

# Combining Appearance and Geometry for Representation of Landmark Scenes

Svetlana Lazebnik  
Department of Computer Science  
University of North Carolina at Chapel Hill

Joint work with Xiaowei Li, Changchang Wu, Rahul Raguram,  
Christopher Zach, and Jan-Michael Frahm

























# Motivation

- Develop efficient representations and algorithms for landmark photo collections on the Internet

## “Statue of Liberty” on Flickr

✓ We found **79,565 results** matching "statue of liberty".

View: [Most relevant](#) • [Most recent](#) • [Most interesting](#)    Show: [Details](#) • [Thumbnails](#)

 From <a href="#">Scandblue</a>	 From <a href="#">saxtopia</a>	 From <a href="#">Odalaiigh</a>	 From <a href="#">Steve W Lee</a>	 From <a href="#">heath_bar</a>	 From <a href="#">pedals</a>
 From <a href="#">pedals</a>	 From <a href="#">Stumblingthr...</a>	 From <a href="#">Scandblue</a>	 From <a href="#">Scandblue</a>	 From <a href="#">Scandblue</a>	 From <a href="#">pixelpixel</a>
 From <a href="#">satosphere</a>	 From <a href="#">Scandblue</a>	 From <a href="#">featherboa</a>	 From <a href="#">Scandblue</a>	 From <a href="#">Scandblue</a>	 From <a href="#">savioriosh</a>
 From <a href="#">gmg_660</a>	 From <a href="#">Scandblue</a>	 From <a href="#">crashcallowa...</a>	 From <a href="#">Scandblue</a>	 From <a href="#">docnad</a>	 From <a href="#">docnad</a>

# Motivation

- Develop efficient representations and algorithms for landmark photo collections on the Internet
  - Key concept: *iconic images*



# Motivation

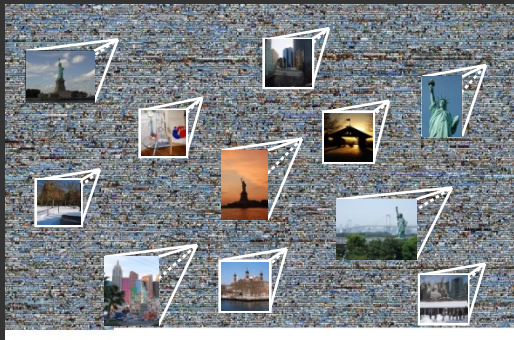
- Develop efficient representations and algorithms for landmark photo collections on the Internet
  - Key concept: *iconic images*
  - Applications: summarization, browsing, reconstruction, recognition

# Previous Work

- **Appearance-based:** Berg & Forsyth (2007), Kennedy & Naaman (2008)
- **Geometry-based:** Snavely et al. (2006), Simon et al. (2007)
- **Our approach:** combine loose 2D appearance constraints and strong 3D geometry constraints to get both efficiency and accuracy

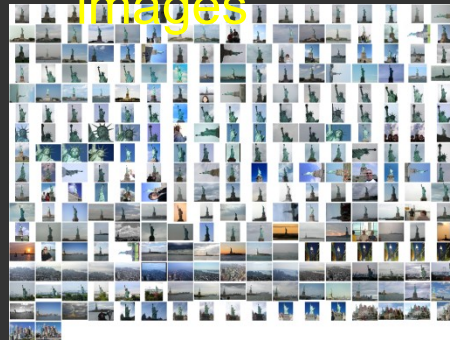
# Our Approach

All images



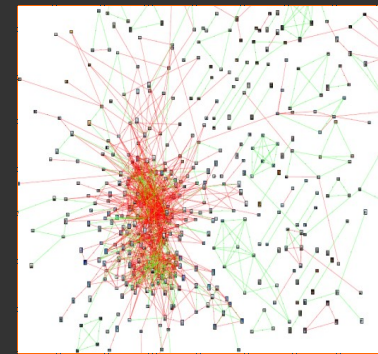
Appearance-based clustering, geometric verification

