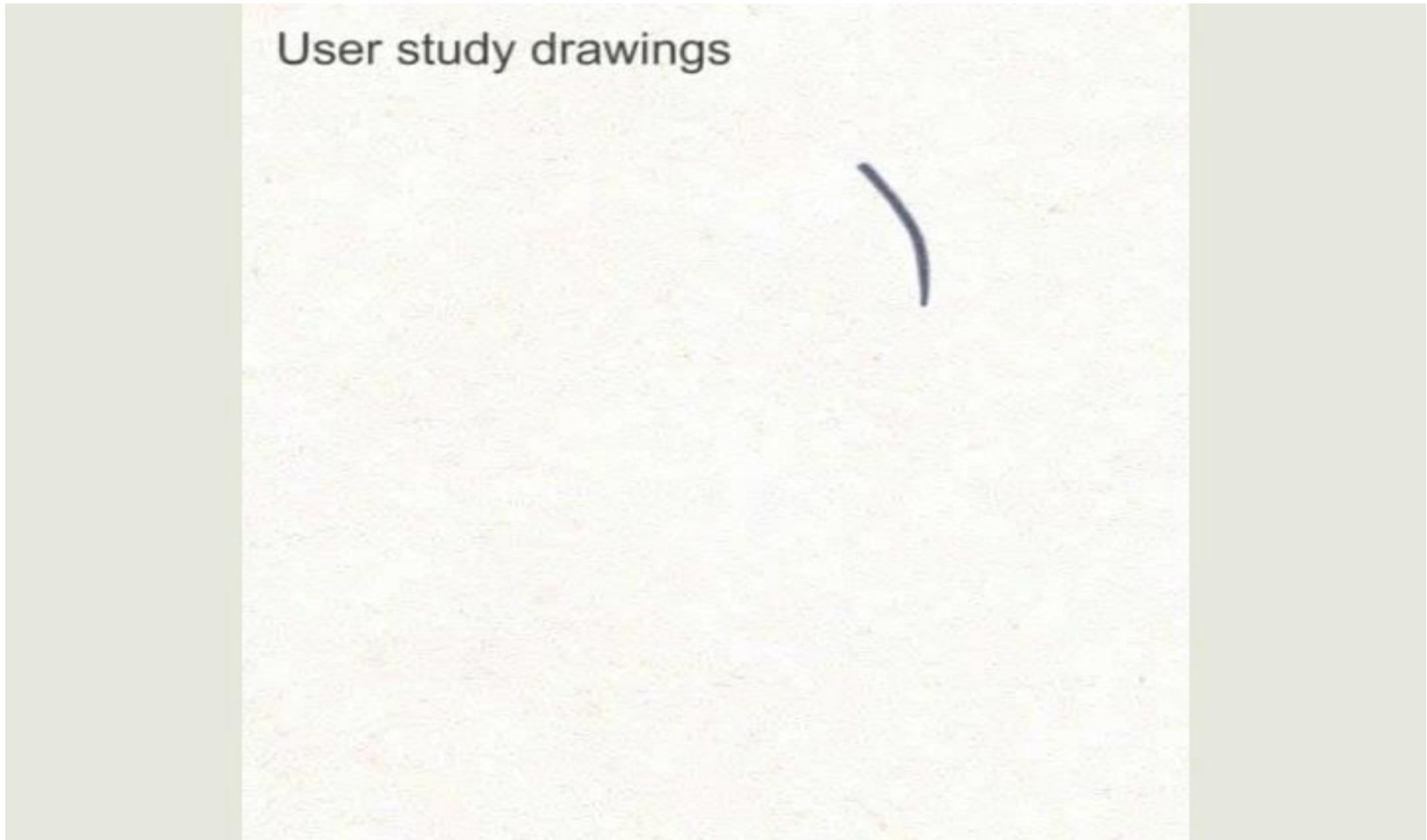


2D

# 主要内容

- 基于互联网图像的颜色空间降维技术
  - Data-Driven Color Manifolds. @SIGGRAPH 2015. Chuong H. Nguyen et. al.
- 基于互联网图像的辅助绘制技术
  - ShadowDraw: Real-Time User Guidance for Freehand Drawing @SIGGRAPH 2011. Yong Jae Lee et. al.
- 基于互联网图像的场景漫游
  - Photo tourism: Exploring photo collections in 3D. @SIGGRAPH 2006. Noah Snavely.et.al.
- 基于互联网数据的图像补全
  - Scene Completion Using Millions of Photographs. @SIGGRAPH 2007. J. Hays & A. A. Efros.
- 基于互联网的图像检索与融合
  - Sketch2Photo: Internet Image Montage. @SIGGRAPH ASIA 2009. Tao Chen et. al.

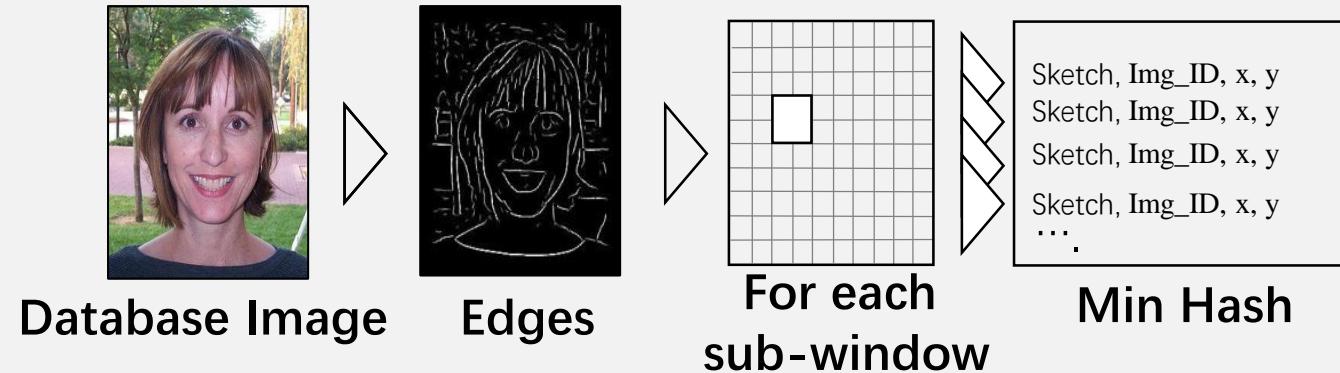
# Our Idea: ShadowDraw



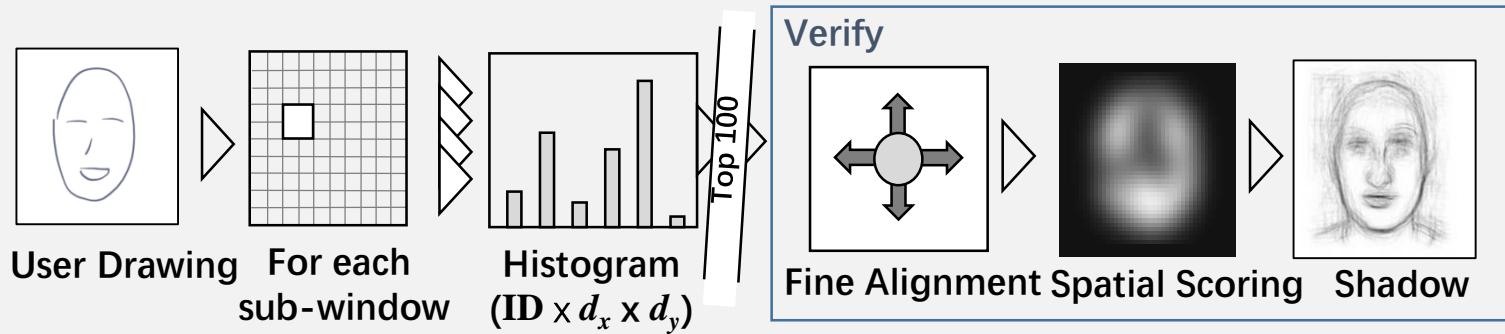
**ShadowDraw**

# Overview

## Database (offline)



## Query time (online)

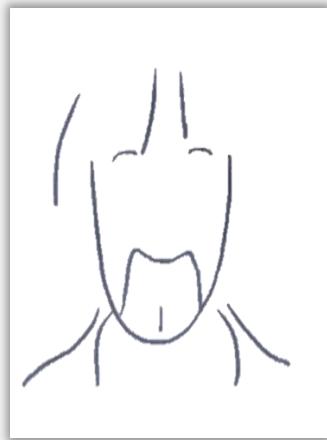


# Database

## Sample database images



# Blending weights



Pen Strokes



\*



+



\*



Top matches

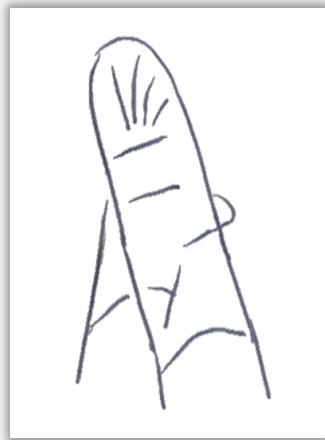
Weights

=

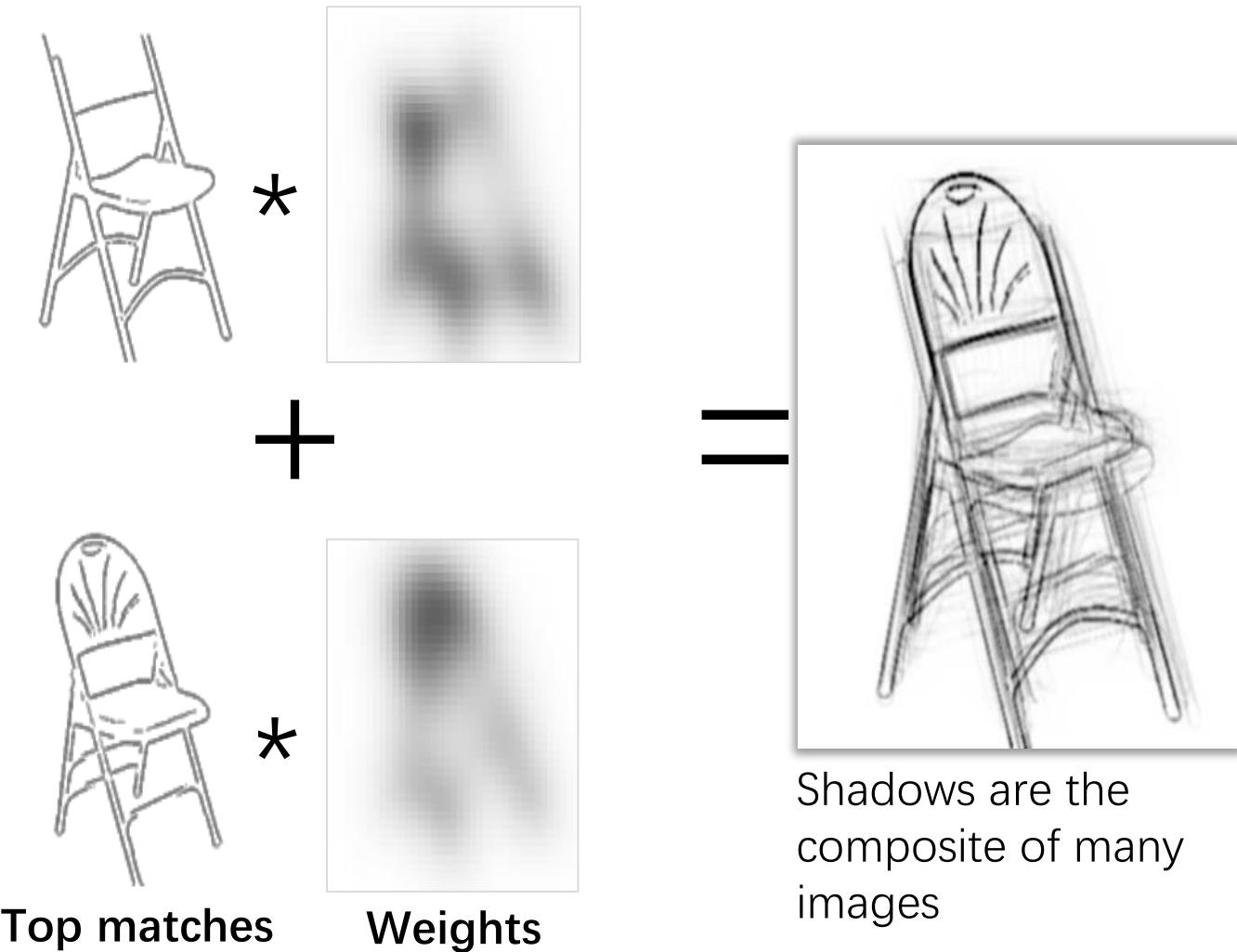


Shadows are the  
composite of many  
images

# Blending weights

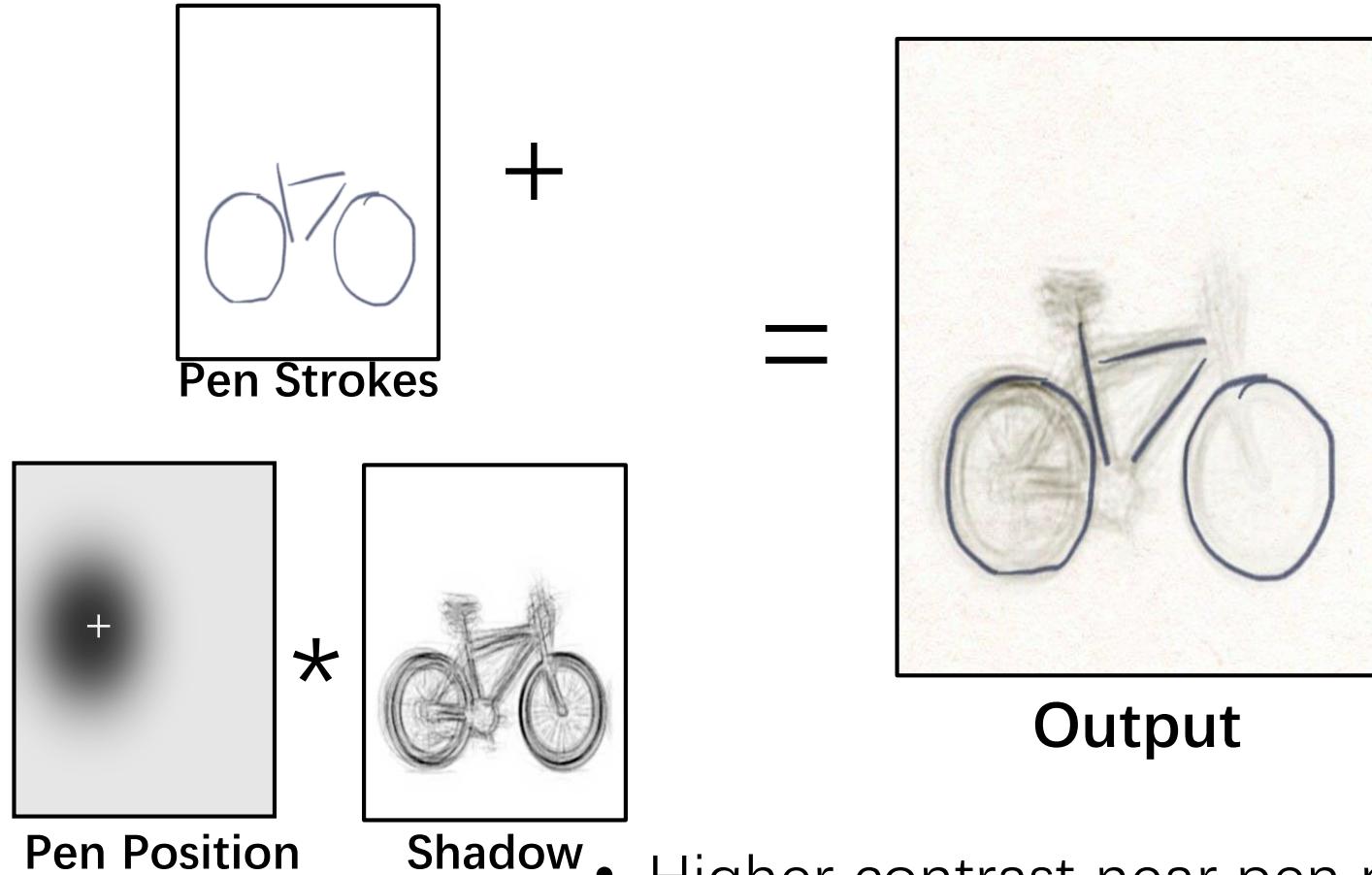


Pen Strokes



Shadows are the  
composite of many  
images

# Rendering



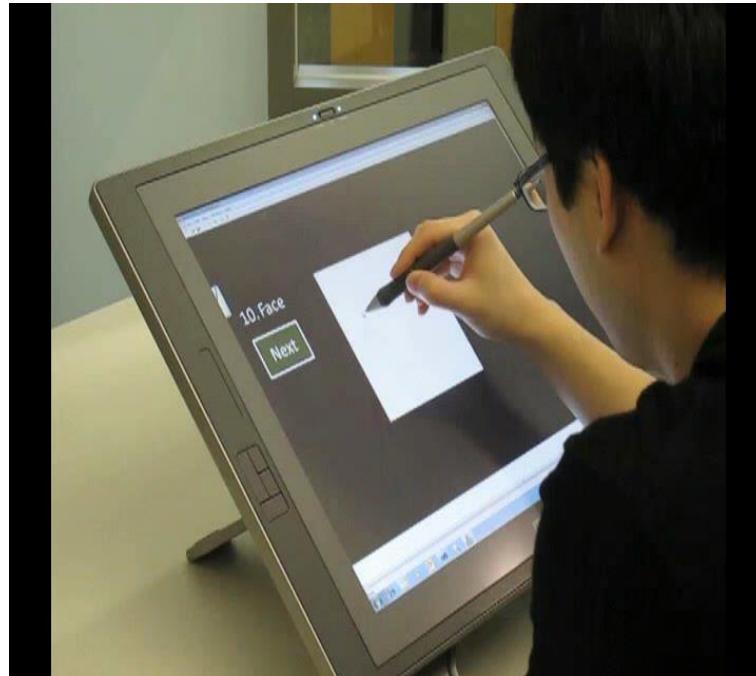
Pen Position

Shadow

- Higher contrast near pen position
- Main focus is user's drawing

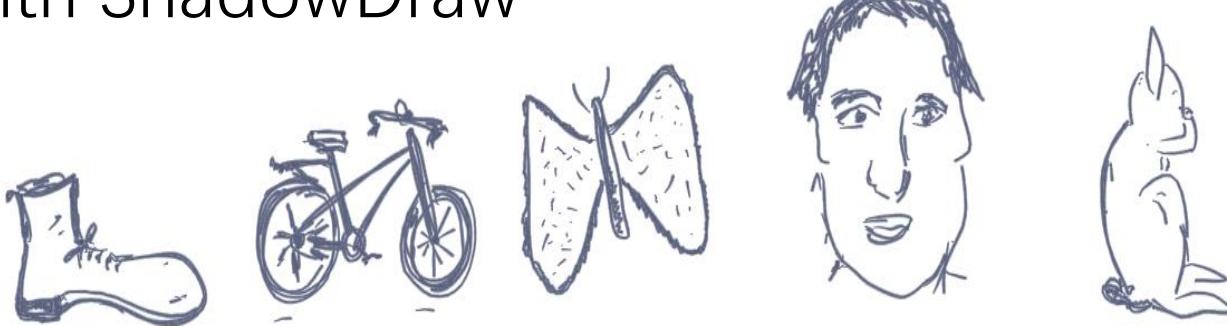
# User studies

- 30,000 images, 20 categories
- 16 drawers (8 men, 8 women), 8 evaluators
- 5 objects (shoe, face, bicycle, butterfly, rabbit)