Iconic images



Pairwise matching of iconic images

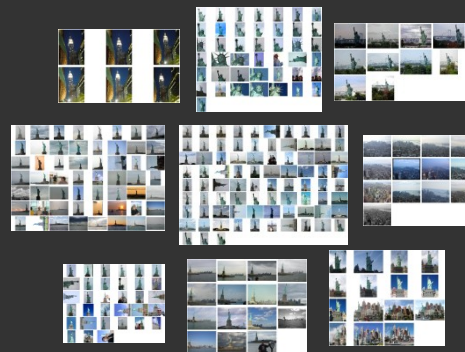
Scene summary



Iconic scene graph

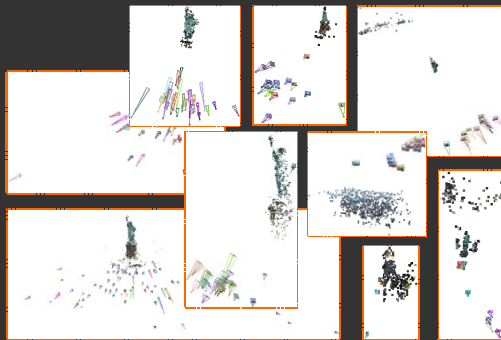
Graph cut

Components of iconic scene graph



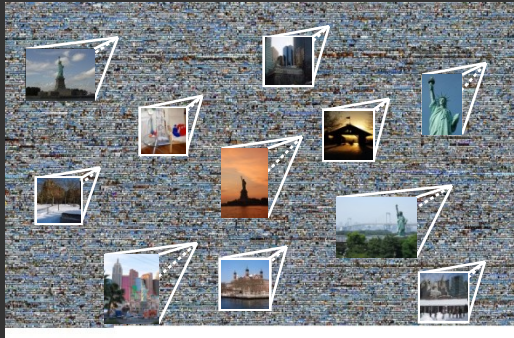
SFM

Component models



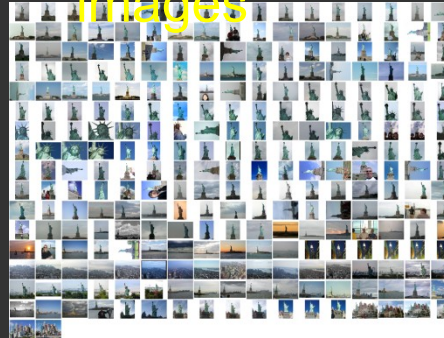
# Our Approach

All images

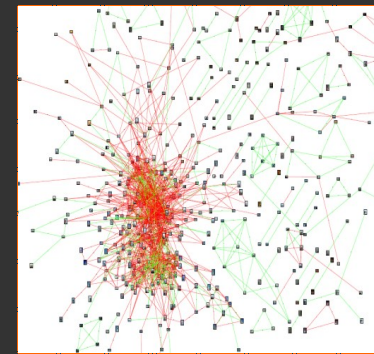


Appearance-based clustering, geometric verification

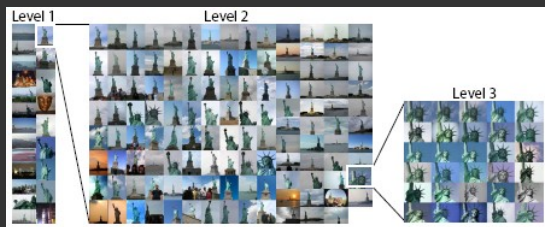
Iconic images



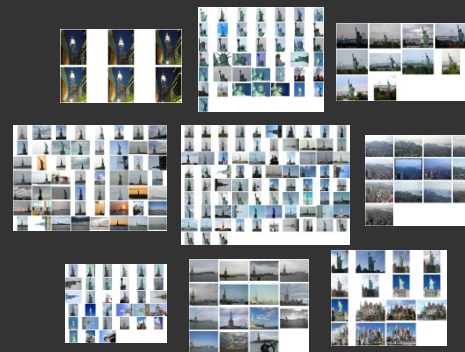
Pairwise matching of iconic images



Scene summary  
Location recognition



Hierarchical browsing



Components of iconic scene graph

Graph cut

Iconic scene graph

# Details

1. **Appearance-based clustering**
  - k-means with GIST descriptor (Oliva & Torralba, 2001)