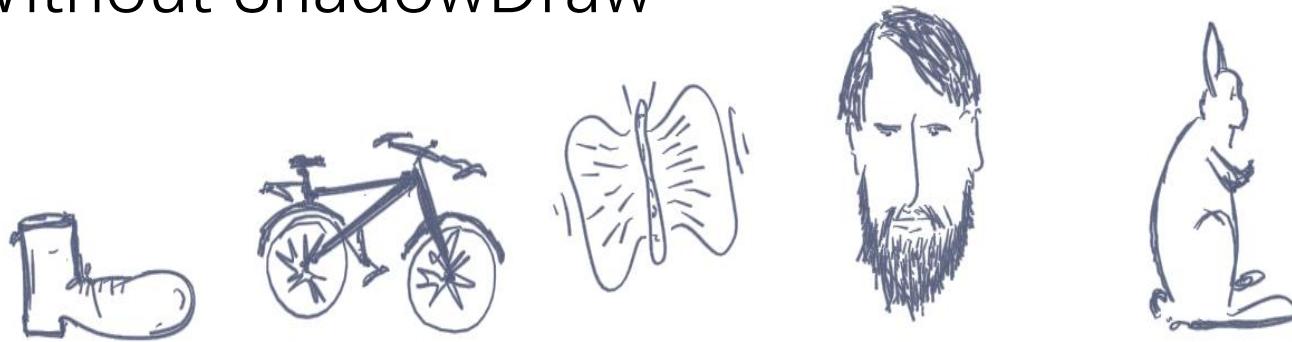


# User studies

With ShadowDraw



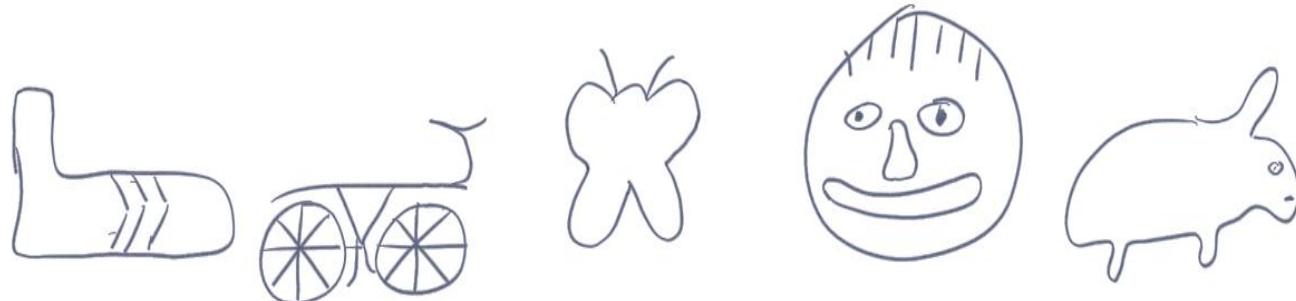
Without ShadowDraw



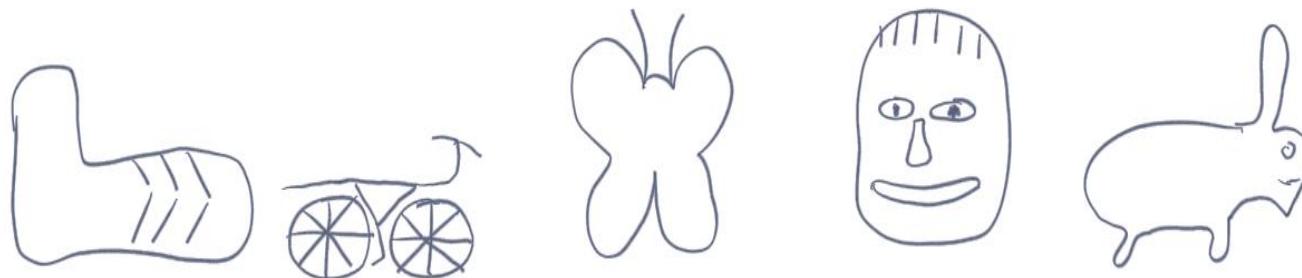
**Good drawers**

# User studies

With ShadowDraw



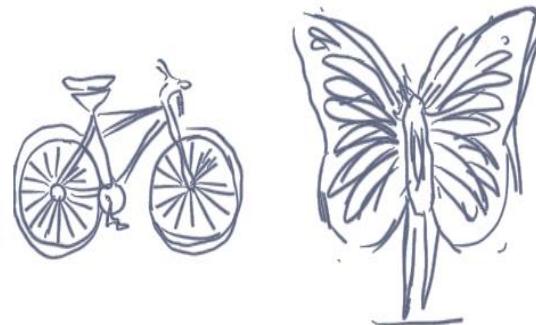
Without ShadowDraw



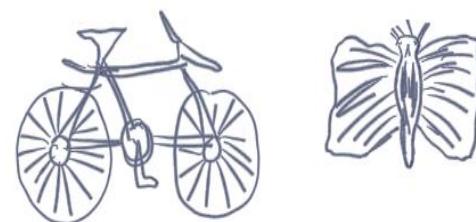
**Bad drawers**

# User studies

With ShadowDraw

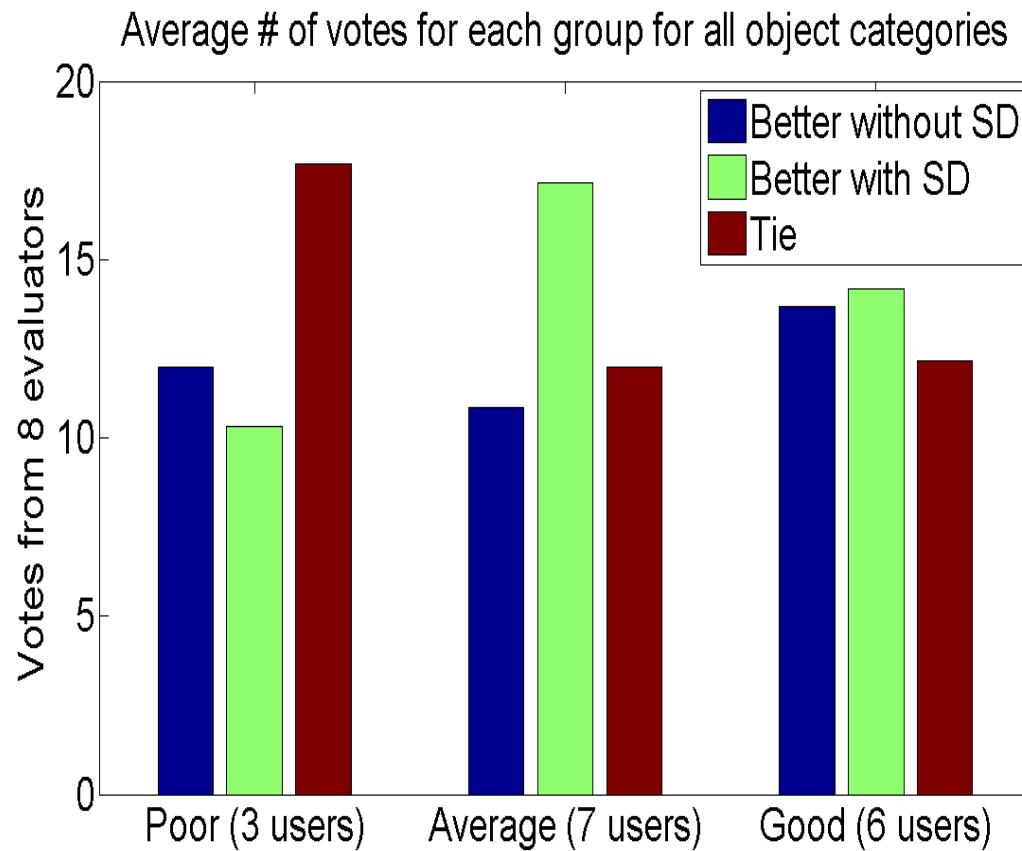


Without ShadowDraw



Average drawers

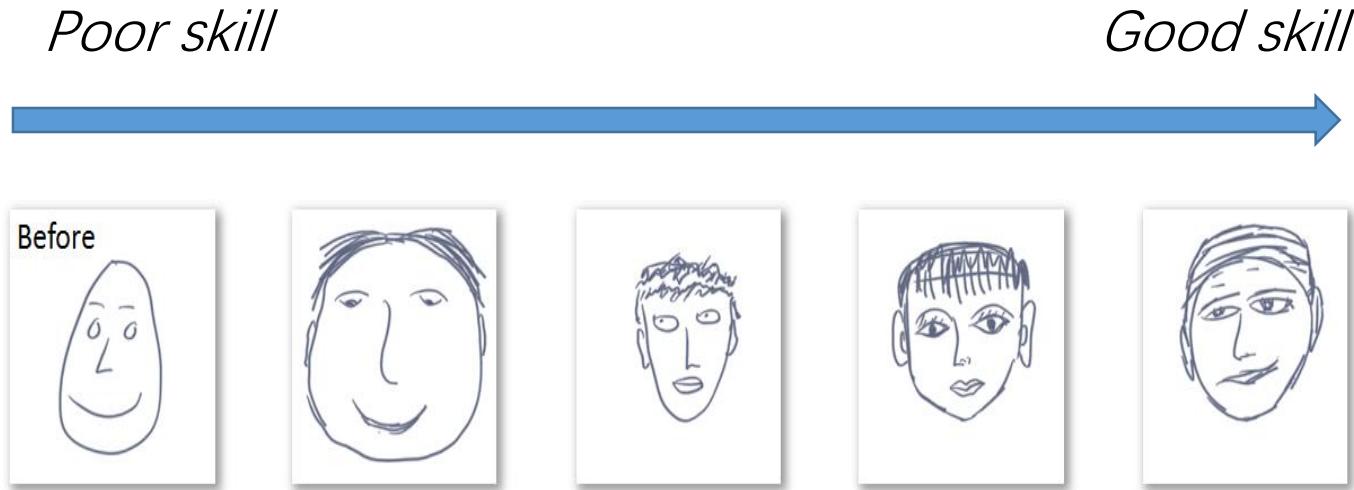
# User studies



- Significant improvement for “Average” group

# User studies

After training all users improved:



Subject's personal style is maintained!

# 主要内容

- 基于互联网图像的颜色空间降维技术
  - Data-Driven Color Manifolds. @SIGGRAPH 2015. Chuong H. Nguyen et. al.
- 基于互联网图像的辅助绘制技术
  - ShadowDraw: Real-Time User Guidance for Freehand Drawing @SIGGRAPH 2011. Yong Jae Lee et. al.
- 基于互联网图像的场景漫游
  - Photo tourism: Exploring photo collections in 3D. @SIGGRAPH 2006. Noah Snavely.et.al.
- 基于互联网数据的图像补全
  - Scene Completion Using Millions of Photographs. @SIGGRAPH 2007. J. Hays & A. A. Efros.
- 基于互联网的图像检索与融合
  - Sketch2Photo: Internet Image Montage. @SIGGRAPH ASIA 2009. Tao Chen et. al.



15,464



37,383



76,389

Flickr: Creative Commons - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.flickr.comcreativecommons/by-nc-nd-2.0/

Home Tags Groups People Invite

Logged in as Jimantha | Your Account | Help | Sign Out

Photos: Yours Upload Organize Your Contacts Explore

flickr<sup>BETA</sup>

## Creative Commons / Attribution-NonCommercial-NoDerivs License

(Or, [browse popular tags](#))

SEARCH

SOME RIGHTS RESERVED

Here are the **100 most recent** licensed photos:

 From <a href="#">mclarnell</a>	 From <a href="#">mugsy1274</a>	 From <a href="#">darren 'djp' paine</a>	 From <a href="#">darren 'djp' paine</a>	 From <a href="#">dizz</a>
 From <a href="#">alpha_zone</a>	 From <a href="#">darren 'djp' paine</a>	 From <a href="#">331</a>	 From <a href="#">mugsy1274</a>	 From <a href="#">dalekq</a>
 From <a href="#">mugsy1274</a>	 From <a href="#">darren 'djp' paine</a>	 From <a href="#">mugsy1274</a>	 From <a href="#">darren 'djp' paine</a>	 From <a href="#">mugsy1274</a>

Done Adblock

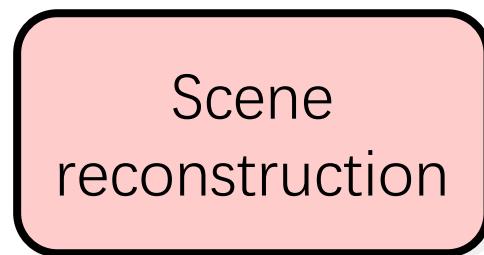
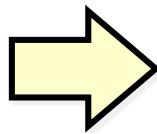
# Photo Tourism



# Photo Tourism overview



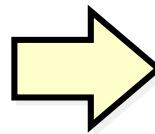
Input photographs



Relative camera positions  
and orientations

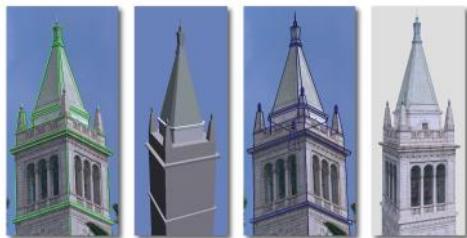
Point cloud

Sparse correspondence



# Related work

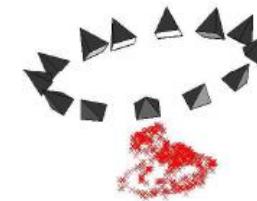
- Image-based modeling



Debevec, *et al.*  
SIGGRAPH 1996

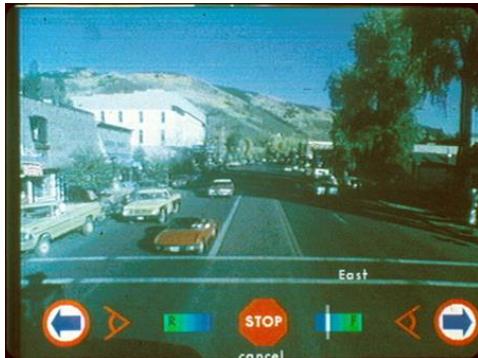


Schaffalitzky and Zisserman  
ECCV 2002



Brown and Lowe  
3DIM 2005

- Image-based rendering

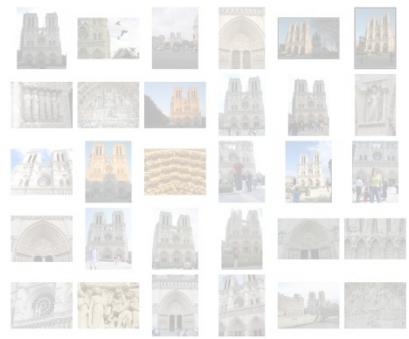


Aspen Movie Map  
Lippman, *et al.*, 1978

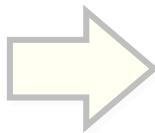
## Photorealistic IBR:

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

# Photo Tourism overview



Input photographs



Scene  
reconstruction

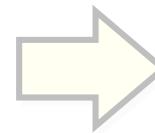
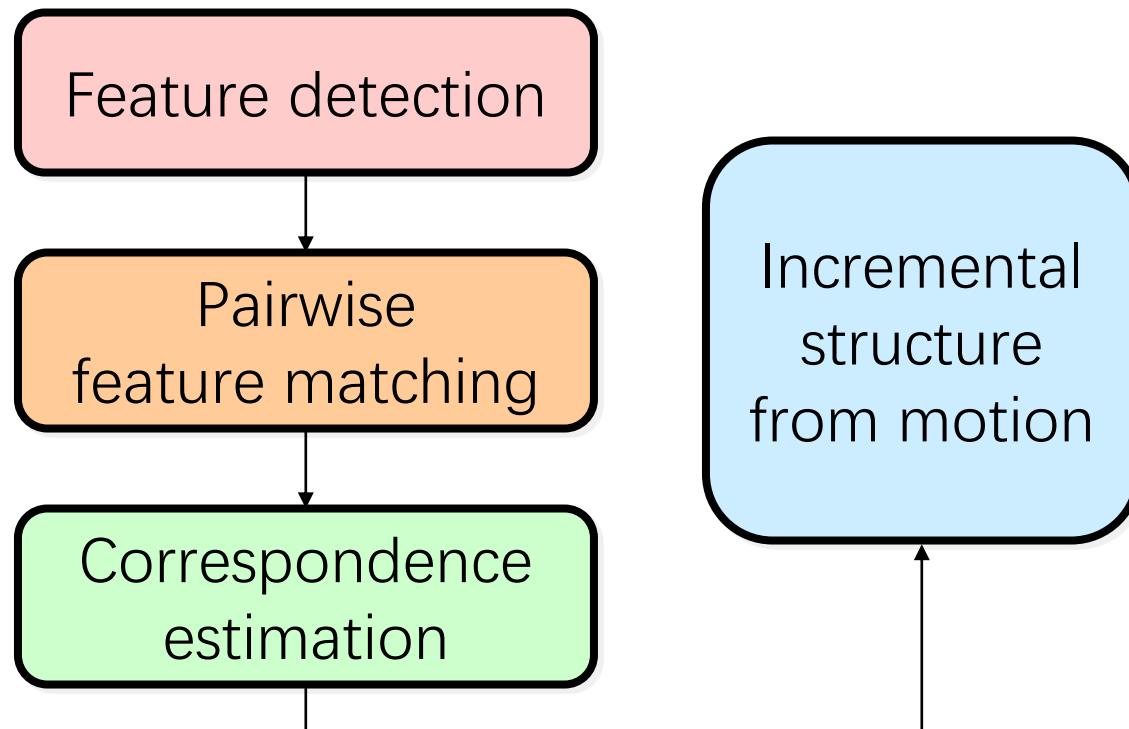


Photo  
Explorer

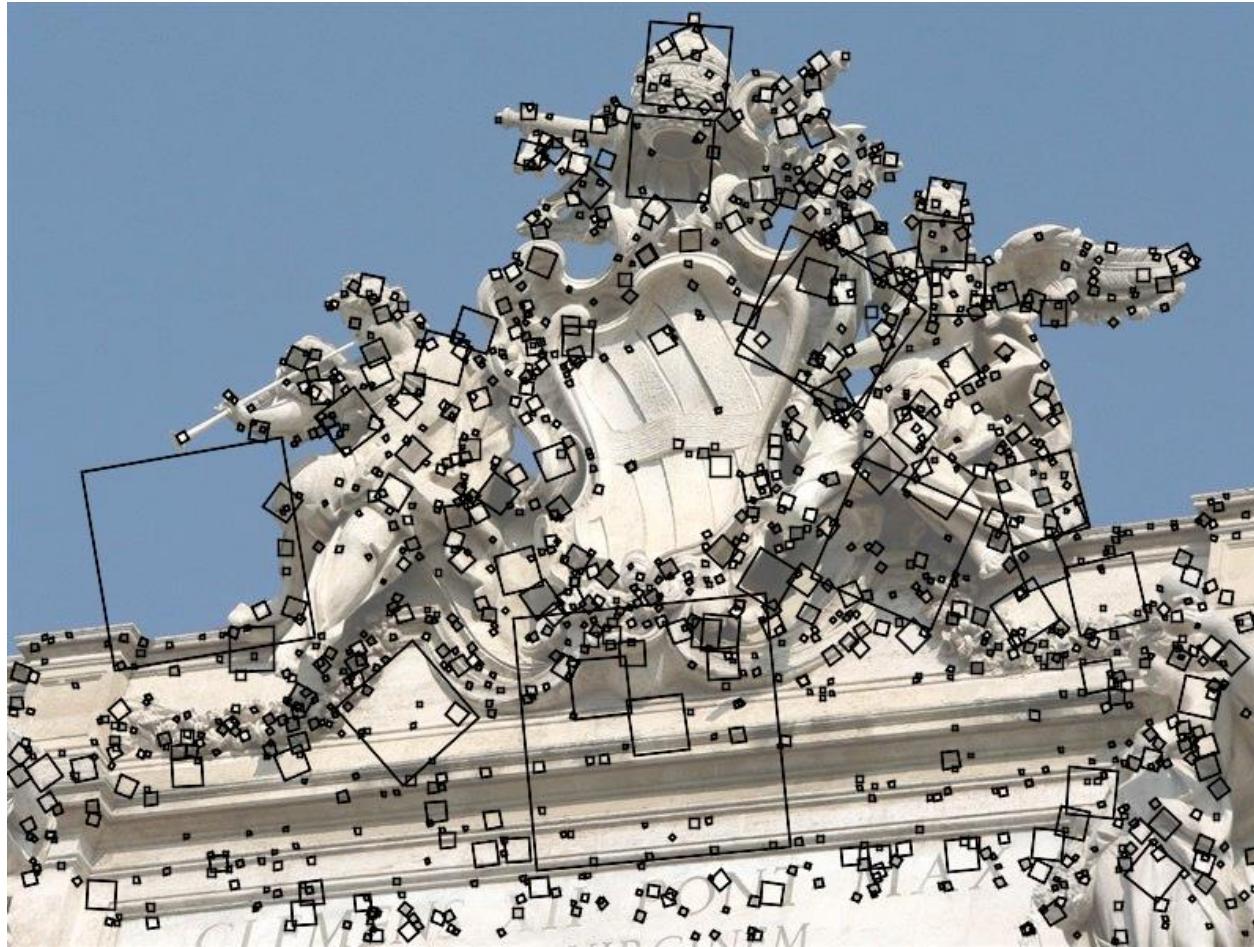
# Scene reconstruction

- Automatically estimate
  - position, orientation, and focal length of cameras
  - 3D positions of feature points



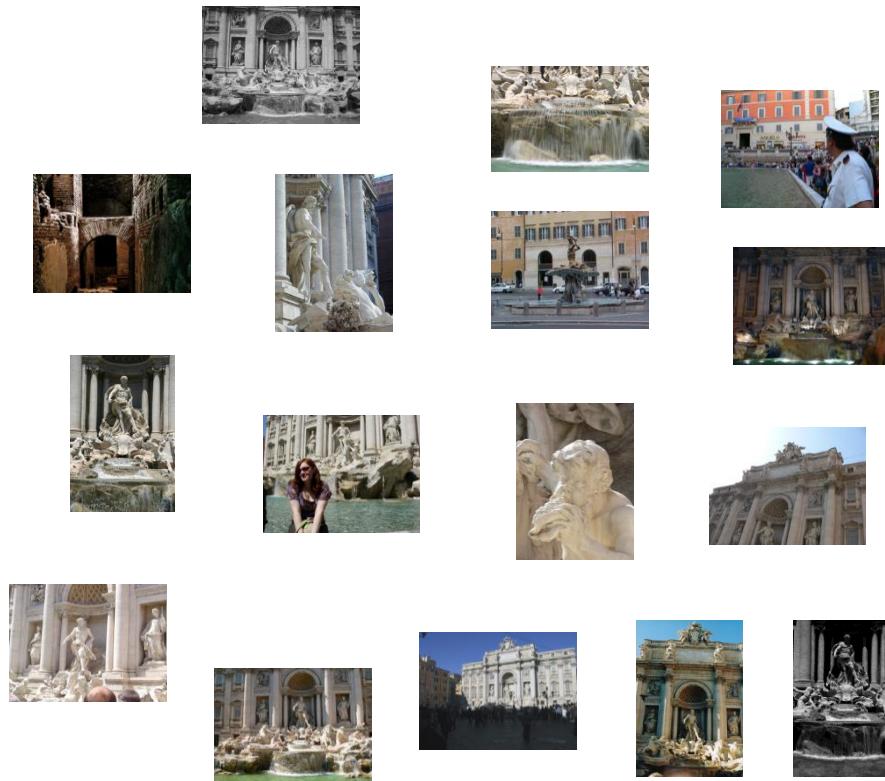
# Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



# Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



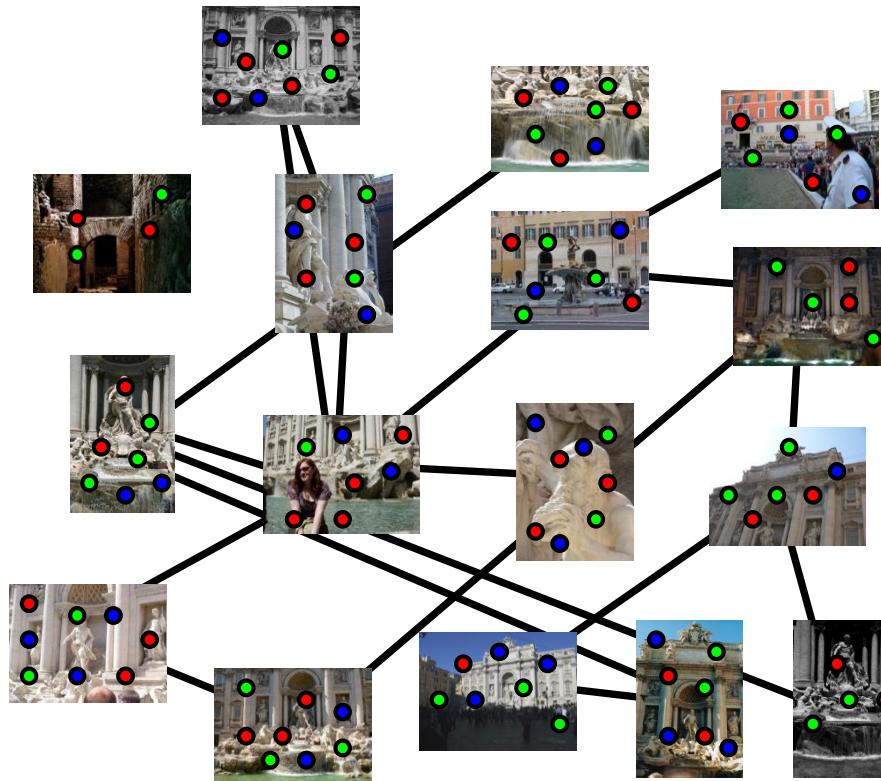
# Feature detection

Detect features using SIFT [Lowe, IJCV 2004]



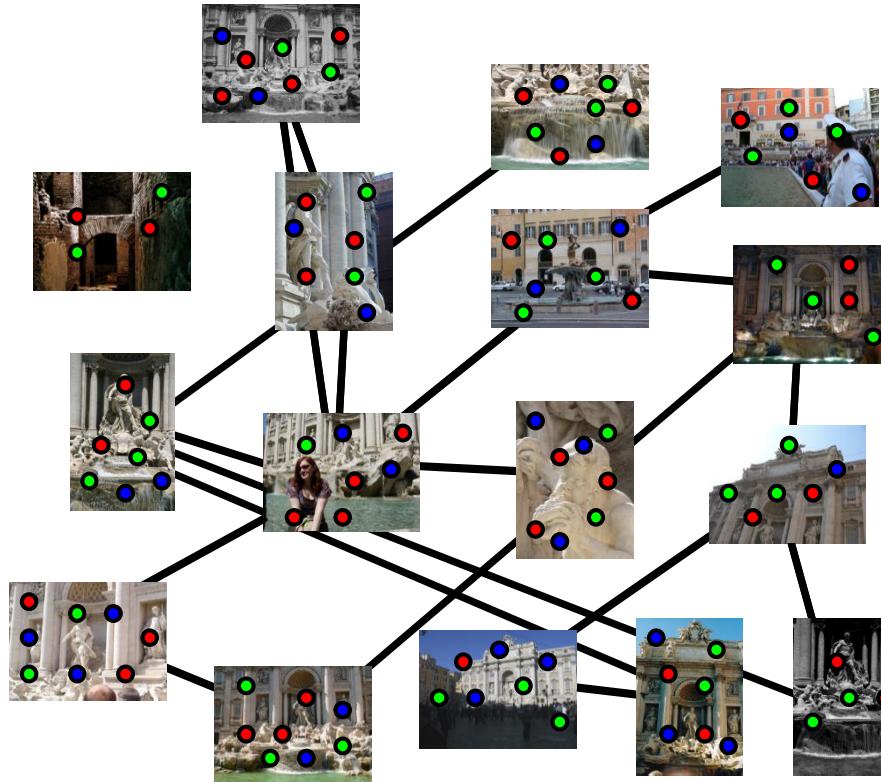
# Feature matching

Match features between each pair of images

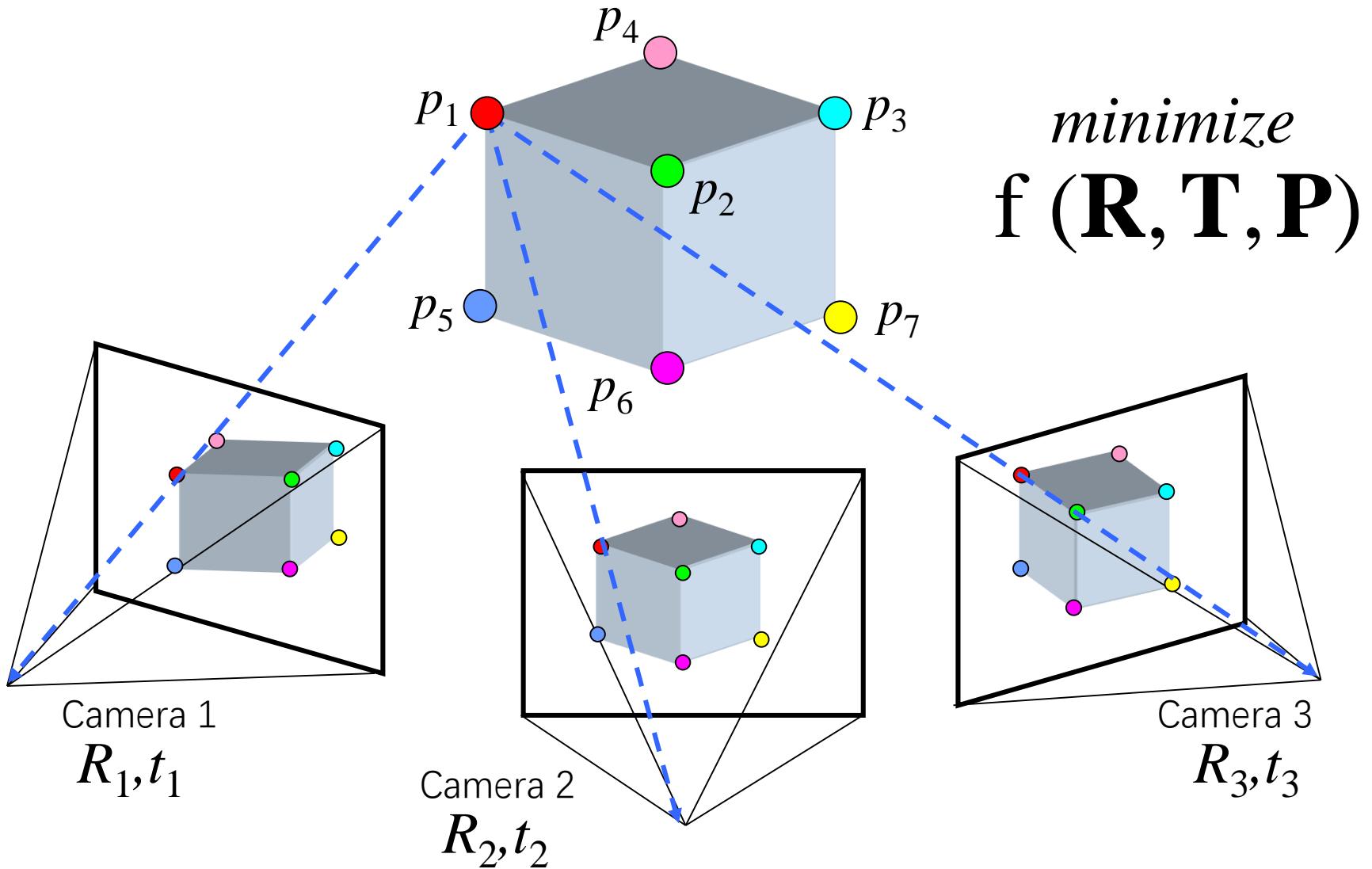


# Feature matching

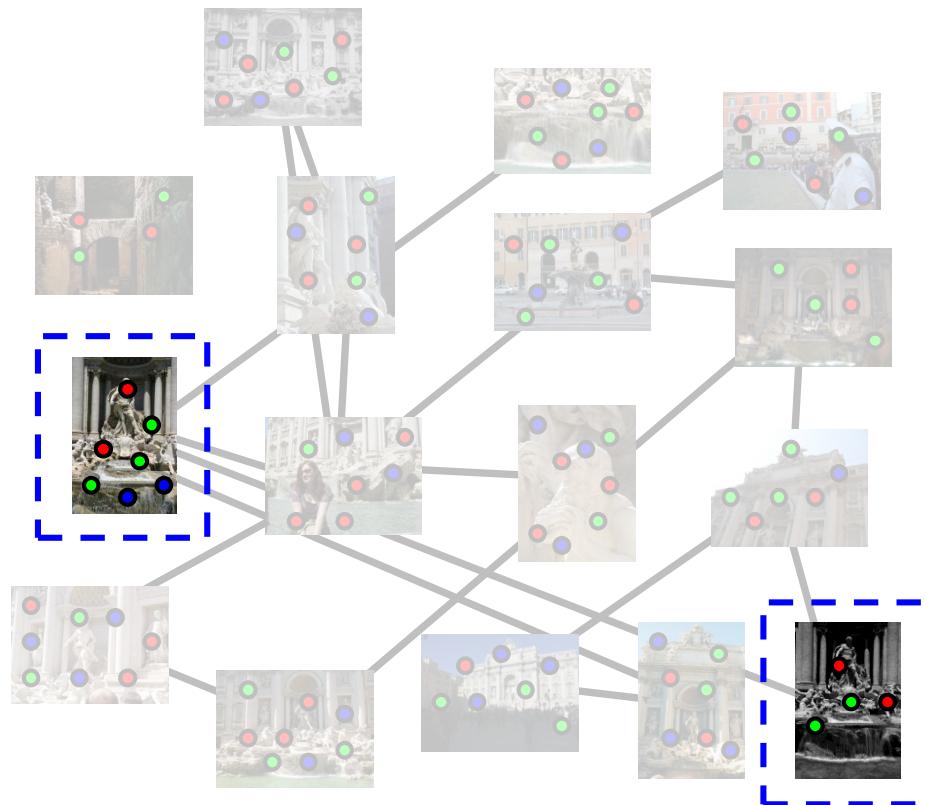
Refine matching using RANSAC [Fischler & Bolles 1987] to estimate fundamental matrices between pairs



# Structure from motion



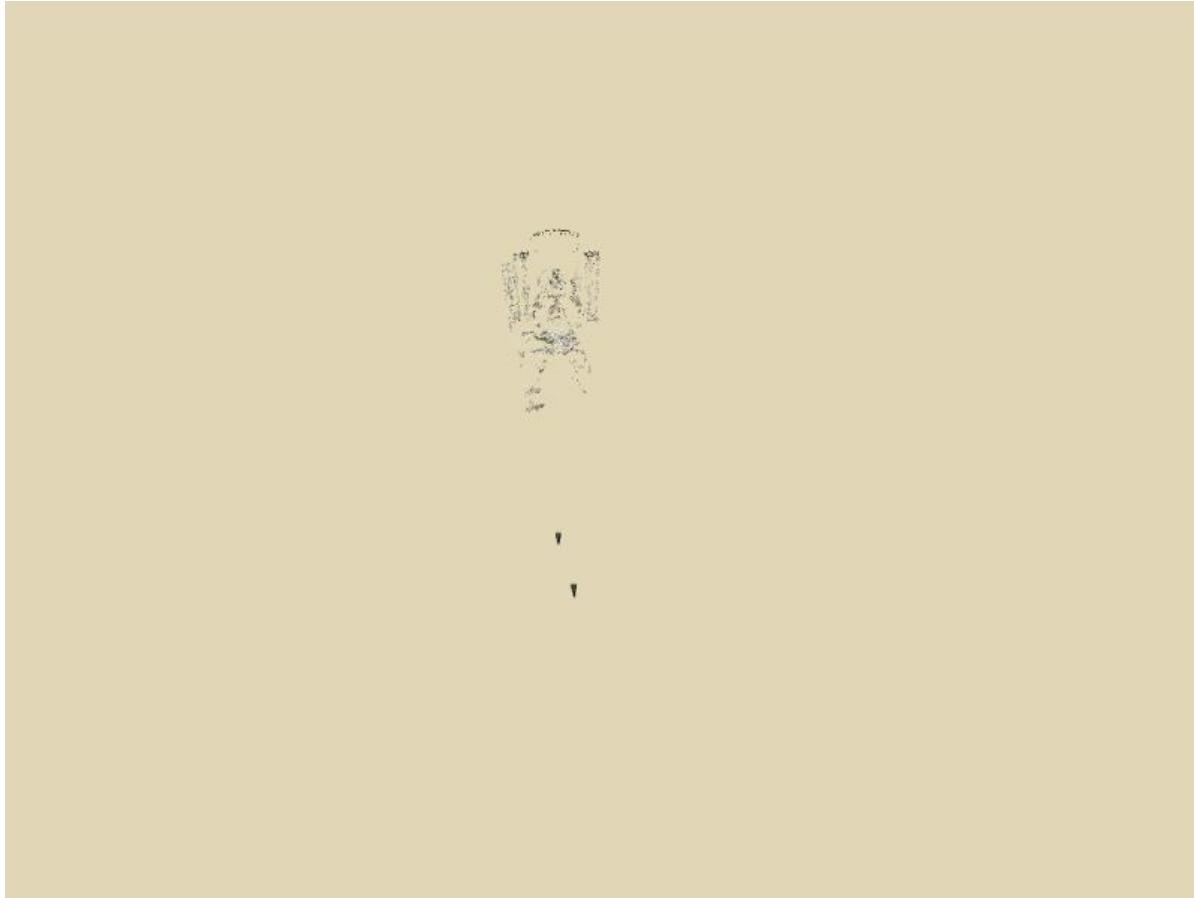
# Incremental structure from motion



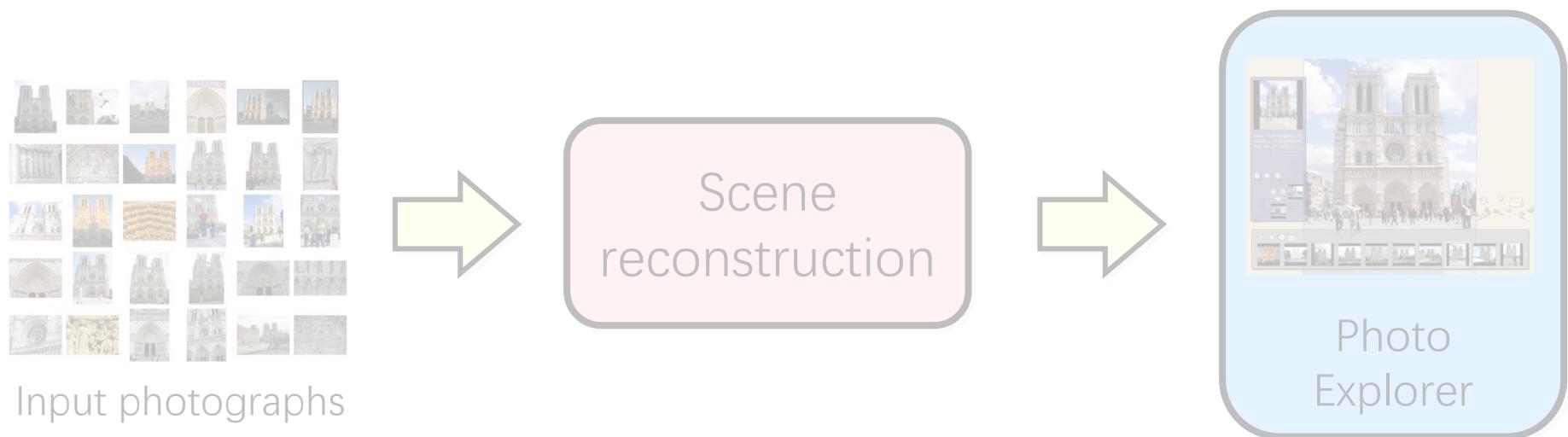
# Incremental structure from motion



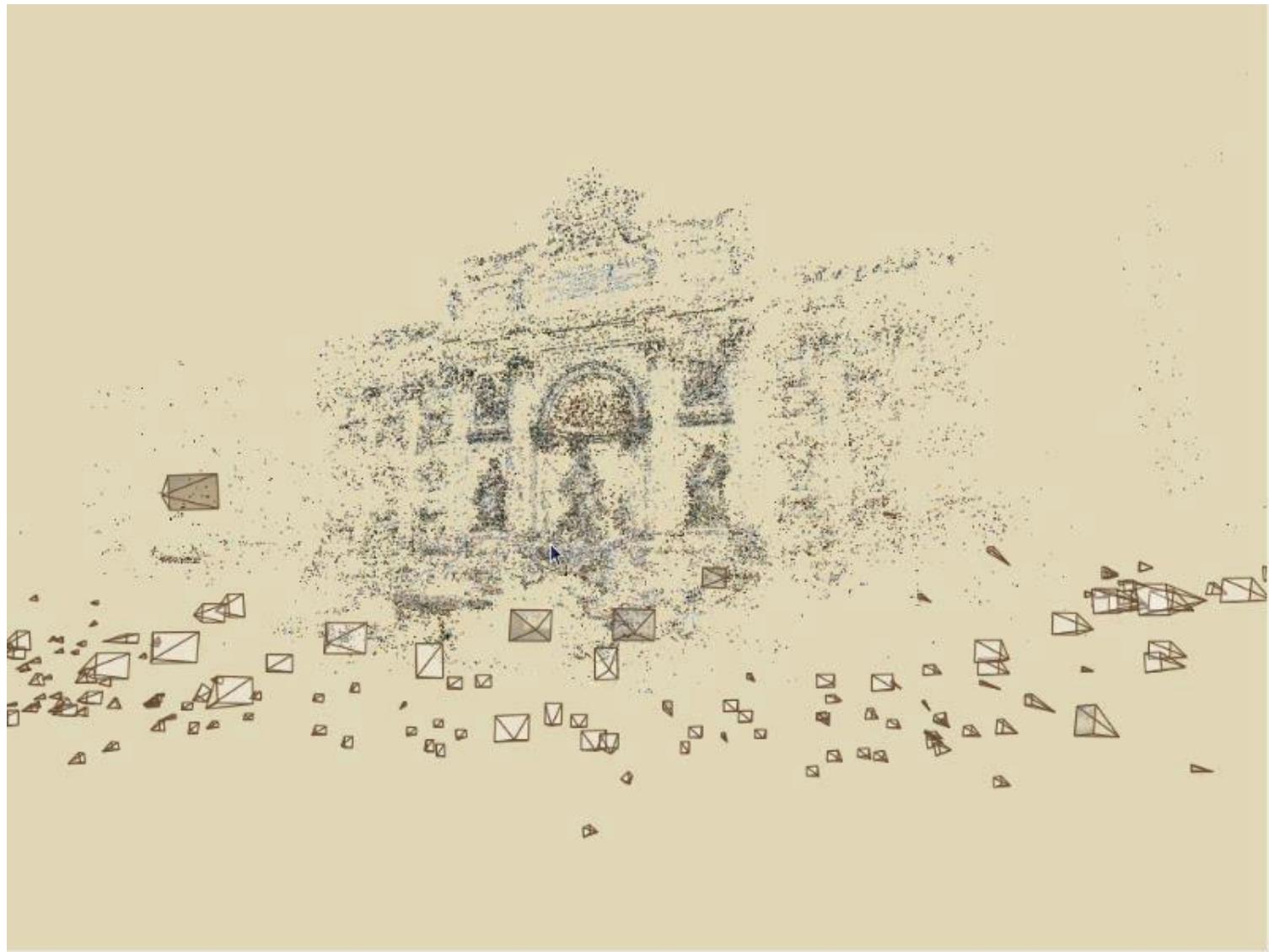
# Incremental structure from motion



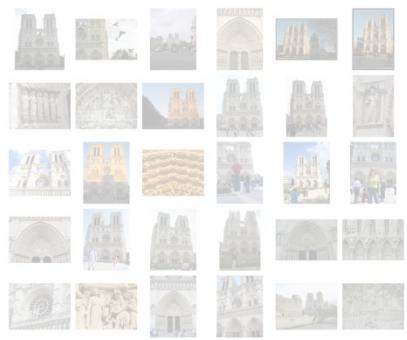
# Photo Tourism overview



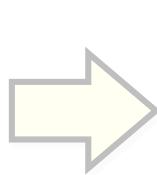
# Photo Explorer



# Photo Tourism overview



Input photographs



Scene  
reconstruction

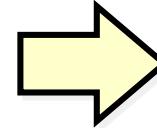
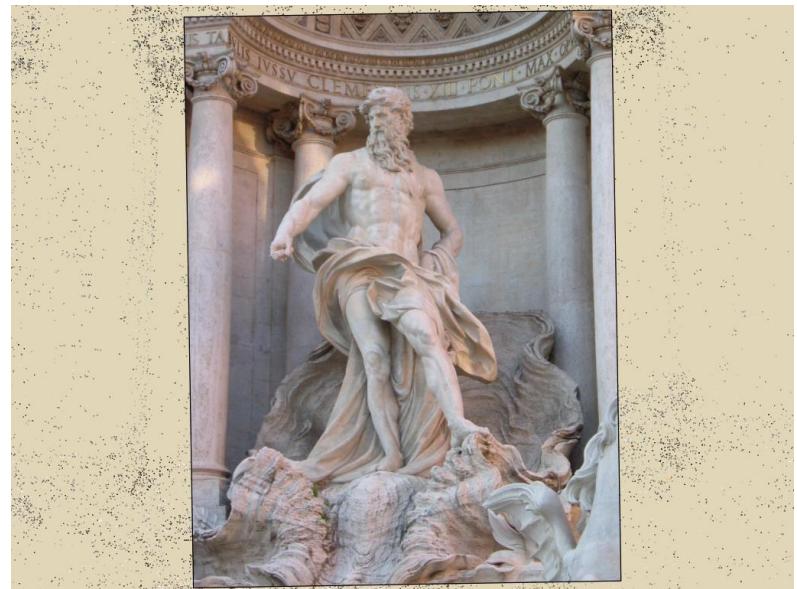
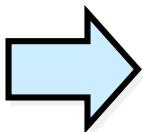