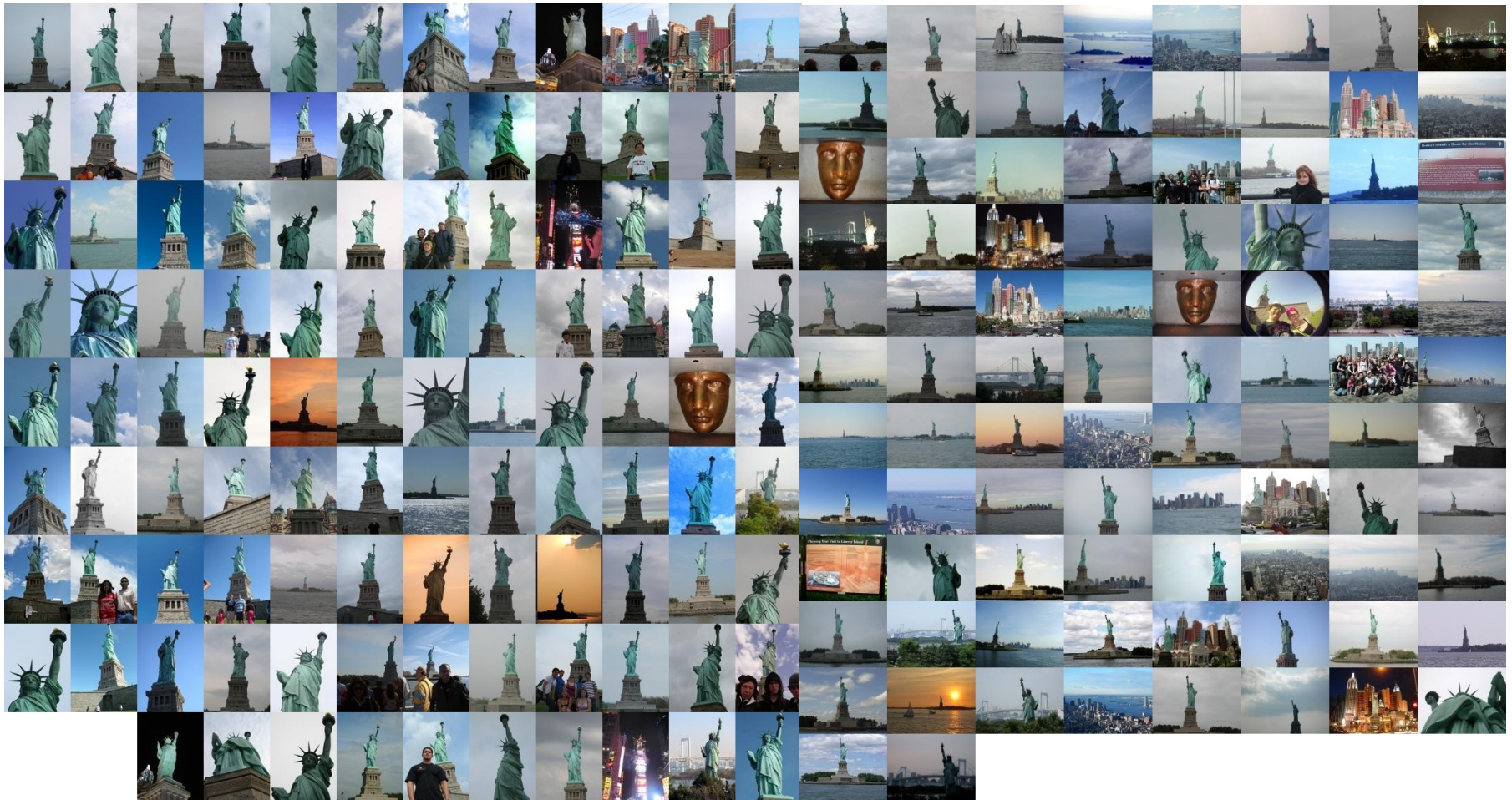
# Details

1. Appearance-based clustering
2. Geometric verification and iconic image selection
  - Perform feature-based geometric matching between a few “top” images from each cluster
  - QDEGSAC (Frahm & Pollefeys, 2006) for robust estimation of fundamental matrix or homography
  - Select an *iconic image* for each cluster as the image with the most total inliers

# Details

1. Appearance-based clustering
2. Geometric verification and iconic image selection

Statue of Liberty: 45284 images, 196 iconics

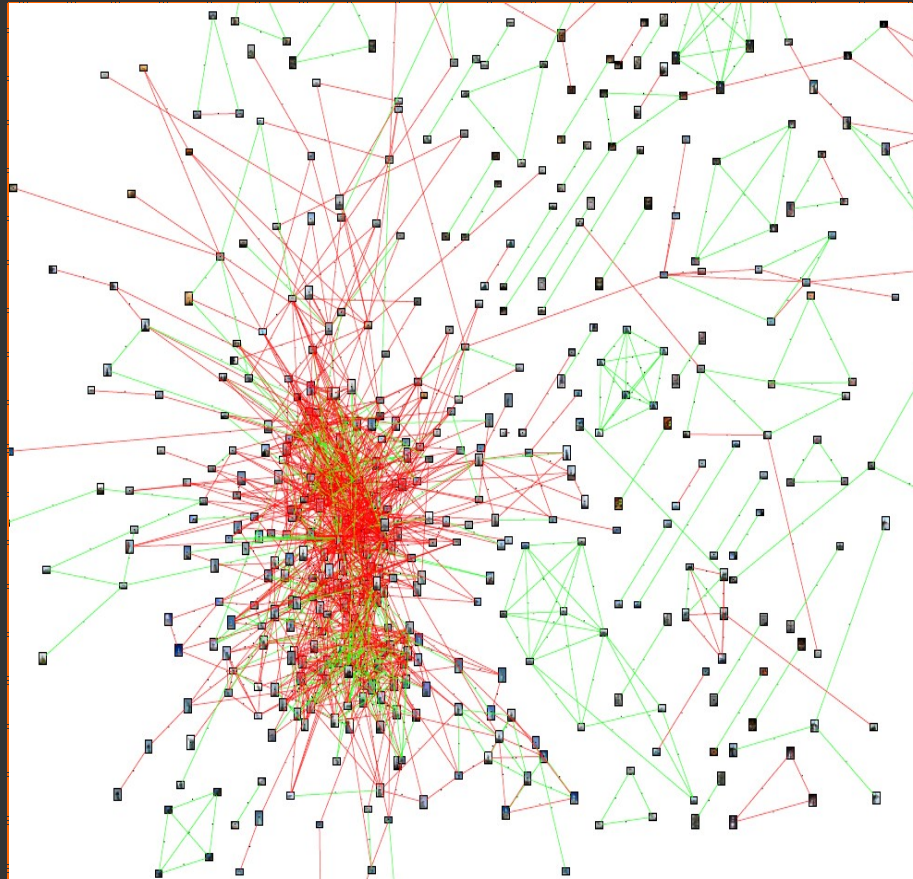


# Details

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph
  - Perform geometric matching between every pair of iconic images
  - Create a weighted edge for every pair related by a homography or a fundamental matrix

# Details

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph

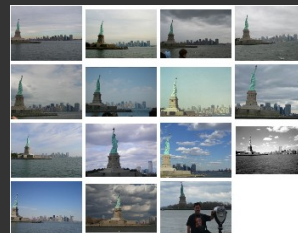
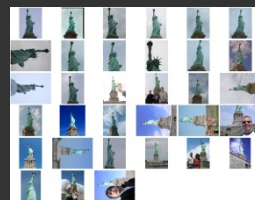
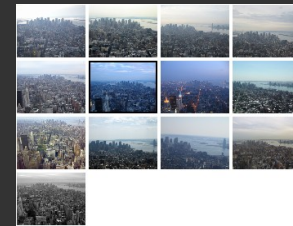
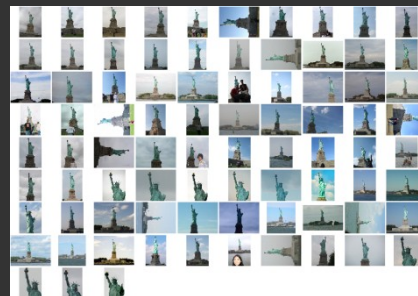
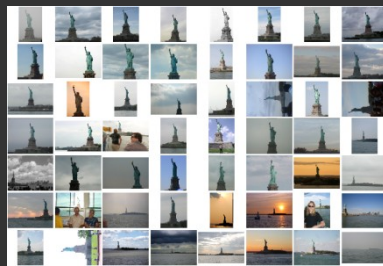
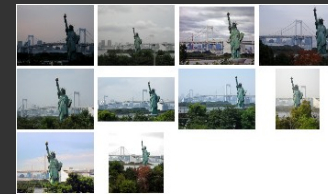


# Details

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph
4. Finding graph components
  - Run normalized cuts to break up the rest of the graph into smaller tightly connected sub-graphs

# Details

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph
4. Finding graph components



# Details

1. Appearance-based clustering
2. Geometric verification and iconic image selection
3. Construction of iconic scene graph
4. Finding graph components
5. Structure from motion
  - Perform SFM separately on each component
  - Maximum-weight spanning tree determines the order of incorporating images into the 3D model
  - If possible, merge component models using geometric relationships along edges that were originally cut
  - Register additional non-iconic images to the models

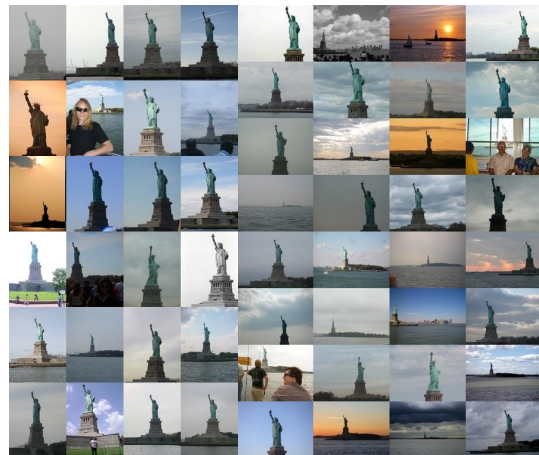
# Structure from Motion: Statue of Liberty

45284 images

## Las Vegas



## Tokyo



## New York

Registered images in largest model: 871

Points visible in 3+ views: 18675

# Hierarchical browsing

- Level 1: components of iconic scene graph
- Level 2: iconic images belonging to each component
- Level 3: images inside the gist cluster of each iconic

Level 1



Level 2



Level 3



# Location Recognition

- Given a new test image, we want to efficiently register it to the model and find out whether or not it contains the landmark
  - Retrieve  $k$  iconics most similar to the test image using either GIST or vocabulary tree (bag-of-features matching) and perform geometric verification
  - The final recognition score is the number of inliers to the most similar iconic