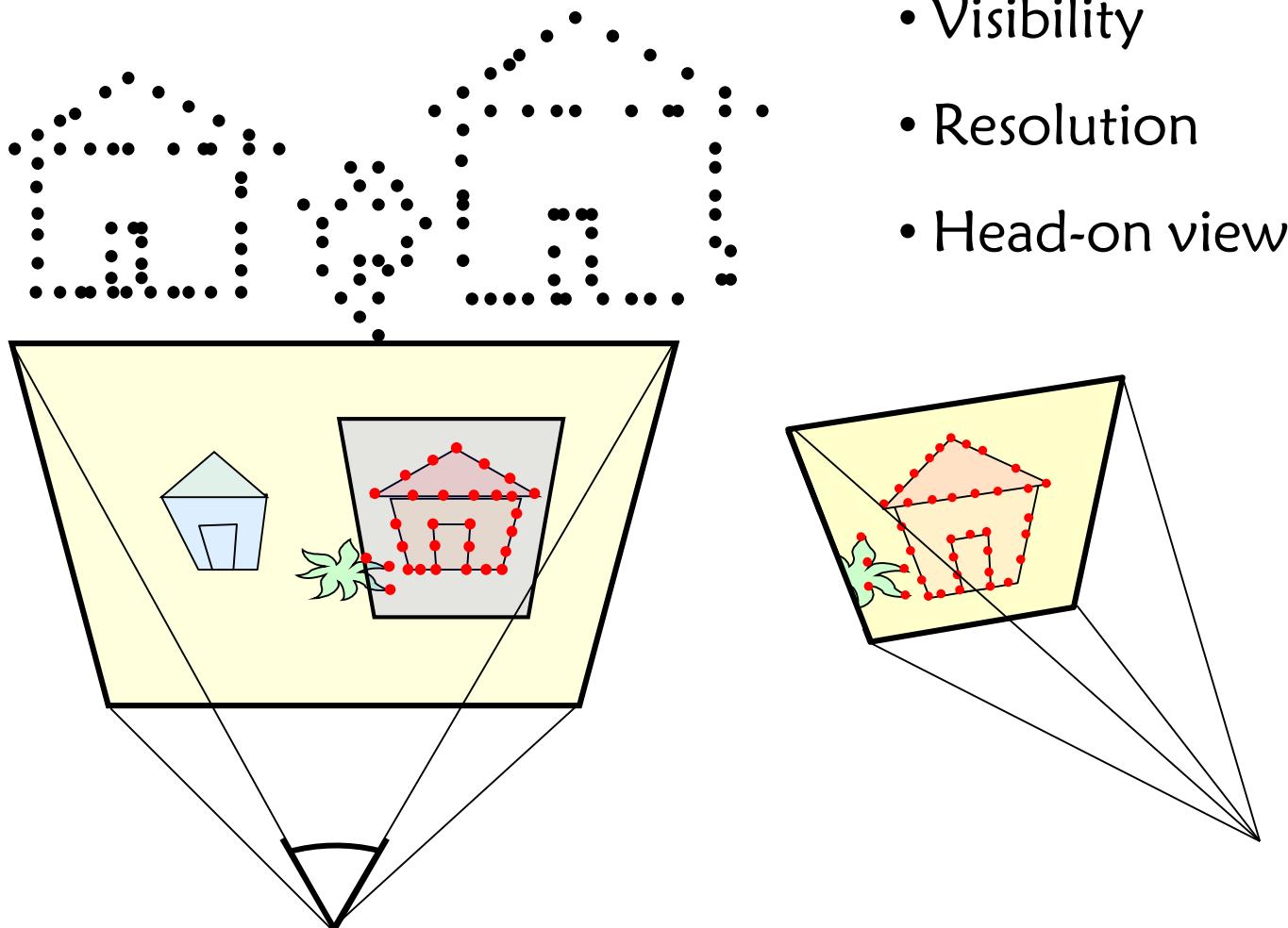
Photo  
Explorer

- Navigation
- Rendering
- Annotations

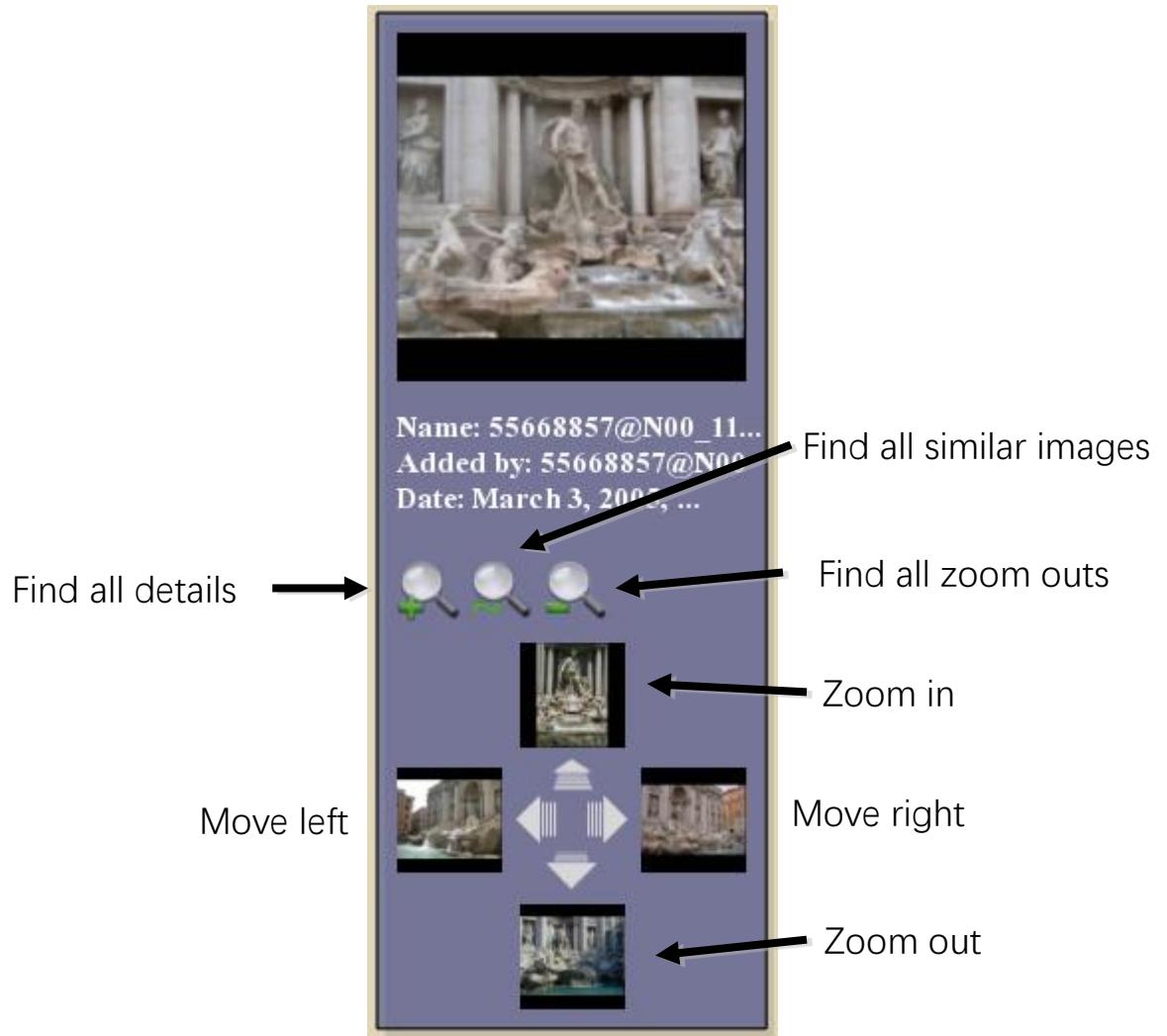
# Object-based browsing



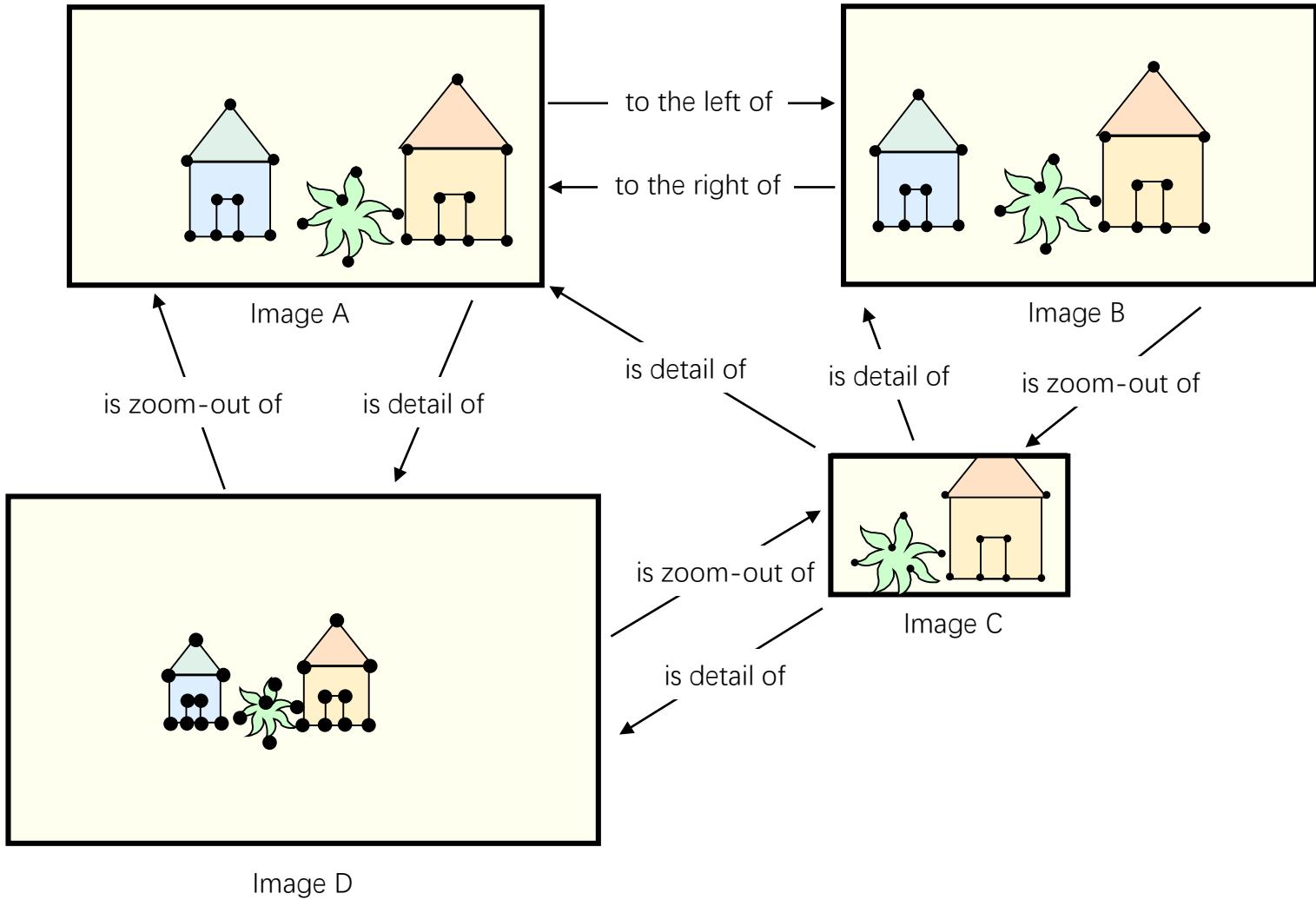
# Object-based browsing



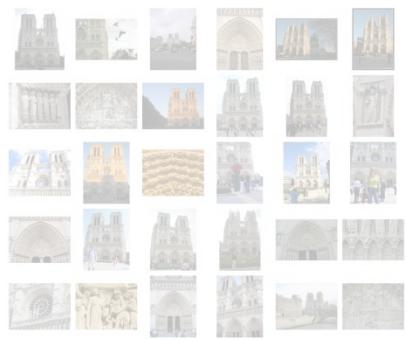
# Relation-based browsing



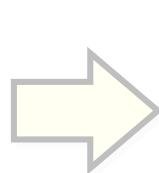
# Relation-based browsing



# Photo Tourism overview



Input photographs



Scene  
reconstruction

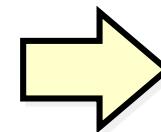
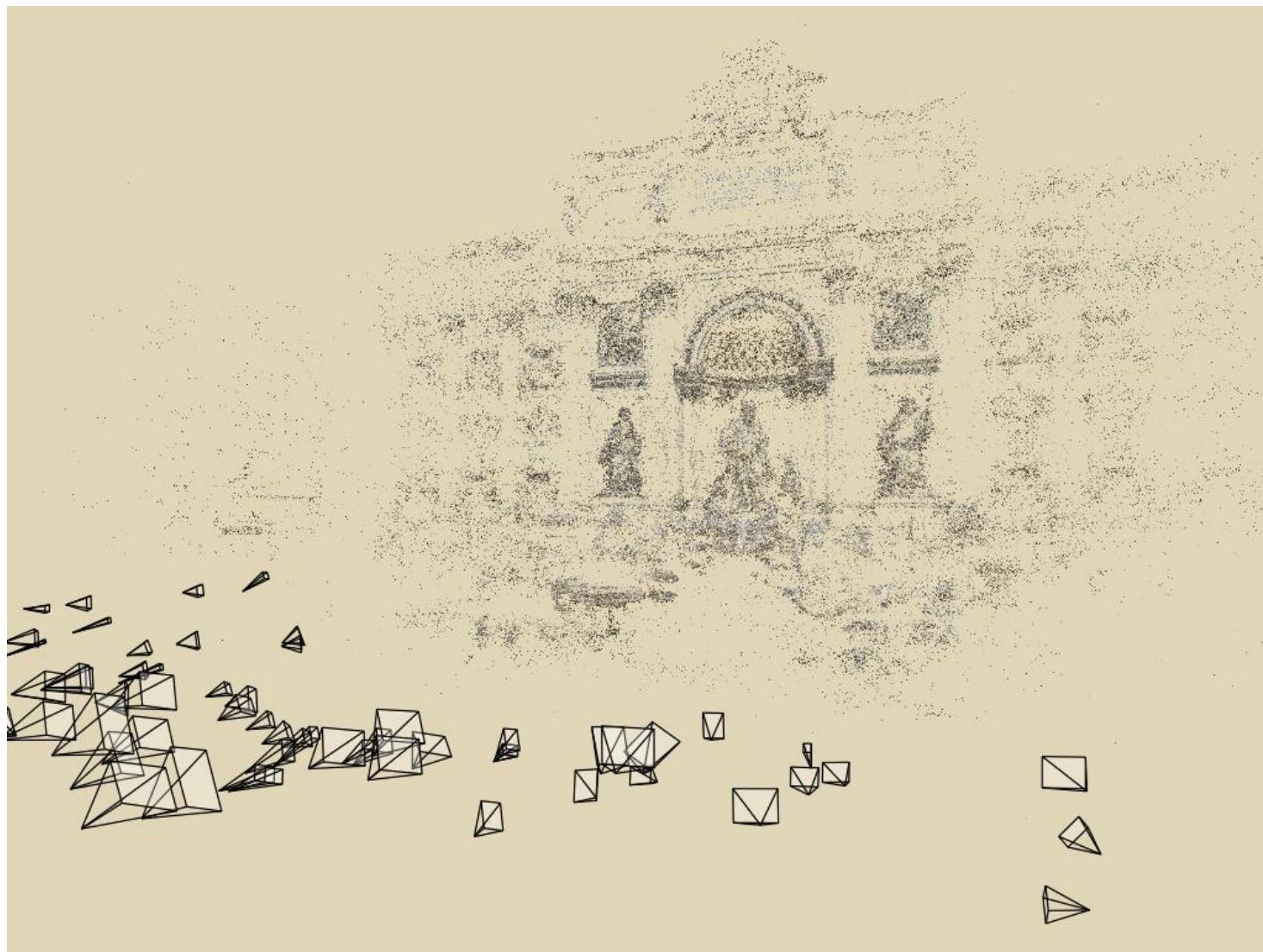


Photo  
Explorer

- Navigation
- Rendering
- Annotations

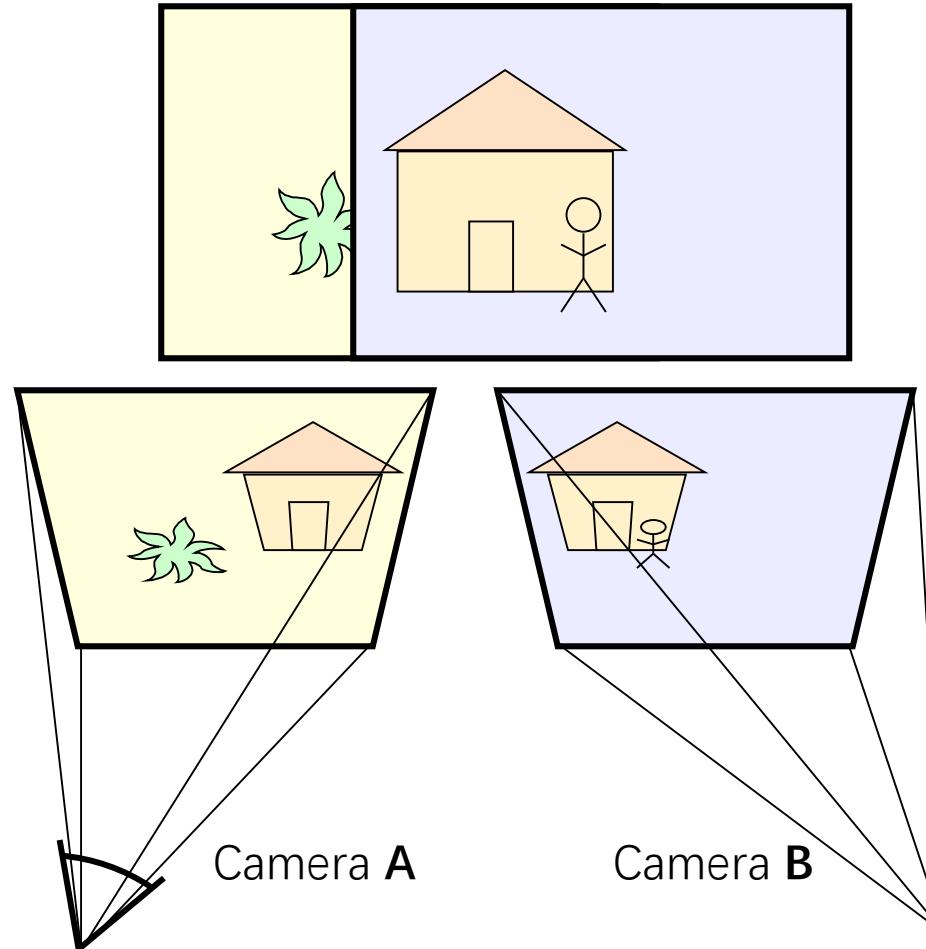
# Rendering



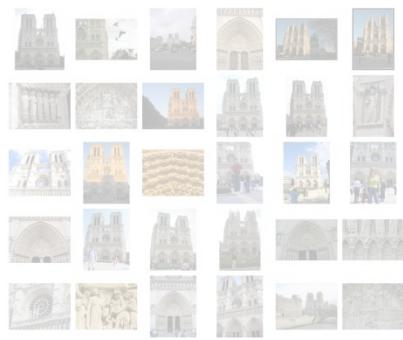
# Rendering



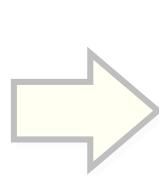
# Rendering transitions



# Photo Tourism overview



Input photographs



Scene  
reconstruction

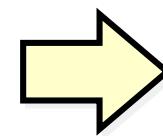


Photo  
Explorer

- Navigation
- Rendering
- Annotations

# Annotations



# Annotations

The screenshot shows a Flickr photo page for a photograph of Notre Dame de Paris. The page includes a toolbar at the top, the Flickr logo, navigation links (Home, Learn More, Sign Up, Explore), a search bar, and a user status message. The main content features the title "Notre Dame de Paris", a large image of the cathedral's facade, and a sidebar for the user "Jimantha". The sidebar displays a photostream with one photo, tags (notredame, paris), and additional information like copyright and date taken.

Notre Dame de Paris

Uploaded on January 13, 2006  
by Jimantha

- Jimantha's photostream

You are at the first photo You are at the last photo 1 photo

browse

Tags

notredame paris

Additional Information

All rights reserved

Taken on January 13, 2006

Viewed once

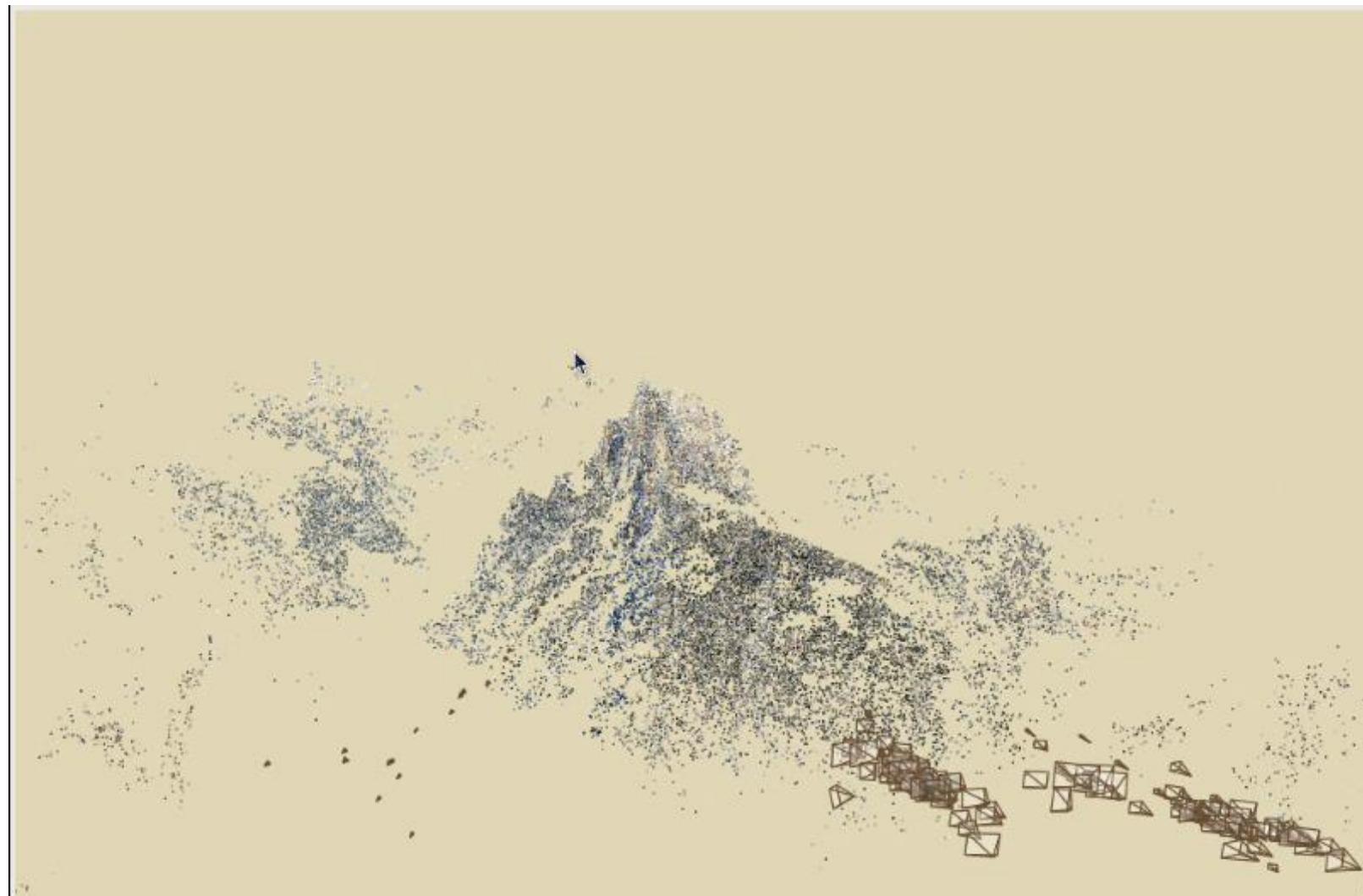
Would you like to comment?

Reproduced with permission of Yahoo! Inc. © 2005 by Yahoo! Inc.  
YAHOO! and the YAHOO! logo are trademarks of Yahoo! Inc.

# Annotations



# Yosemite



# 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

# 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

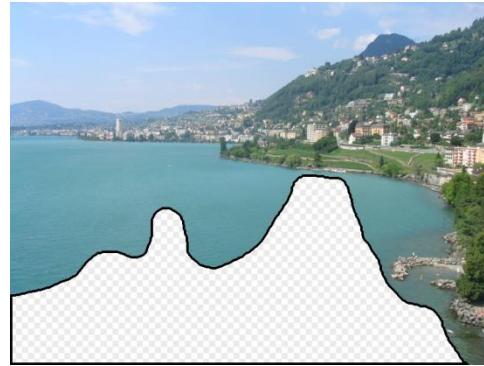


Mount Rushmore

# 主要内容

- 基于互联网图像的颜色空间降维技术
  - Data-Driven Color Manifolds. @SIGGRAPH 2015. Chuong H. Nguyen et. al.
- 基于互联网图像的辅助绘制技术
  - ShadowDraw: Real-Time User Guidance for Freehand Drawing @SIGGRAPH 2011. Yong Jae Lee et. al.
- 基于互联网图像的场景漫游
  - Photo tourism: Exploring photo collections in 3D. @SIGGRAPH 2006. Noah Snavely.et.al.
- 基于互联网数据的图像补全
  - Scene Completion Using Millions of Photographs. @SIGGRAPH 2007. J. Hays & A. A. Efros.
- 基于互联网的图像检索与融合
  - Sketch2Photo: Internet Image Montage. @SIGGRAPH ASIA 2009. Tao Chen et. al.

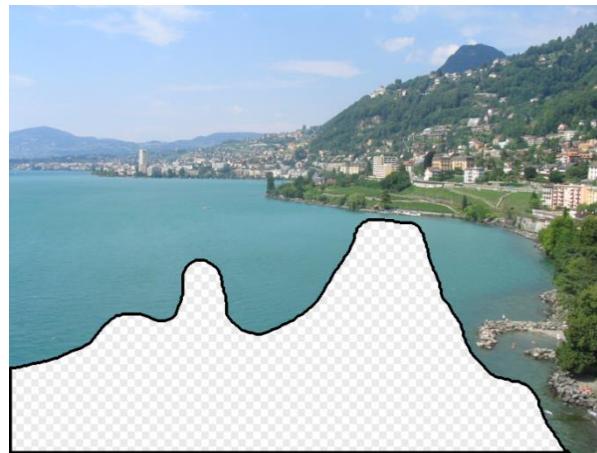
# Scene Completion Using Millions of Photographs



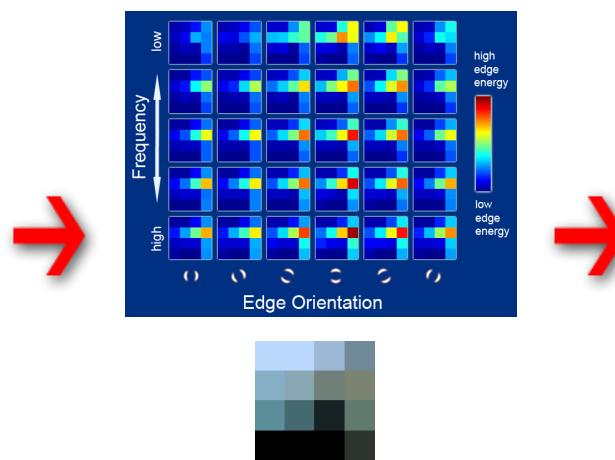
James Hays and Alexei A. Efros  
Carnegie Mellon University



# The Algorithm



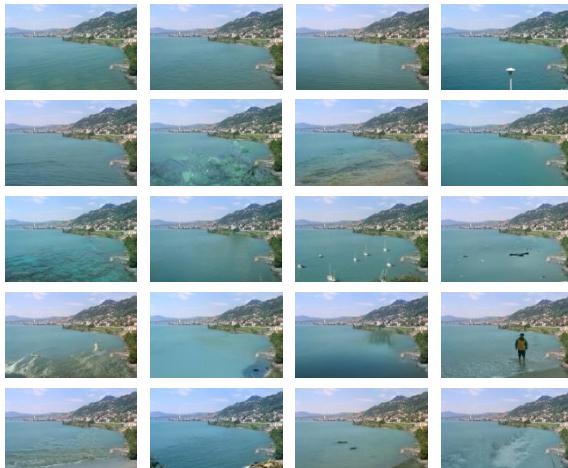
Input image



Scene Descriptor



Image Collection



20 completions



Context matching  
+ blending



200 matches

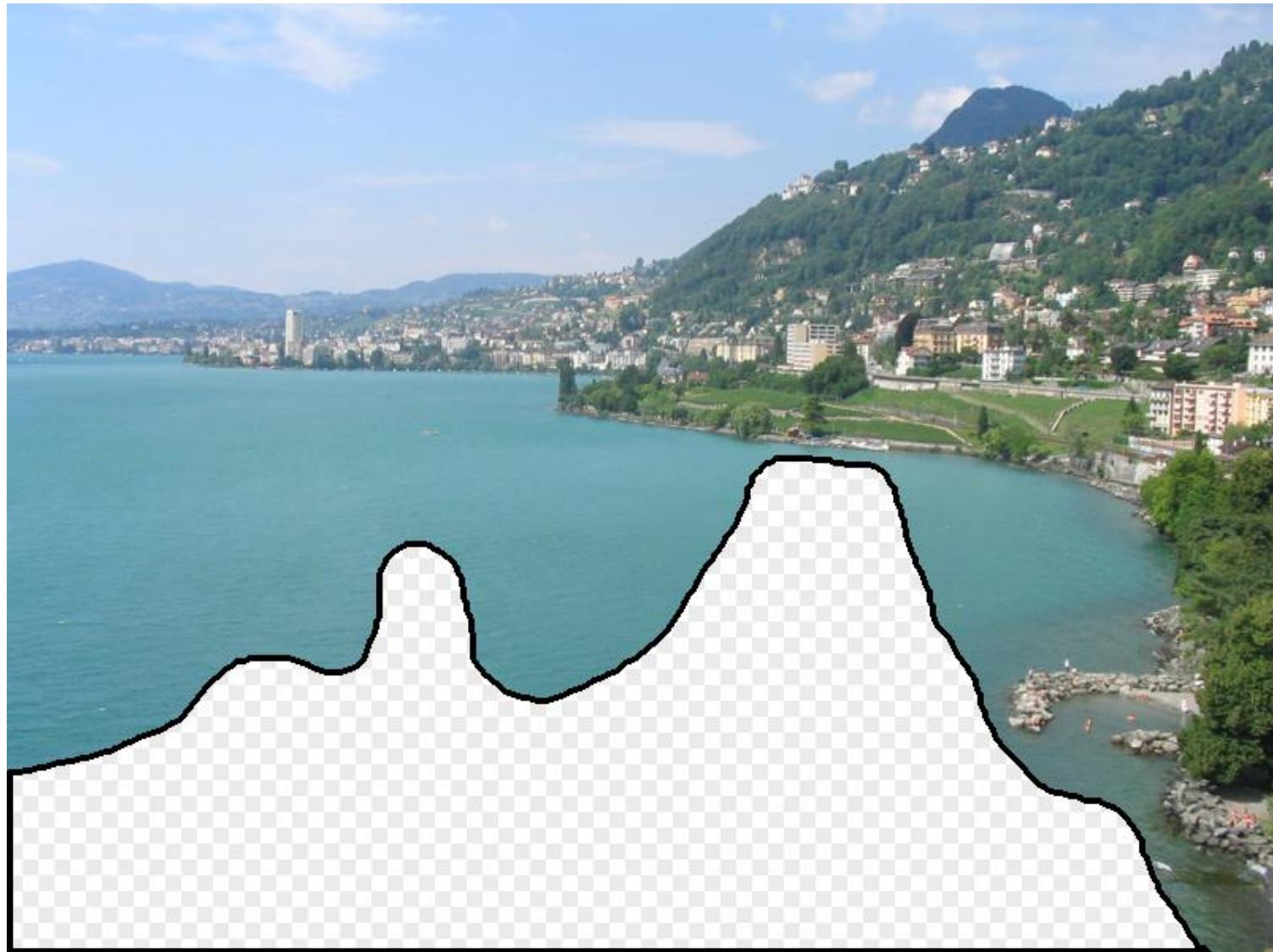
Hays and Efros, SIGGRAPH 2007

# Data

We downloaded 2.3 Million unique images from Flickr groups and keyword searches.