# “Easy” example



Query

GIST



Score: 24



Score: 58



Score: 21



Score: 7



Score: 7



Score: 30



Score: 21



Score: 34



Score: 10



Score: 58

Vocabulary Tree

# Another easy example

GIST



Score: 38



Score: 7



Score: 7



Score: 12



Score: 20

Query



Score: 34



Score: 52



Score: 40



Score: 21



Score: 16

Vocabulary Tree

# Difficult example



Query

GIST



Score: 10



Score: 7



Score: 0



Score: 9



Score: 9



Score: 7



Score: 9



Score: 9



Score: 9



Score: 7

Vocabulary Tree

# Difficult example



Query

GIST



Score: 0



Score: 0



Score: 8



Score: 7



Score: 8



Score: 0



Score: 0



Score: 0



Score: 0

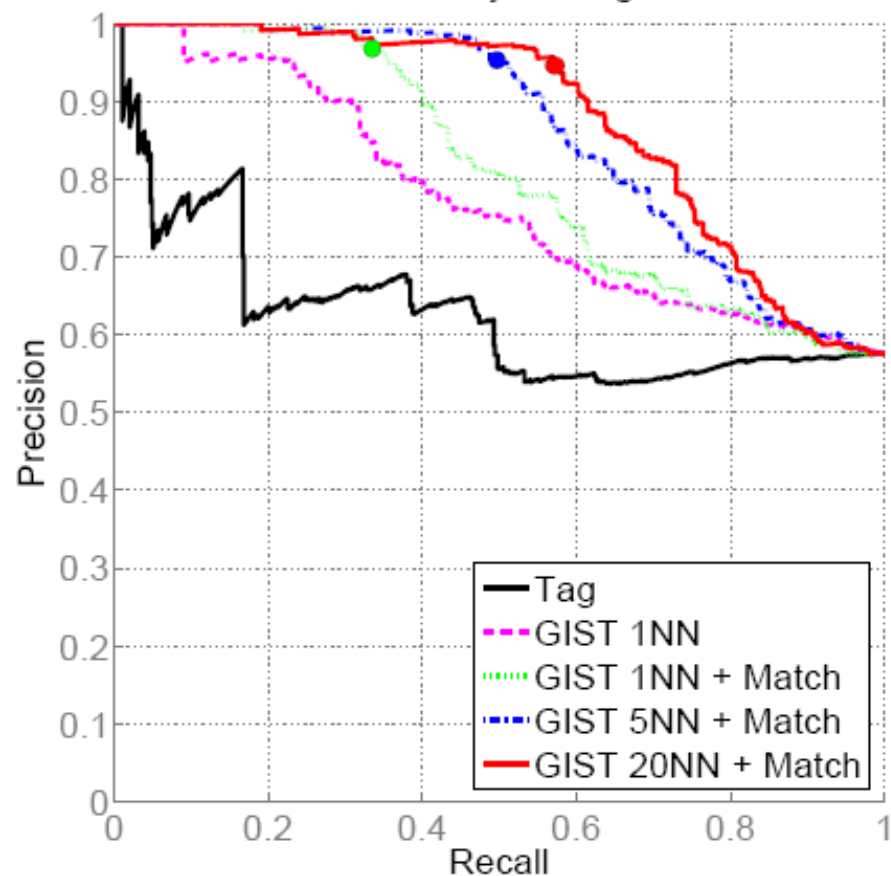


Score: 7

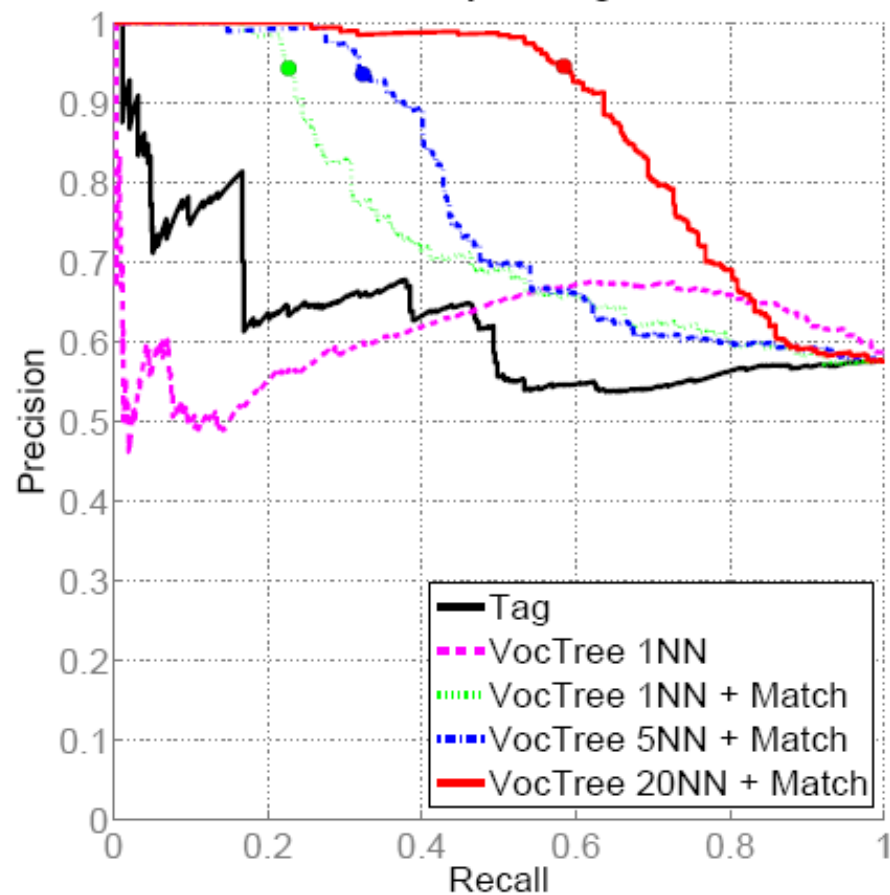
Vocabulary Tree

# Quantitative results

Statue of Liberty Testing - GIST