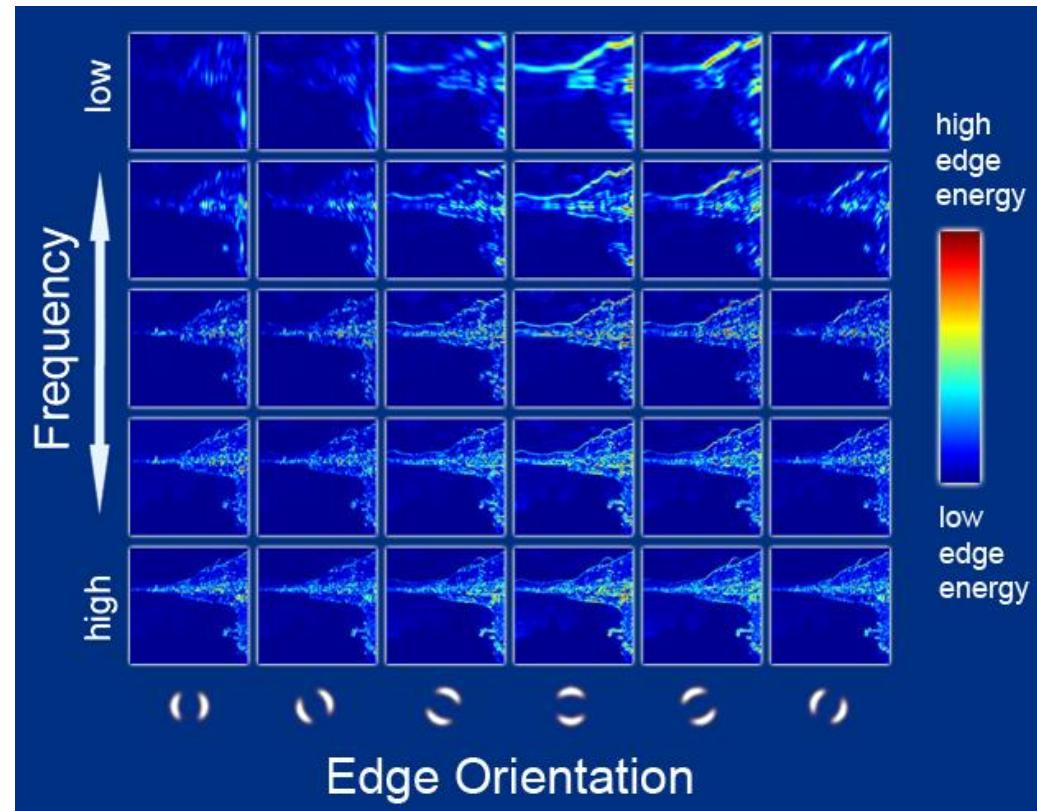
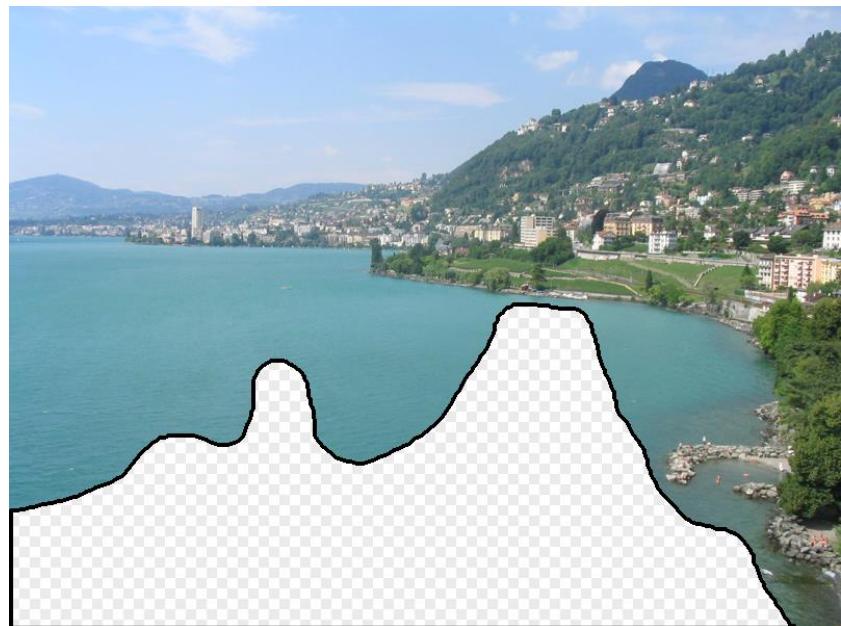


# Scene Matching

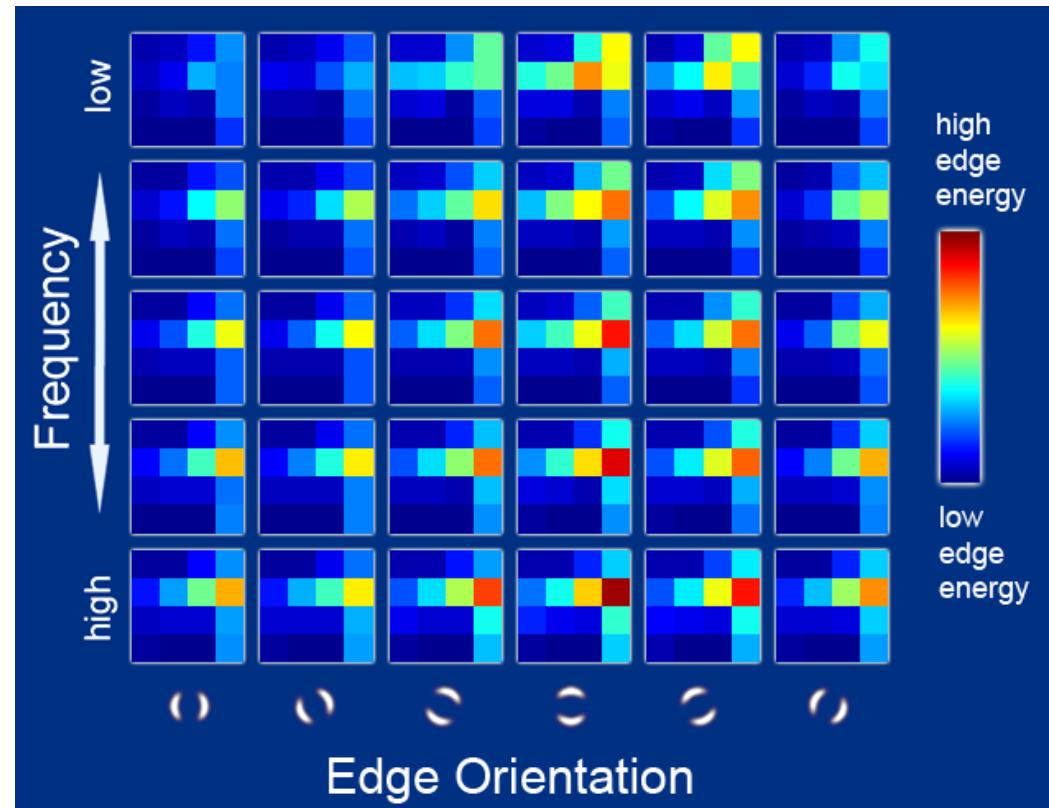
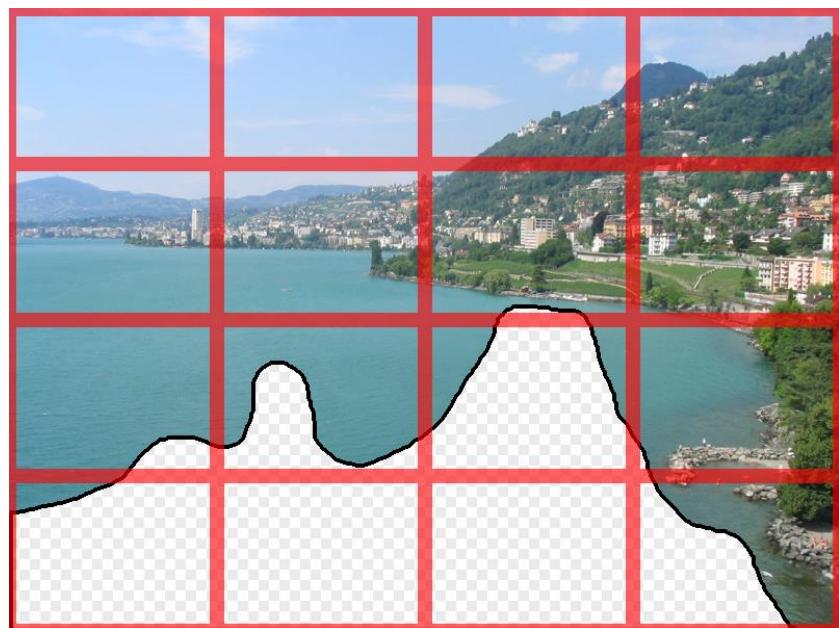


Hays and Efros, SIGGRAPH 2007

# Scene Descriptor

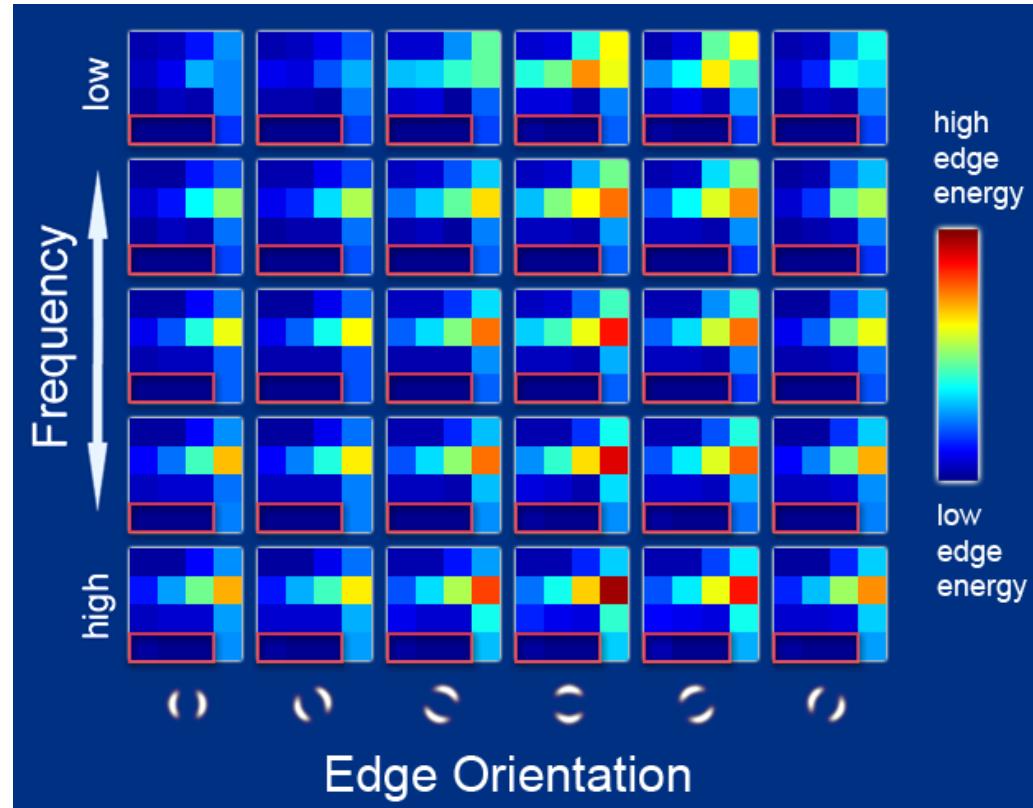
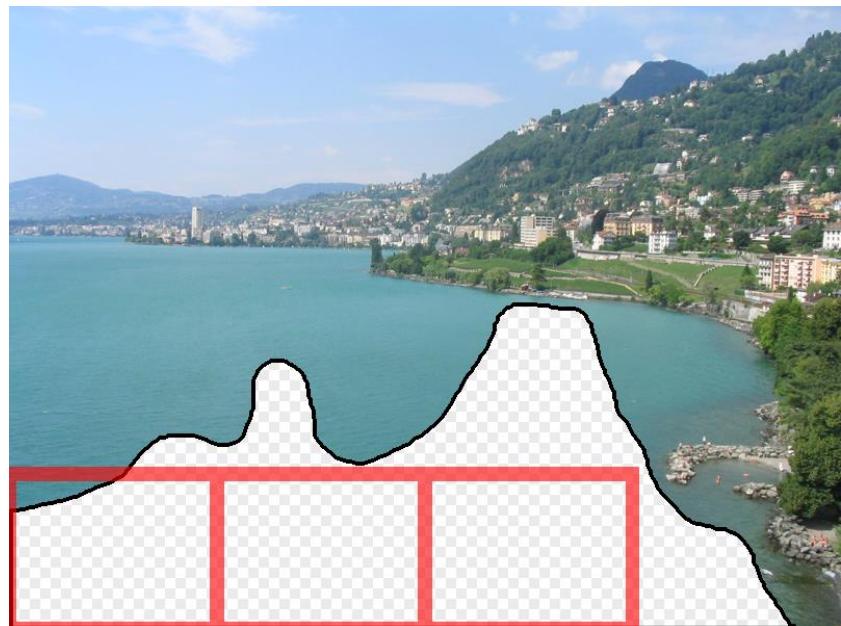


# Scene Descriptor



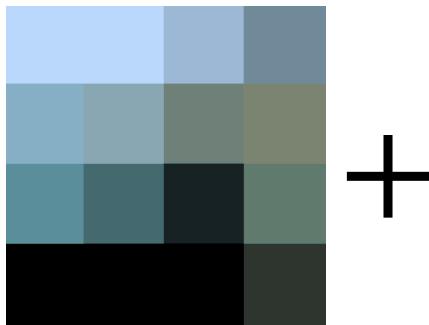
Gist scene descriptor  
(Oliva and Torralba 2001)

# Scene Descriptor

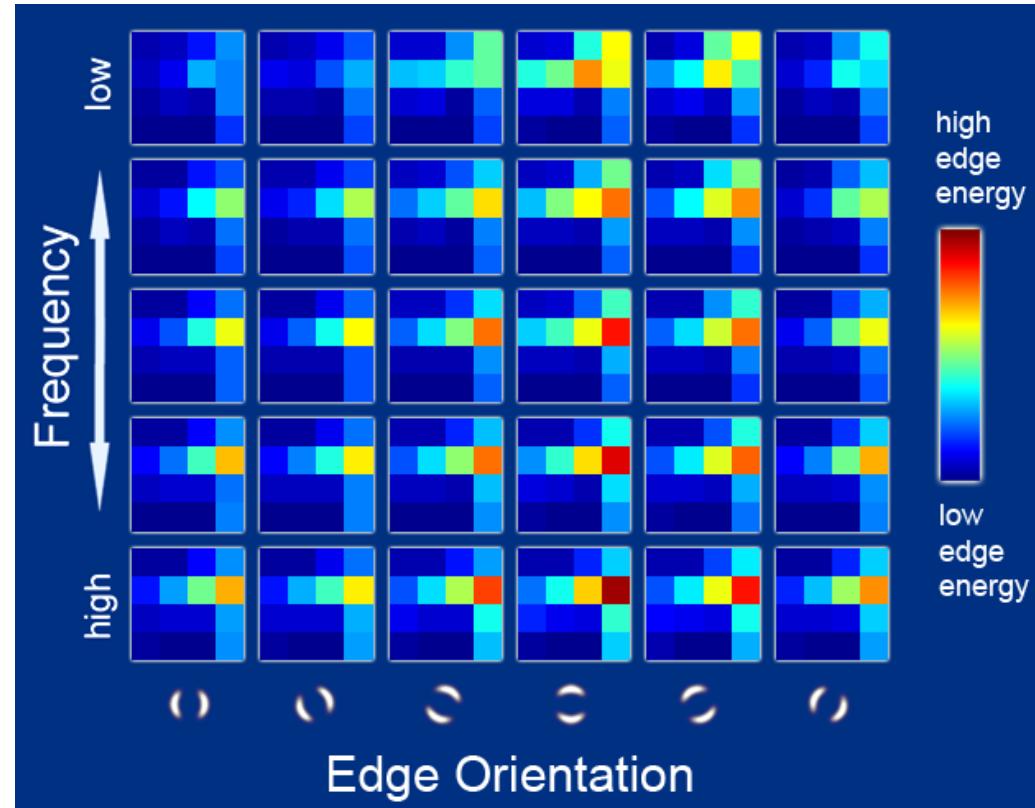


Gist scene descriptor  
(Oliva and Torralba 2001)

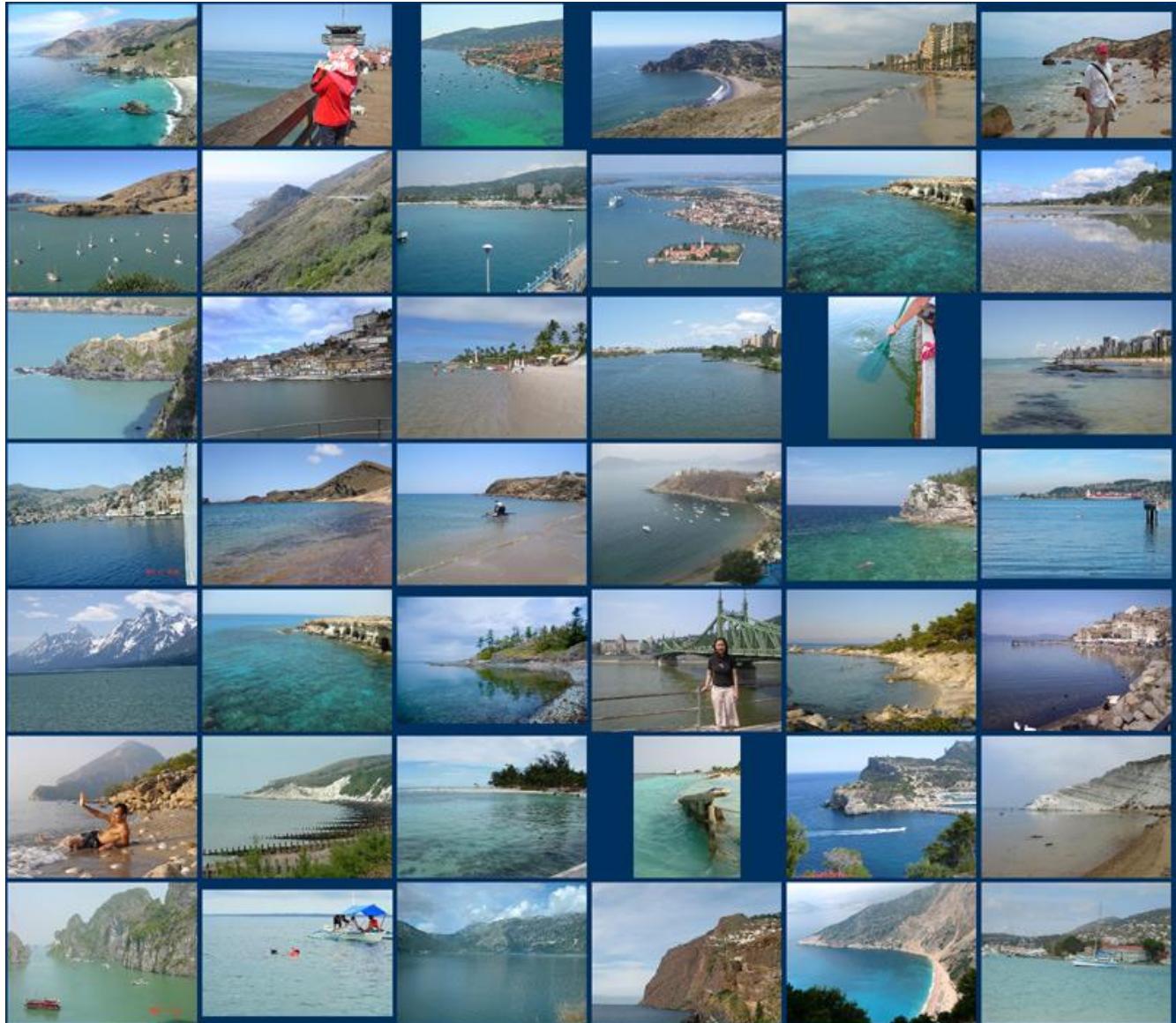
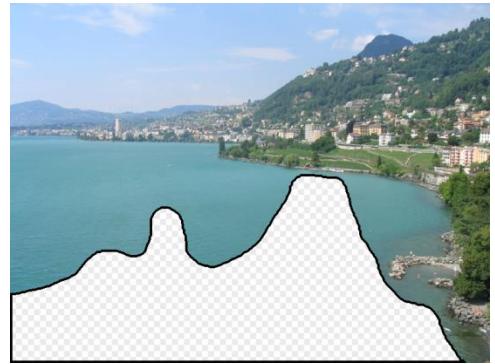
# Scene Descriptor



+

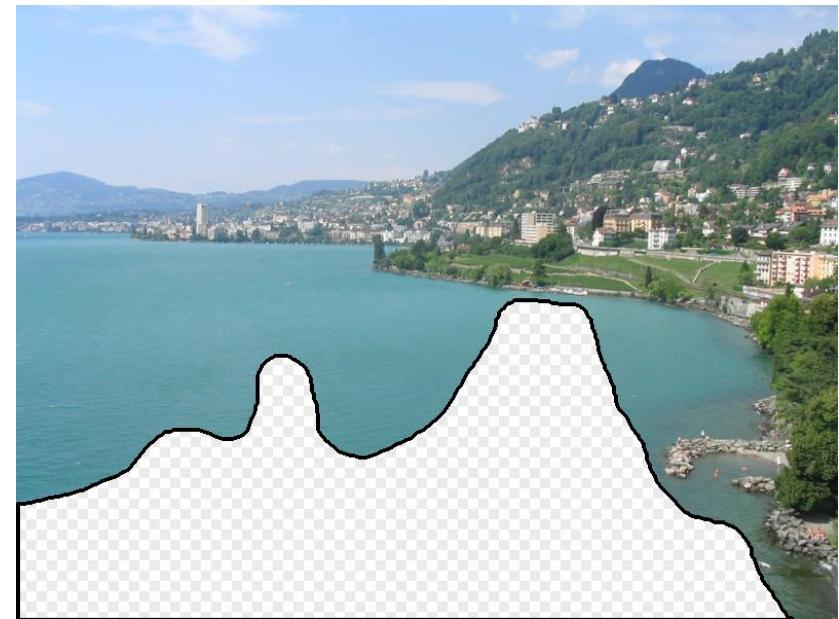


Gist scene descriptor  
(Oliva and Torralba 2001)



... 200 total

# Context Matching



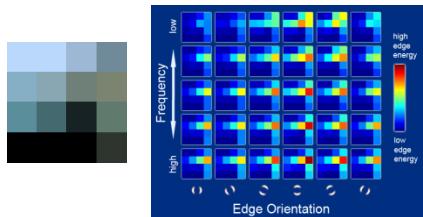


Graph cut + Poisson blending:

Hays and Efros, SIGGRAPH 2007

# Result Ranking

We assign each of the 200 results a score which is the sum of:



The scene matching distance

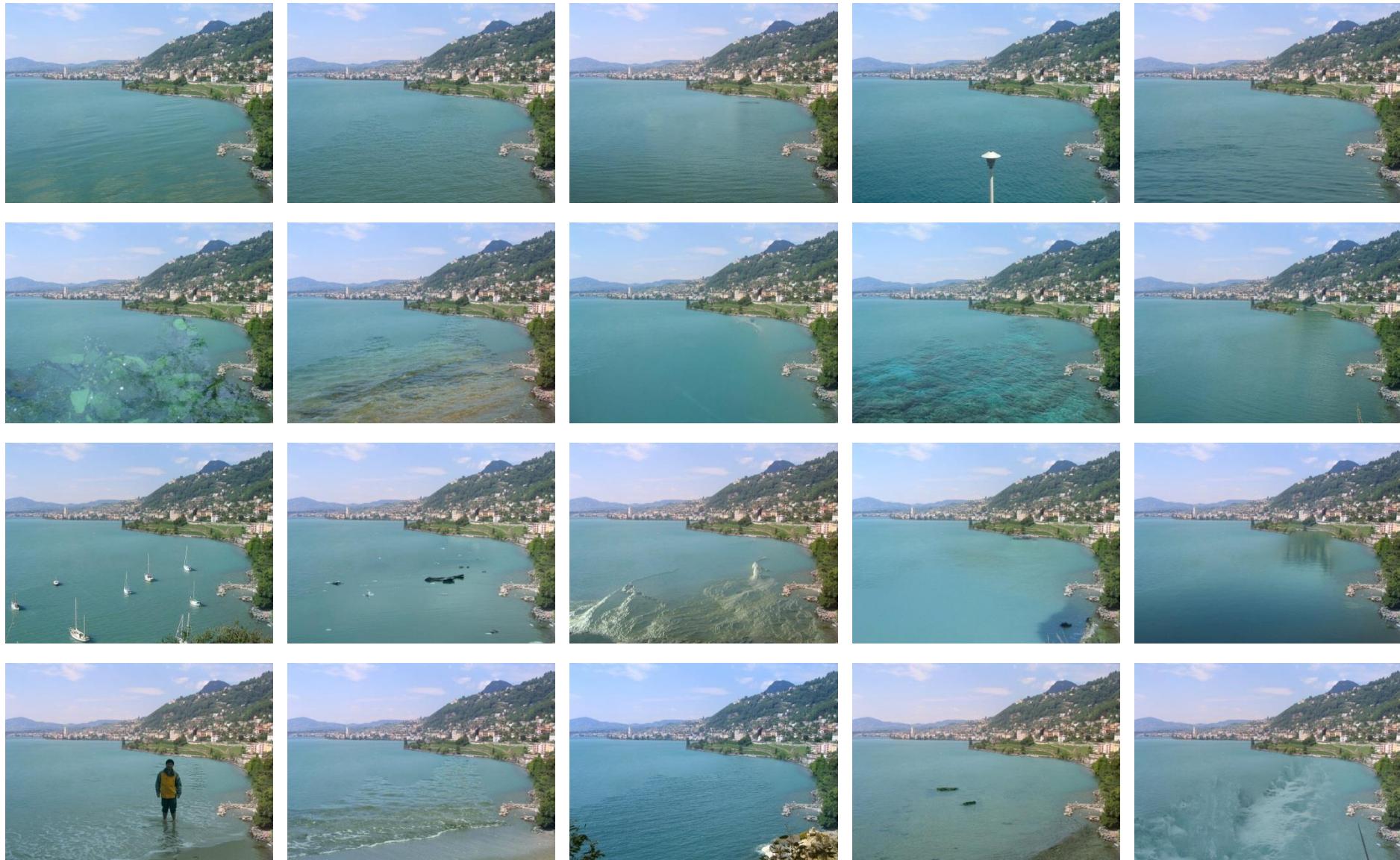


The context matching distance  
(color + texture)



The graph cut cost

# Top 20 Results





Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007

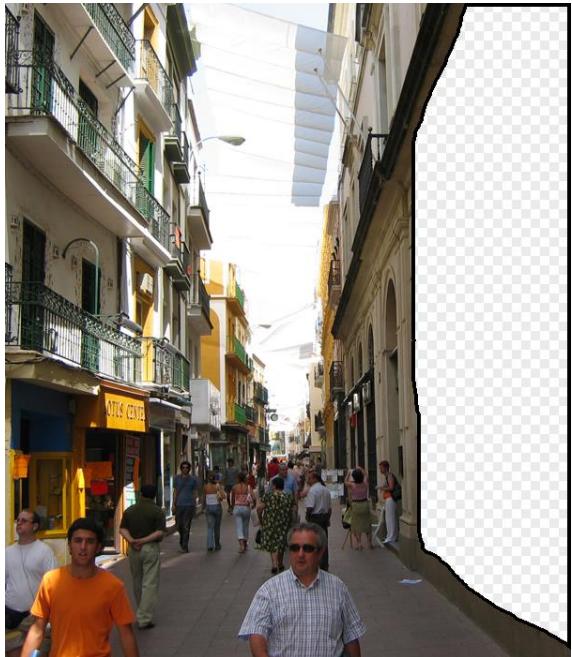


Hays and Efros, SIGGRAPH 2007



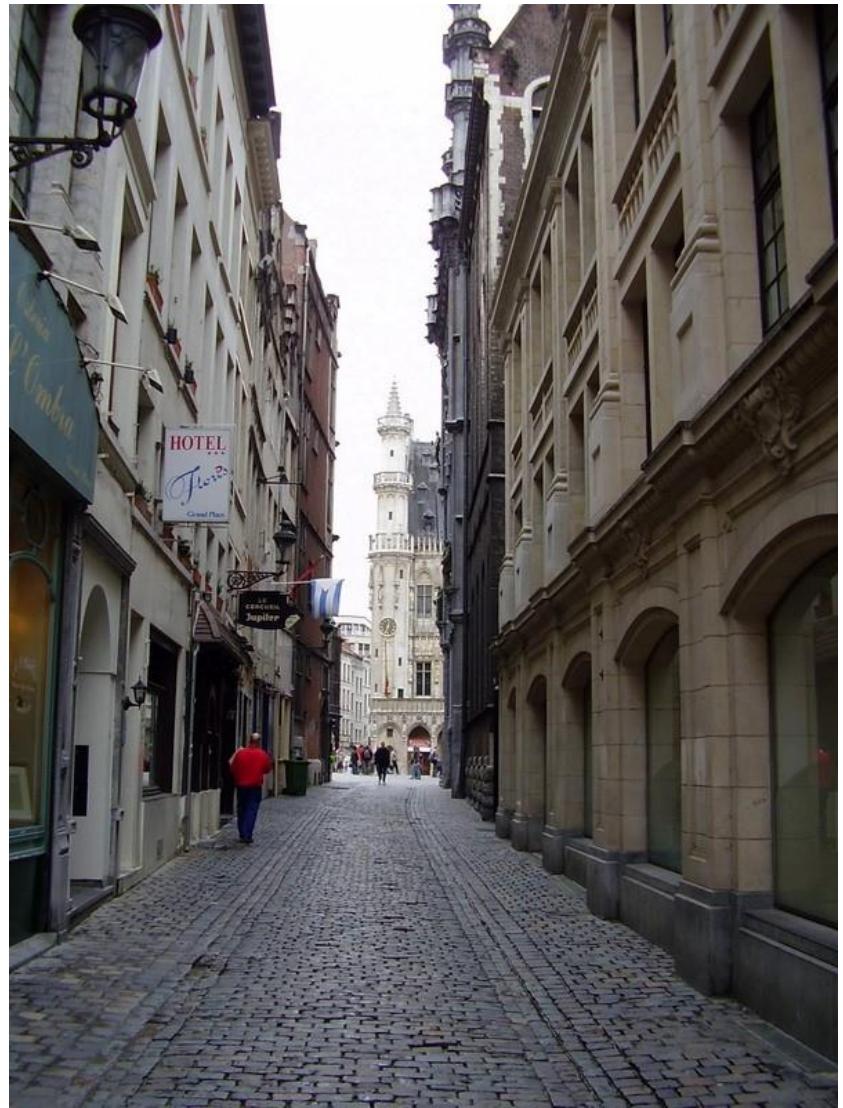
… 200 scene matches

Hays and Efros, SIGGRAPH 2007



... 200 scene matches

Hays and Efros, SIGGRAPH 2007



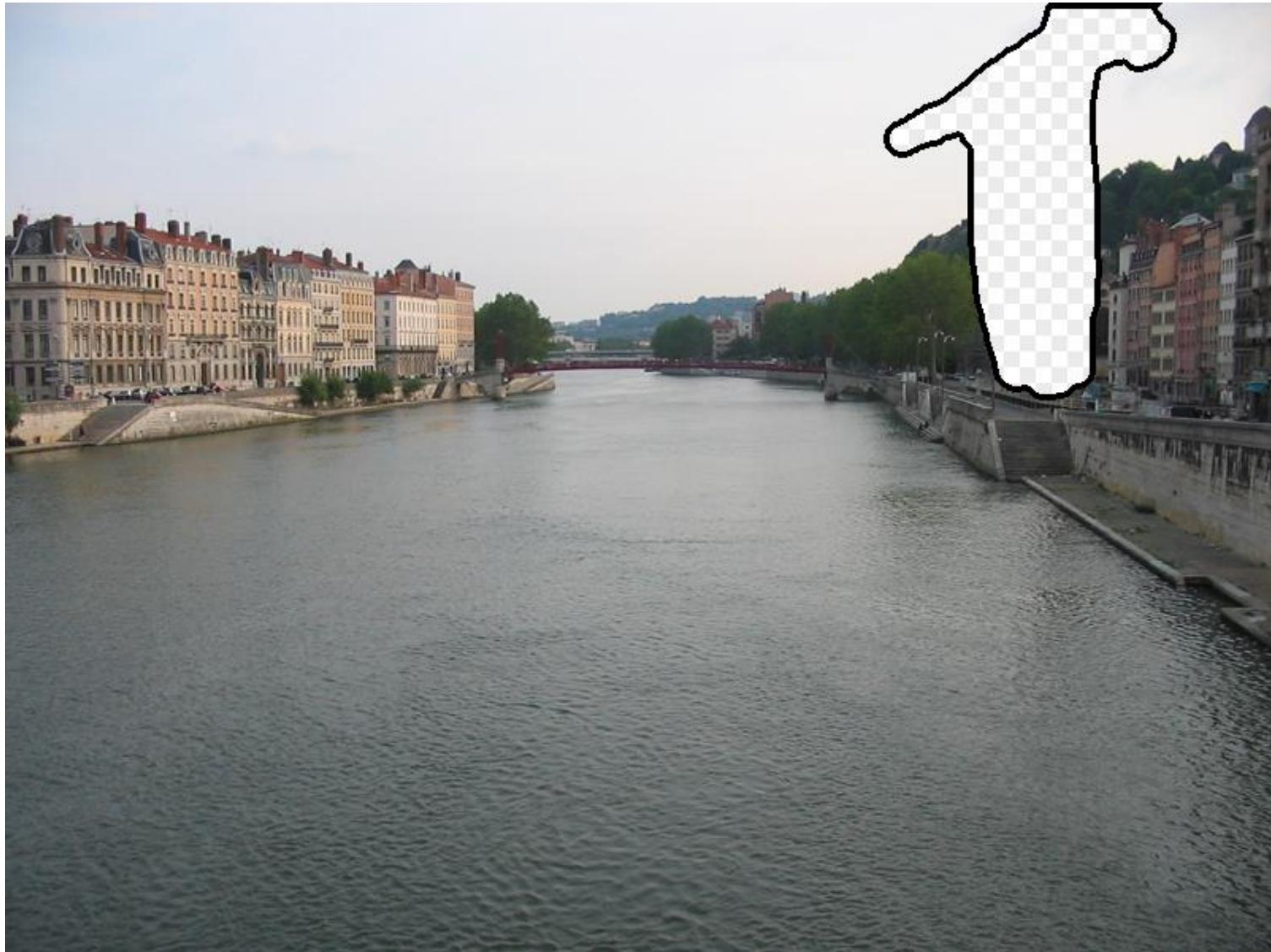
Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007

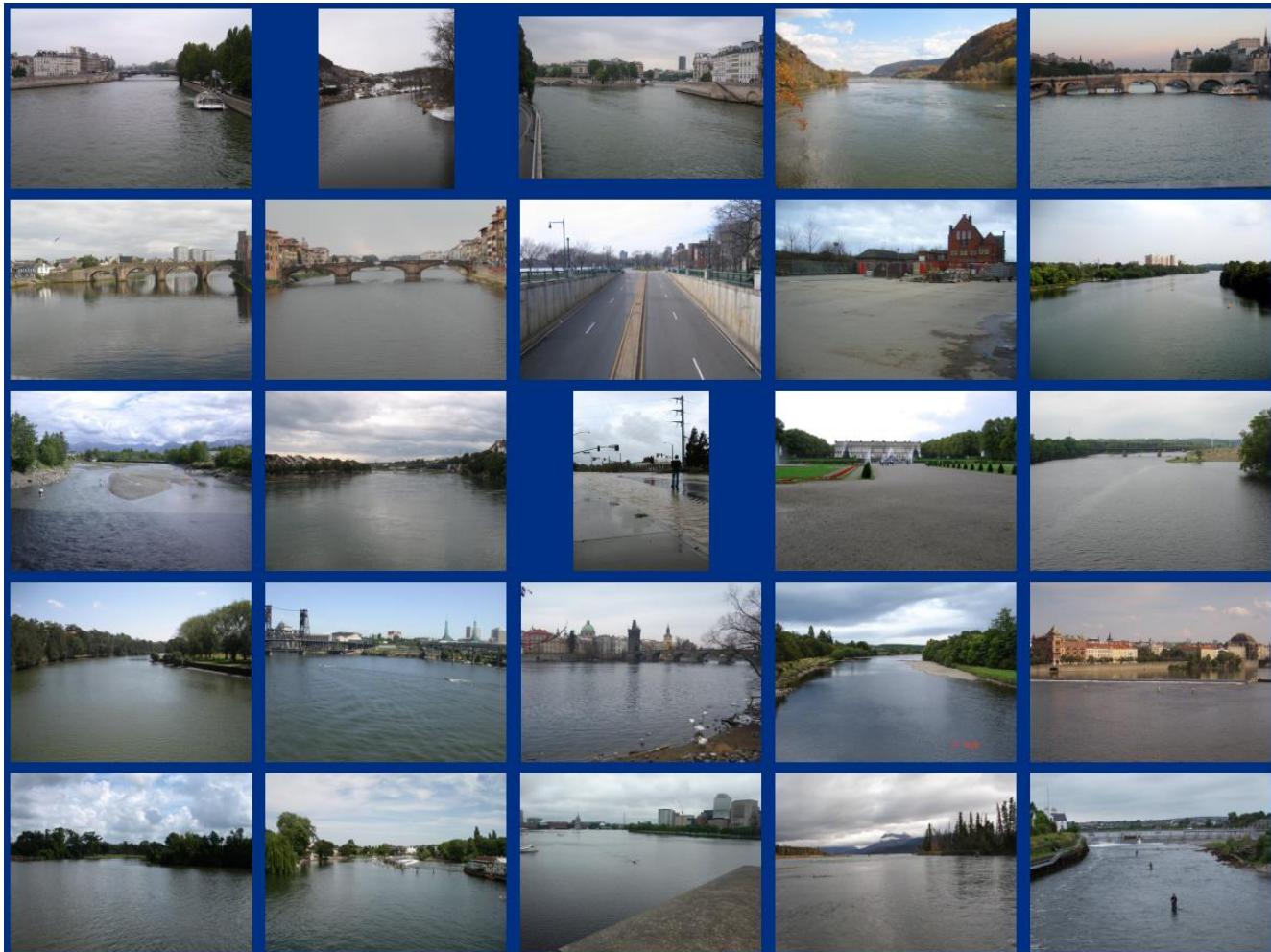
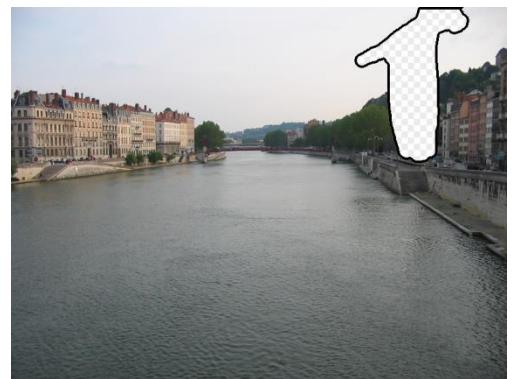


Hays and Efros, SIGGRAPH 2007

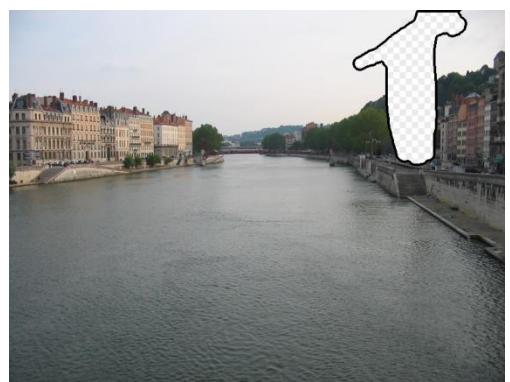




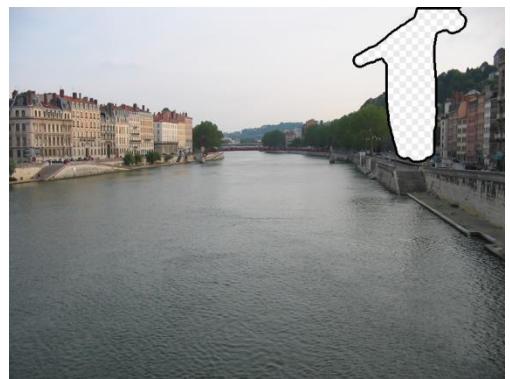
Hays and Efros, SIGGRAPH 2007



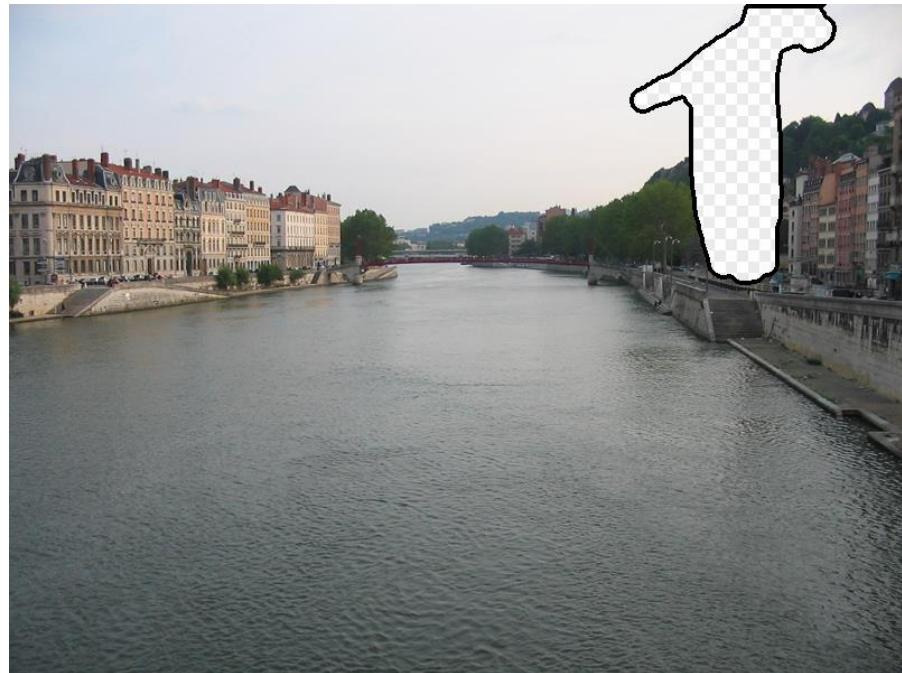
… 200 scene matches



… 200 scene matches



… 200 scene matches





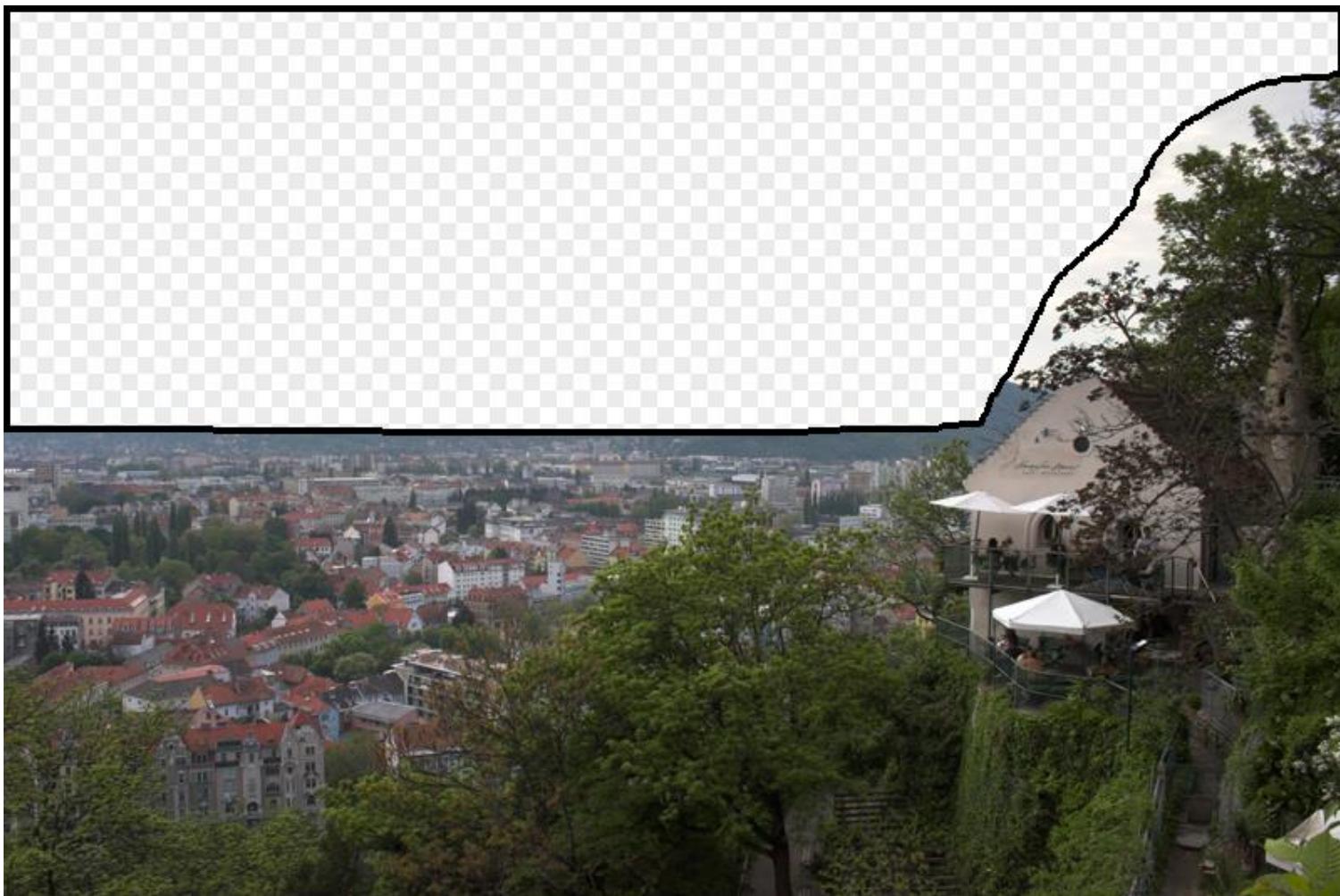
Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007