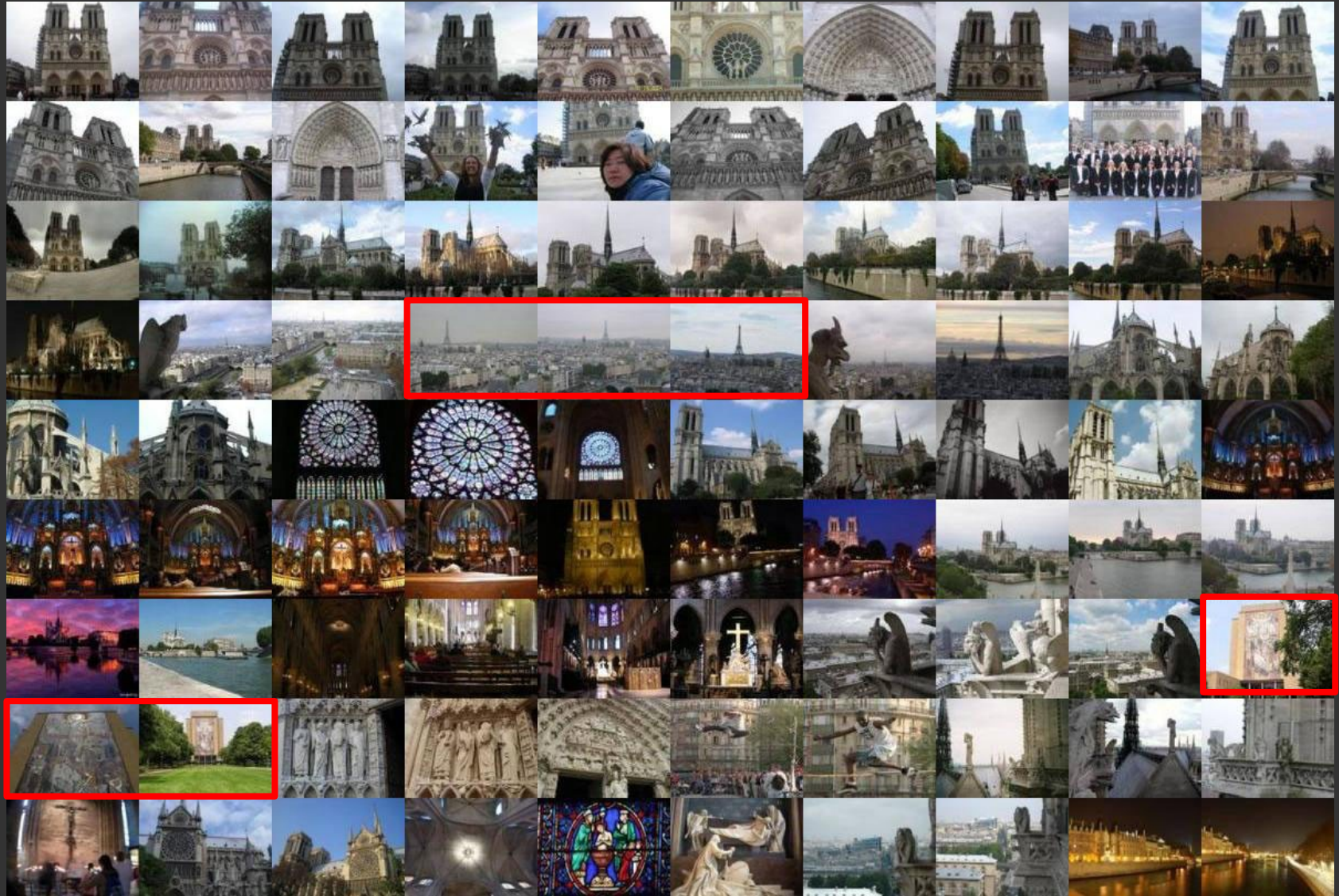
Statue of Liberty Testing - VocTree





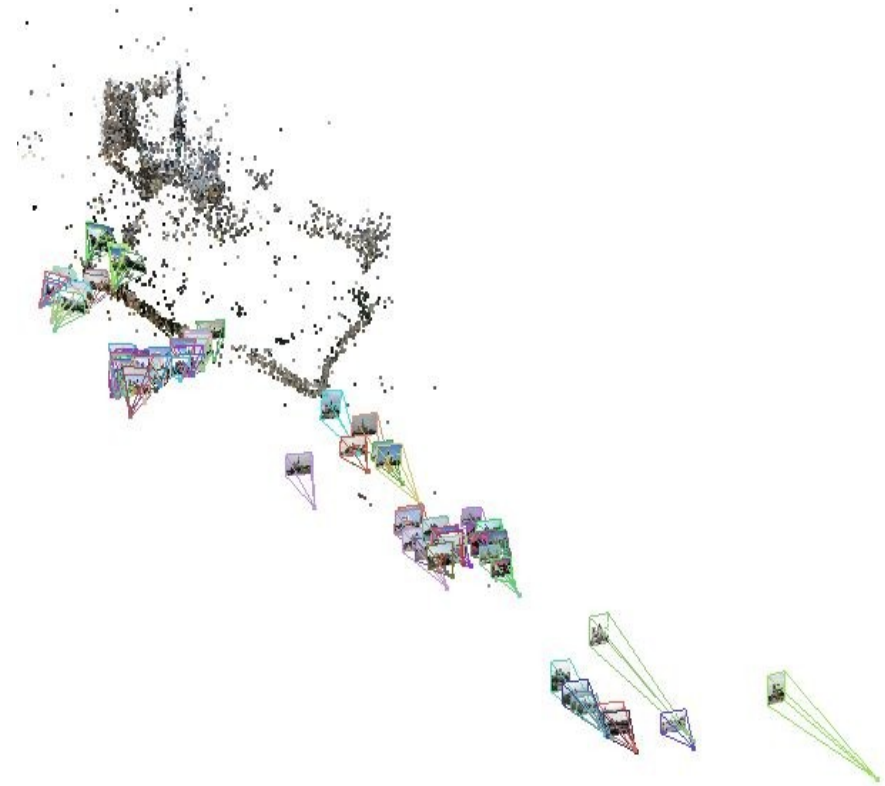
# Notre Dame dataset

10840 images originally, 162 iconics



# Notre Dame dataset

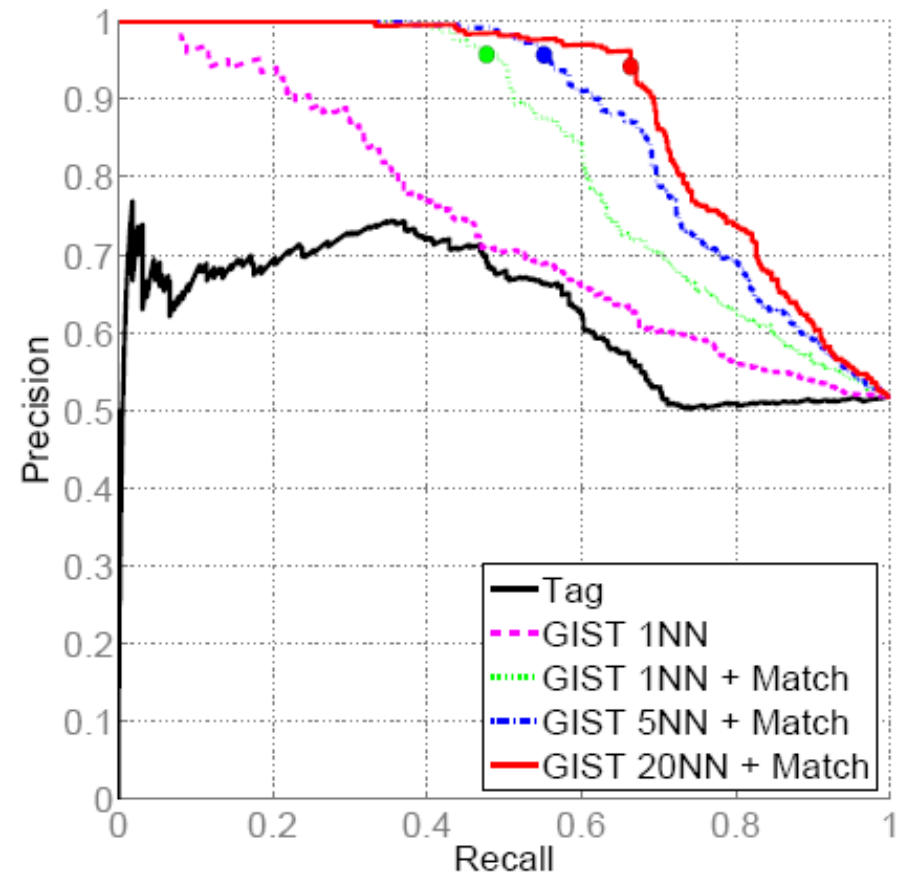
10840 images originally, 162 iconics



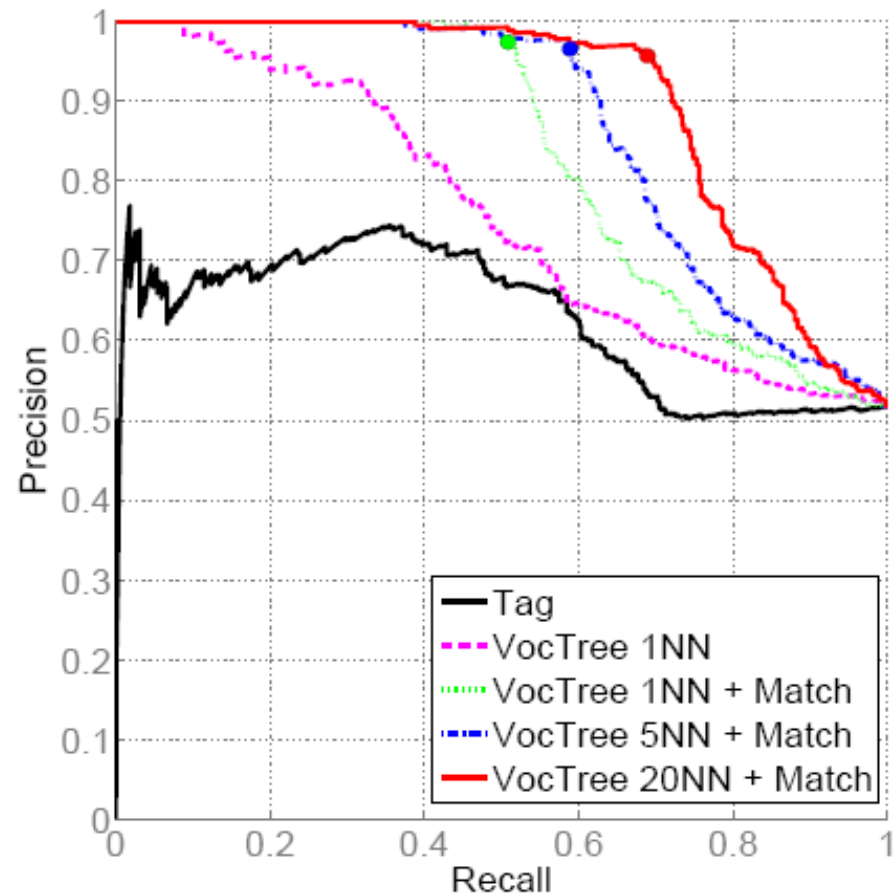
Registered images in largest model: 337  
Points visible in 3+ views: 30802