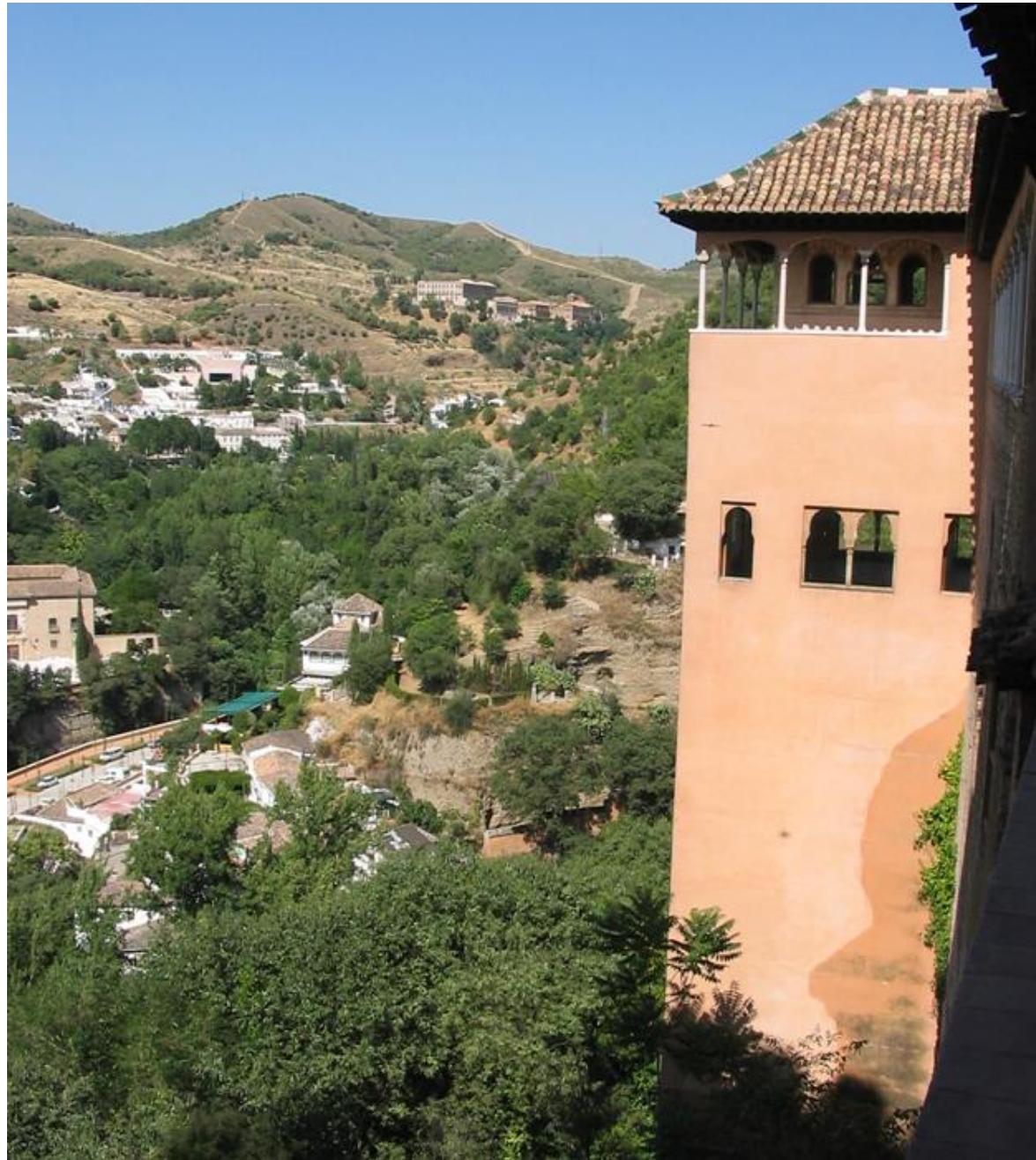


Hays and Efros, SIGGRAPH 2007





Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



Hays and Efros, SIGGRAPH 2007



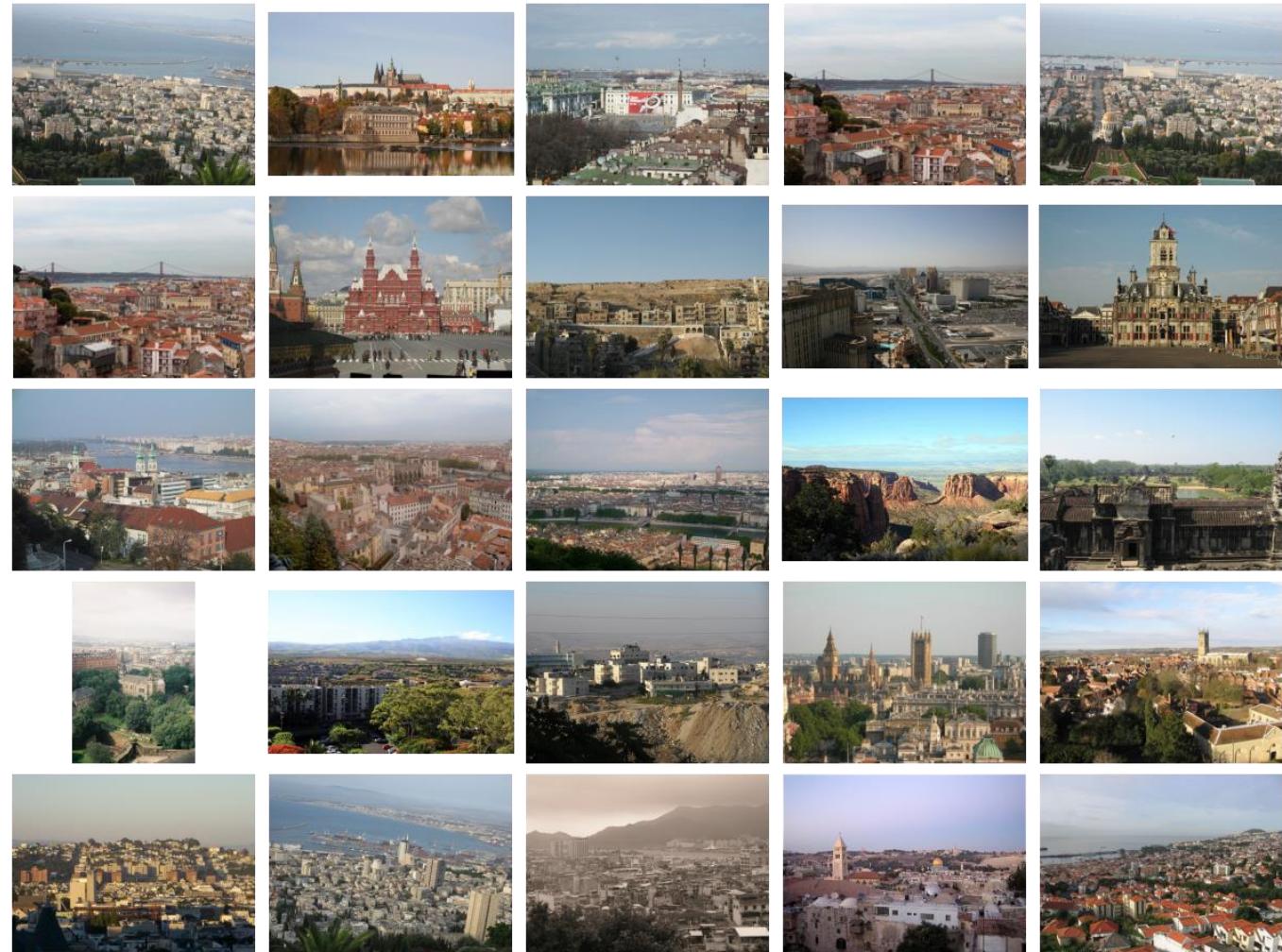
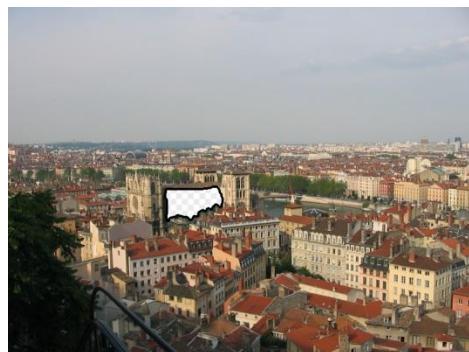
Hays and Efros, SIGGRAPH 2007



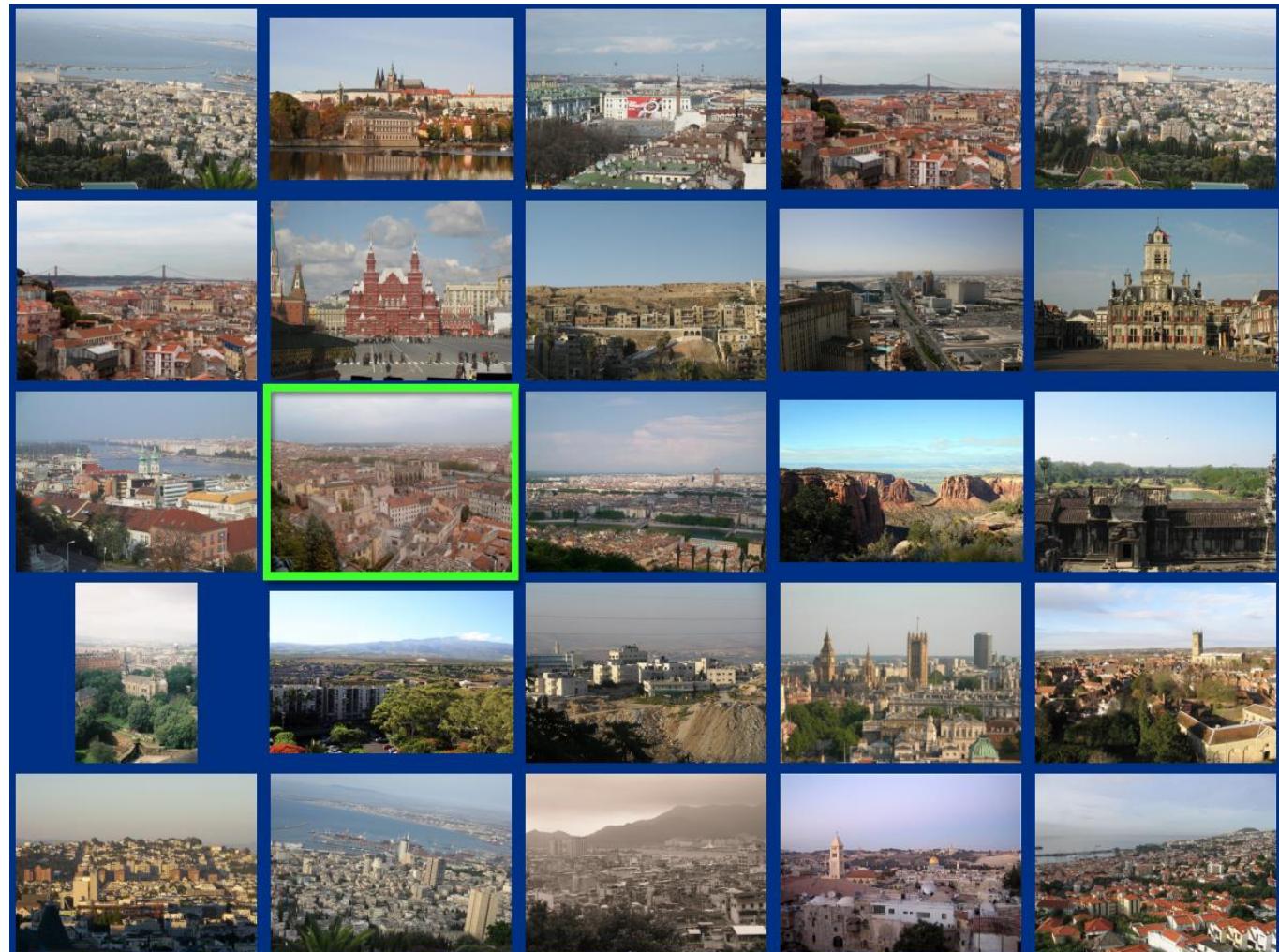
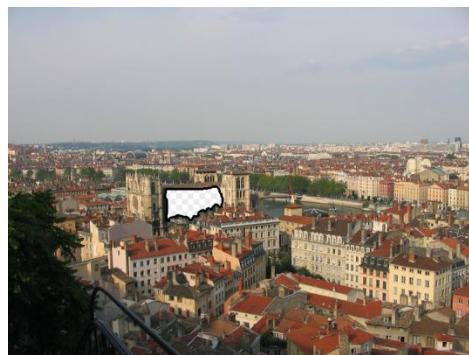
Hays and Efros, SIGGRAPH 2007



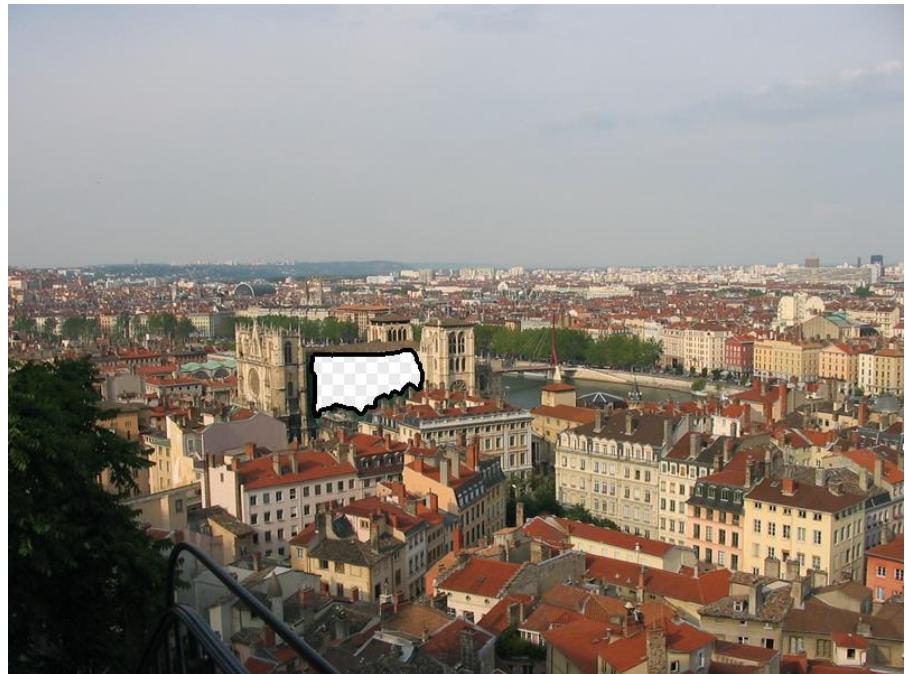
Hays and Efros, SIGGRAPH 2007



… 200 scene matches



… 200 scene matches



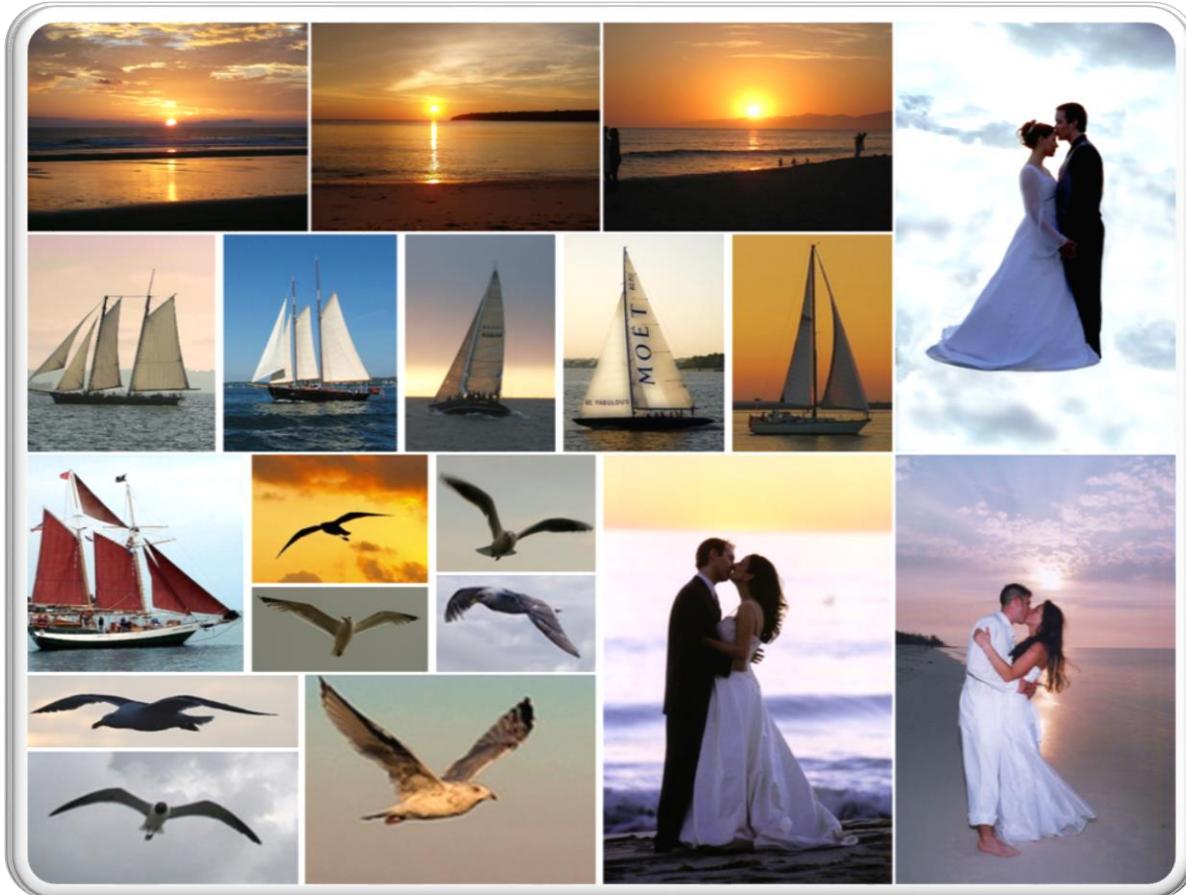


Hays and Efros, SIGGRAPH 2007

# 主要内容

- 基于互联网图像的颜色空间降维技术
  - Data-Driven Color Manifolds. @SIGGRAPH 2015. Chuong H. Nguyen et. al.
- 基于互联网图像的辅助绘制技术
  - ShadowDraw: Real-Time User Guidance for Freehand Drawing @SIGGRAPH 2011. Yong Jae Lee et. al.
- 基于互联网图像的场景漫游
  - Photo tourism: Exploring photo collections in 3D. @SIGGRAPH 2006. Noah Snavely.et.al.
- 基于互联网数据的图像补全
  - Scene Completion Using Millions of Photographs. @SIGGRAPH 2007. J. Hays & A. A. Efros.
- 基于互联网的图像检索与融合
  - Sketch2Photo: Internet Image Montage. @SIGGRAPH ASIA 2009. Tao Chen et. al.

# Motivation



# Motivation



Motivation

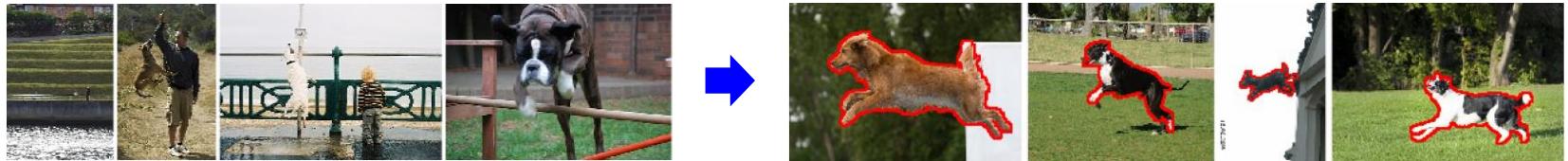
Challenges?

# Challenges

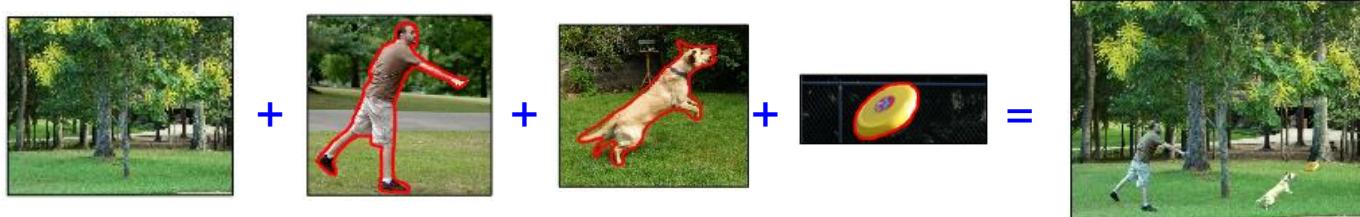
- Automatically find good image components
- Seamlessly compose these components

# Contributions

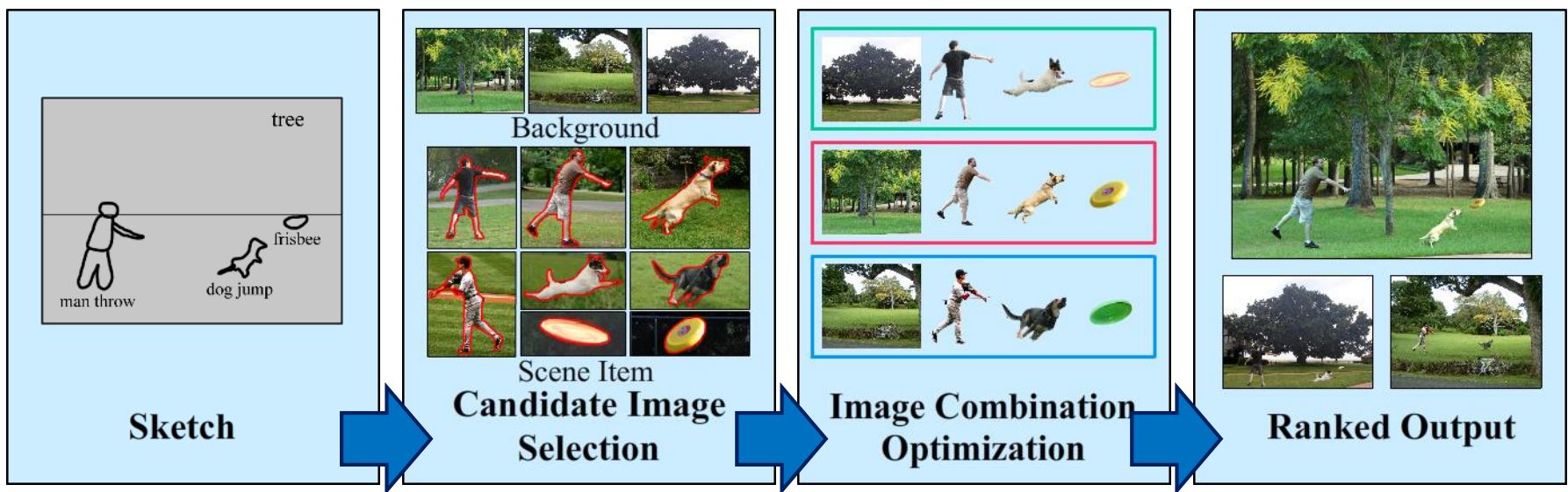
- Automatically find good image components
  - Extensive filtering of internet images



- Seamlessly compose these components
  - Hybrid compositing technique



# System Overview



# Involved Problems

- Image search
- Image segmentation
- Image composition

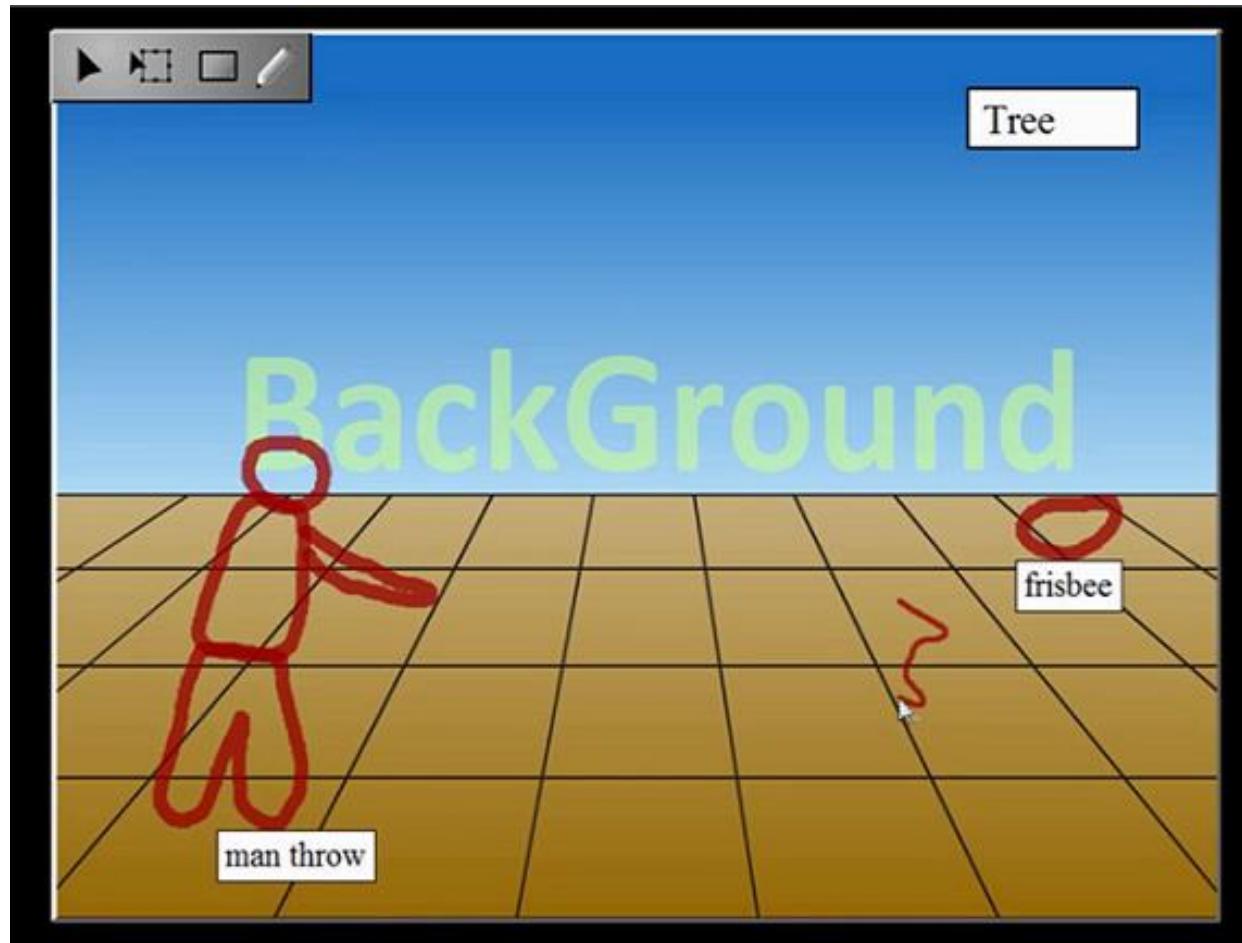
*Difficult problems in general!*

- Key point:

*Use simple images by filtering!*



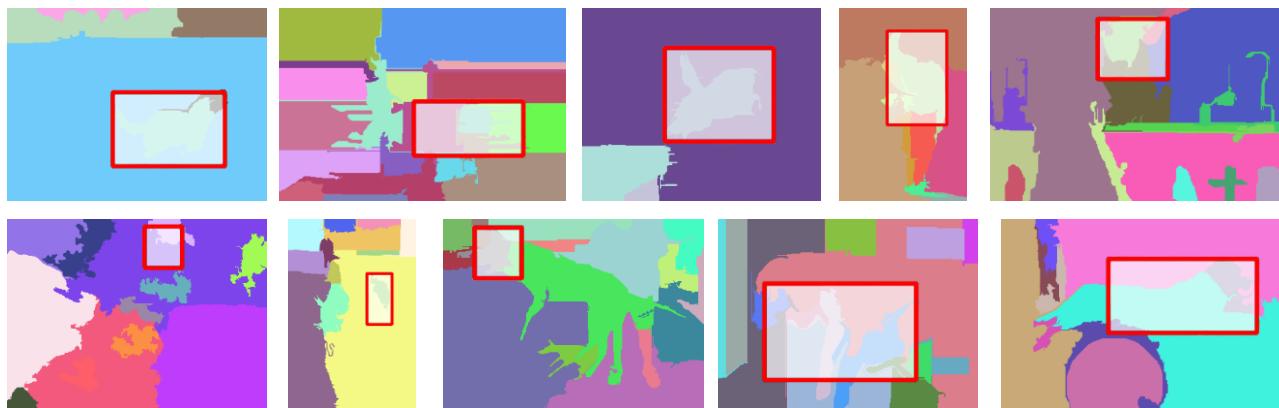
# User Interface



# Image Filtering

## Saliency Filtering

- Identify salient region
- Over-segment all images
- Choose simple images (less segments in *non-salient* region)



# Image Filtering

## Saliency Filtering

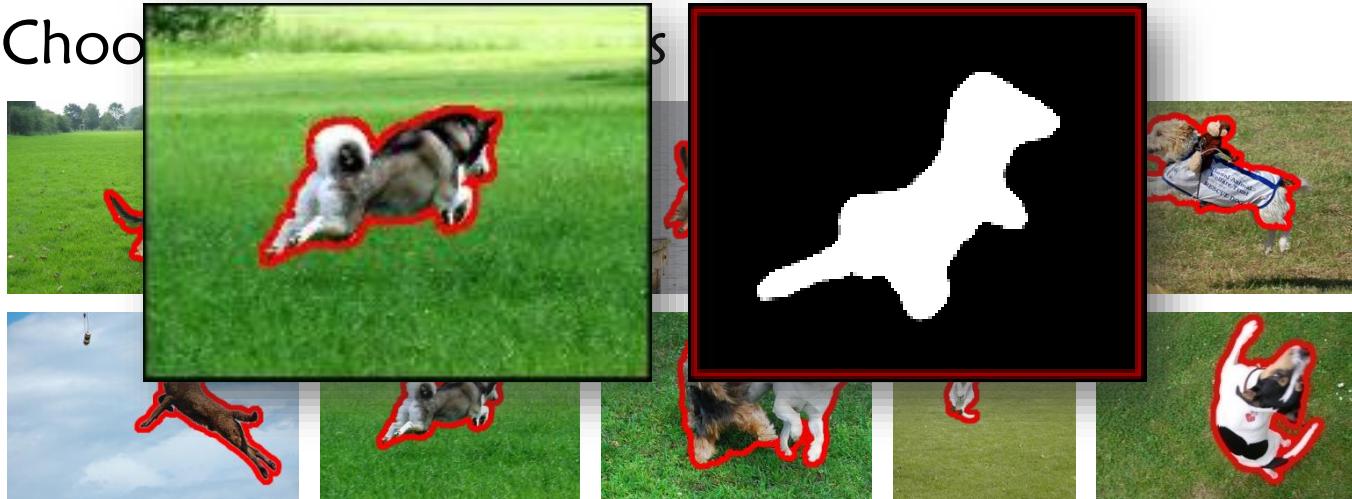
- Identify salient region
- Over-segment all images
- Choose simple images (less segments in *non-salient* region)



# Image Filtering

## Contour filtering

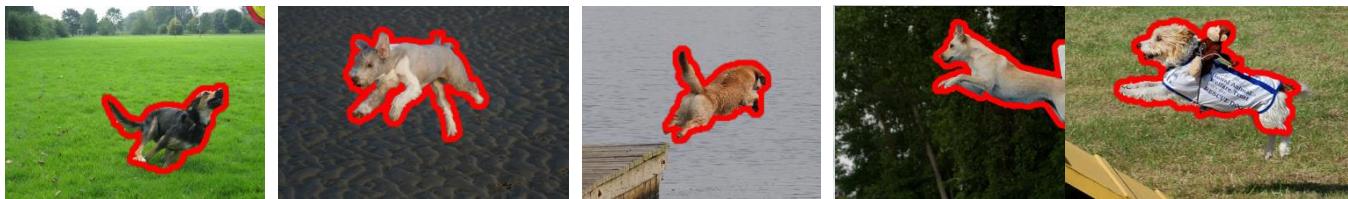
- Image segmentation (by Grabcut [Rother et al 2004])
- Check consistency between segmentation boundary and user sketched contour (by shape-context [Belongie et al. 2002] )
- Choose the best segmentation



# Image Filtering

## Contour filtering

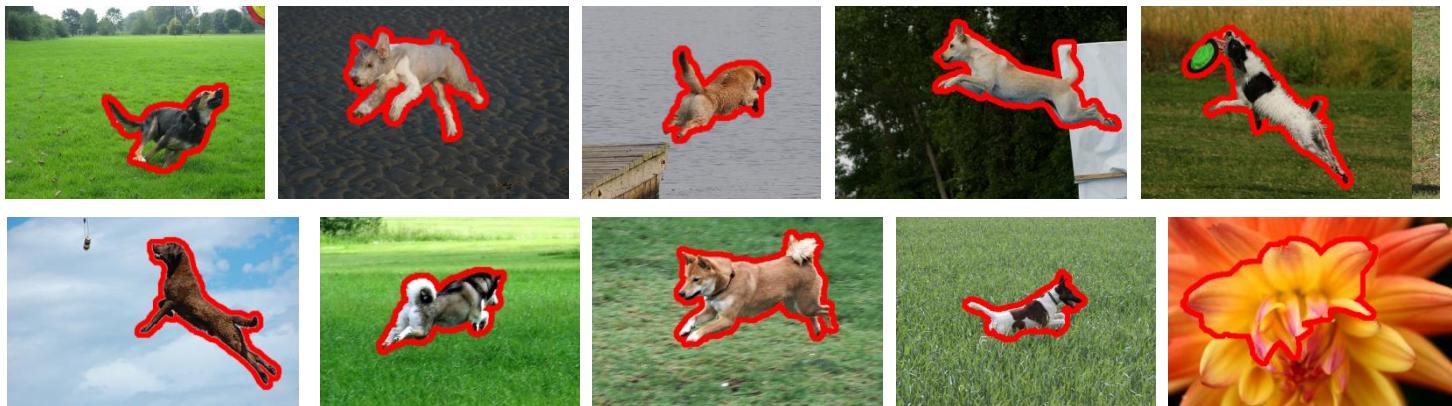
- Image segmentation (by Grabcut [Rother et al 2004])
- Check consistency between segmentation boundary and user sketched contour (by shape-context [Belongie et al. 2002] )
- Choose consistent images



# Image Filtering

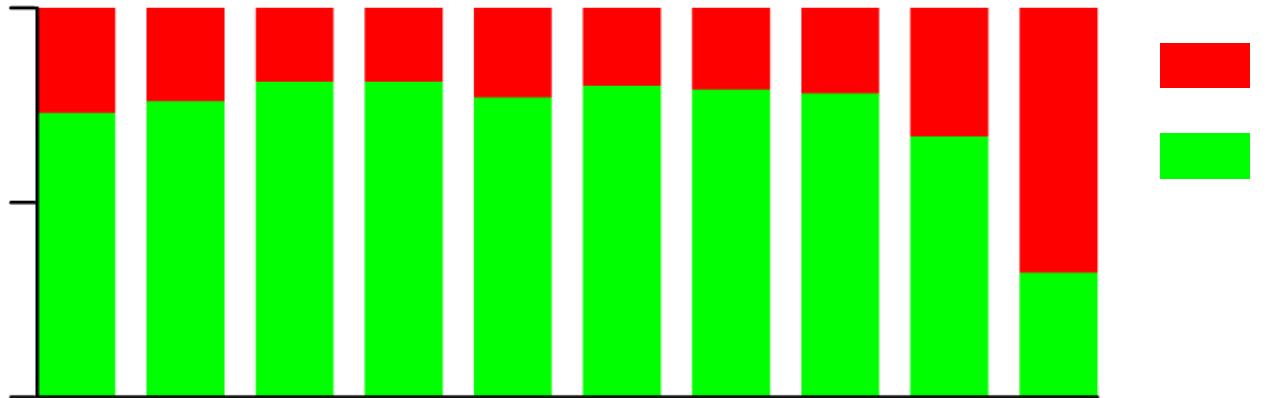
## Content filtering

- Group segmented images in a feature space
- Choose images associated with large ( $>5\%$ ) clusters



# Filtering Performance

## Statistics:



Download Images:



Saliency Filtering:



Contour Filtering:



Content Filtering:



# Filtering Performance

Some observations:

- Contour filtering is critical
- Saliency filtering is important
- Additional verbs are necessary

**False positive rate at each stage of the filtering**

	Man throw	Dog jump	Frisbee	Sailboat	Moto rider	Seagull	Dog
IS (%)	83	65	79	71	86	72	97
SF (%)	81	--	80	61	81	70	97
CF1 (%)	30	47	31	35	27	29	77
CF2 (%)	29	37	27	27	24	23	68

IS = internet search; SF = saliency filtering; CF1 = contour consistency filtering; CF2 = content consistency filtering

# Improved Image Composition

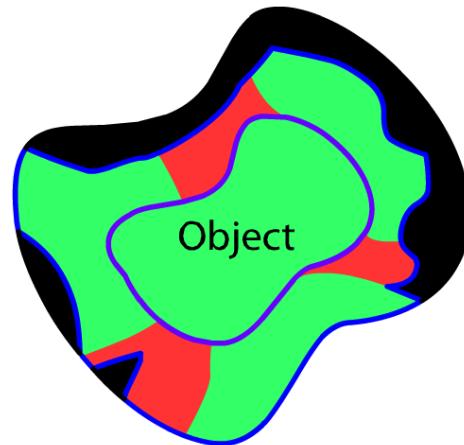
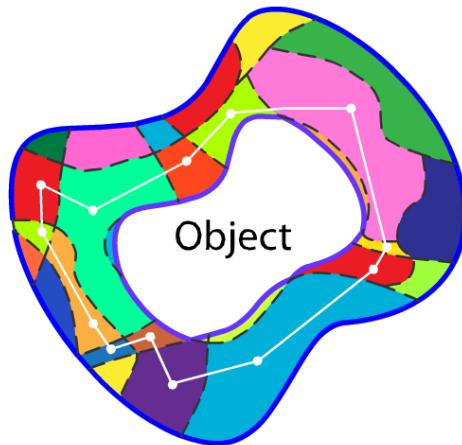
Limitations of previous methods



# Blending Boundary Optimization

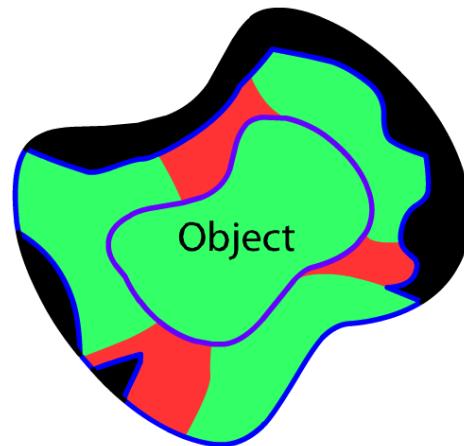
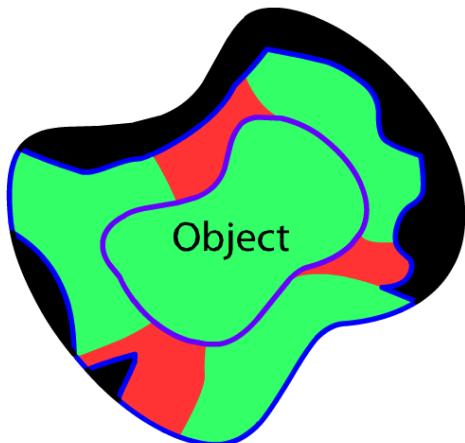
## Super-pixel level boundary optimization

- Evaluate image consistencies at super-pixels
- Find a most consistent chain of super-pixels
- Mark out inconsistent super-pixels (shown in red)



# Blending Boundary Optimization

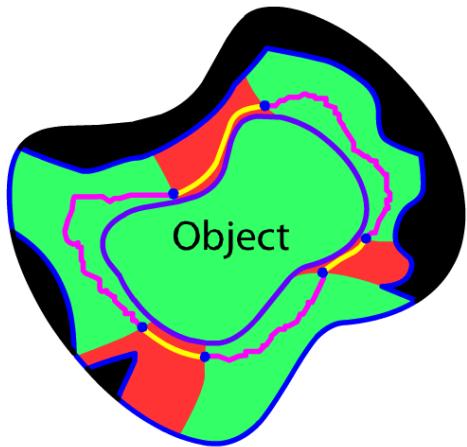
Pixel level boundary optimization



# Blending Boundary Optimization

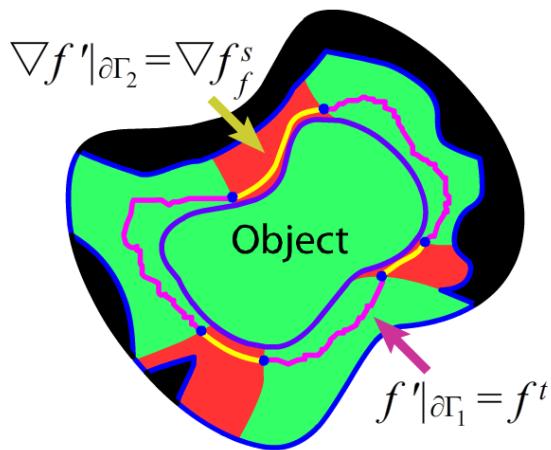
## Pixel level boundary optimization

- Boundary in red cells: matting boundary
- Boundary in green cells: optimized by Drag & Drop Pasting [Jia et al. 2006]



# Hybrid Blending

- Poisson blending with mixed boundary condition

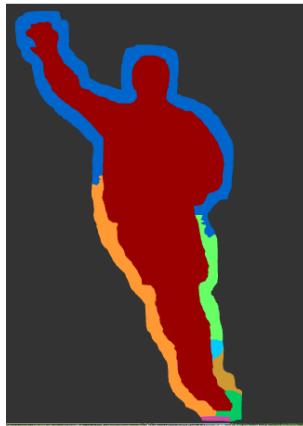


$$\min_{f'} \int_{p \in \Omega} |\nabla f' - v|^2 dp$$

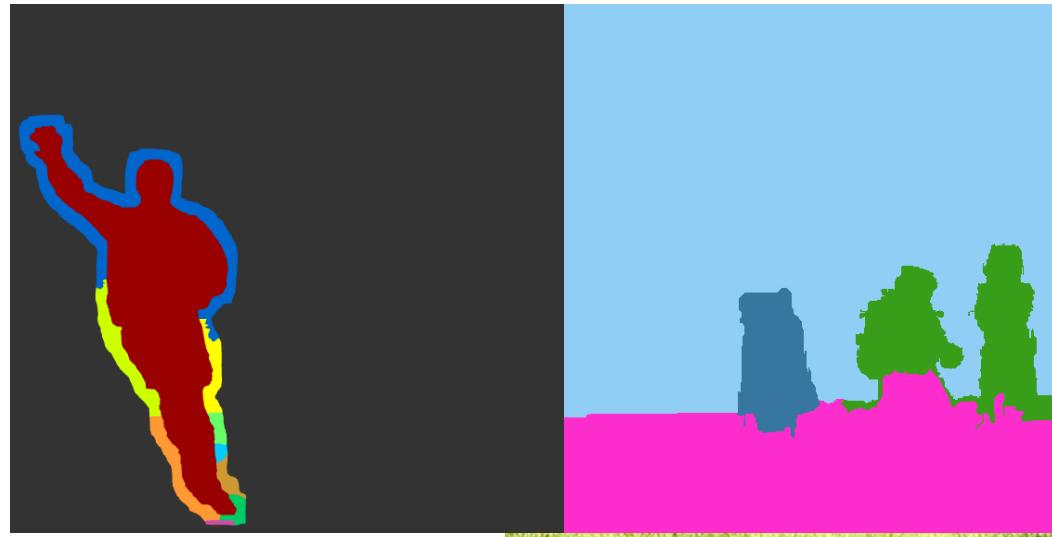
- Fix pixel value on purple boundary
- Fix pixel gradient on yellow boundary
- Alpha blending

$$f(p) = \begin{cases} f' & \text{if } p \in \text{Green Zone} \\ \alpha f'(p) + (1 - \alpha) f^t(p) & \text{if } p \in \text{Red Zone} \end{cases}$$

# Example



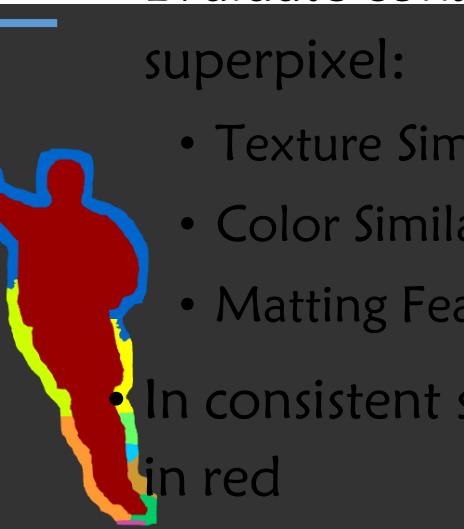
Segments  
background  
blending boundary



Superpixel

Target image  
Segments  
background

# Example



- Evaluate consistency for each superpixel:
  - Texture Similarity
  - Color Similarity
  - Matting Feasibility
- Inconsistent super-pixels are marked in red
- Optimize pixelwise blending  
Superpixel boundary
- Final blending result

# More Blending Comparisons



Input

Matting

Drag&Drop

Photo Clip Art

Our Result

# Image Combination Optimization

Optimize image combination

- Rank by the consistency over the optimized super-pixel boundary



Blending cost is 0.2, 0.4, 0.6 and 0.8 from left to right.

# Image Combination Optimization

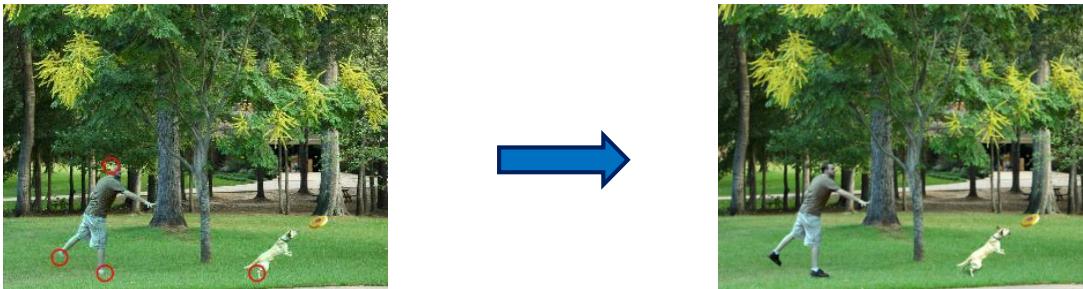
	Tree	Man throw	Dog jump	Frisbee
Image 0				
Image 1				
Image 2				
Image 3				
Image 4				
.....	.....	.....	.....	.....

# Interactive Refinement

- Discard compositions with incorrect scene items



- Refine segmentation



# More Results



# More Results



**Thank you!**