# Notre Dame dataset: Recognition

Notre Dame Testing - GIST



Notre Dame Testing - VocTree



# False positive example



Query



Score: 45



Score: 97

GIST



Score: 9



Score: 8



Score: 8



Score: 8



Score: 45



Score: 10



Score: 10



Score: 8

Vocabulary Tree

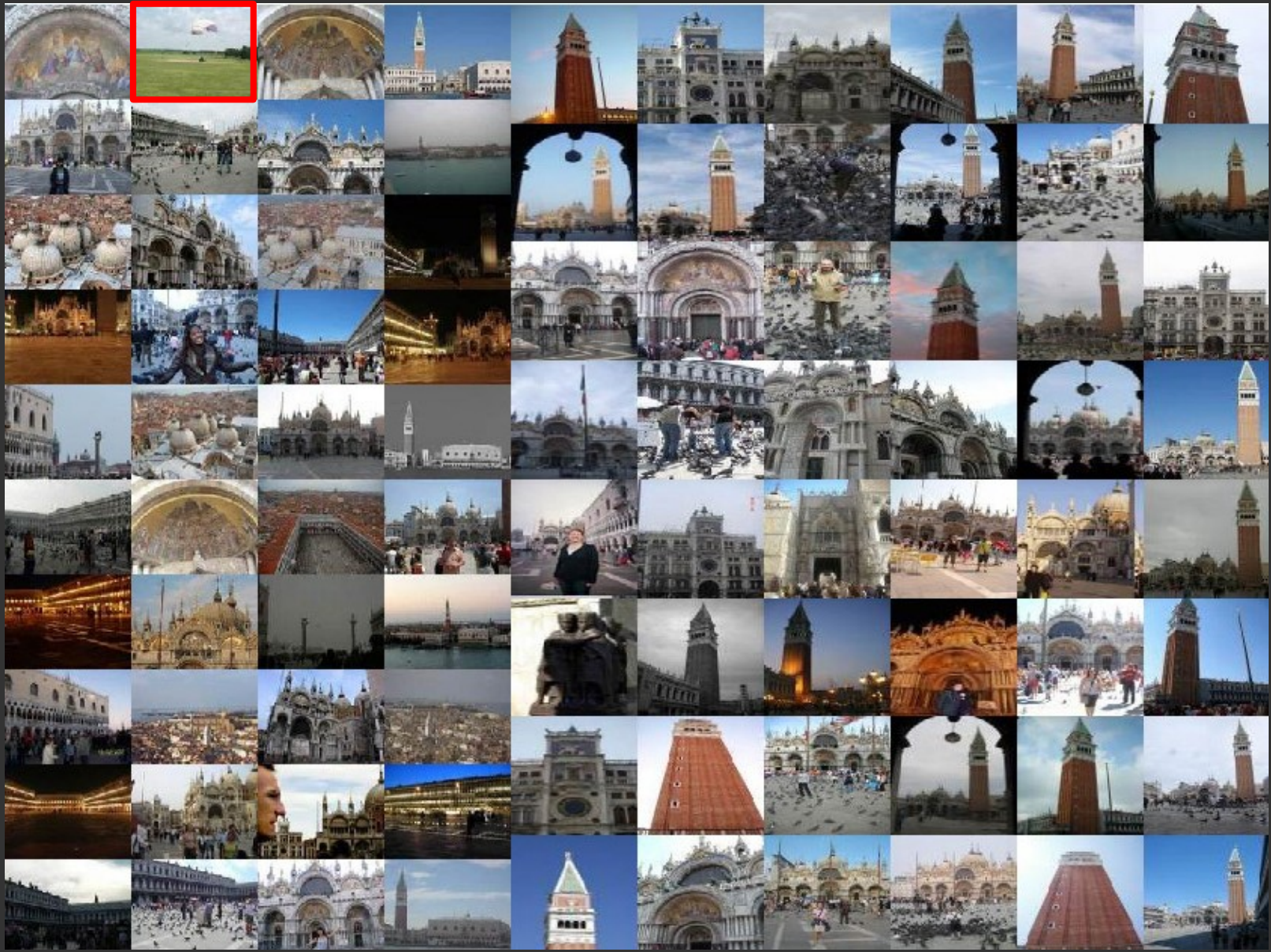
# San Marco dataset

43557 images originally, 198 iconics



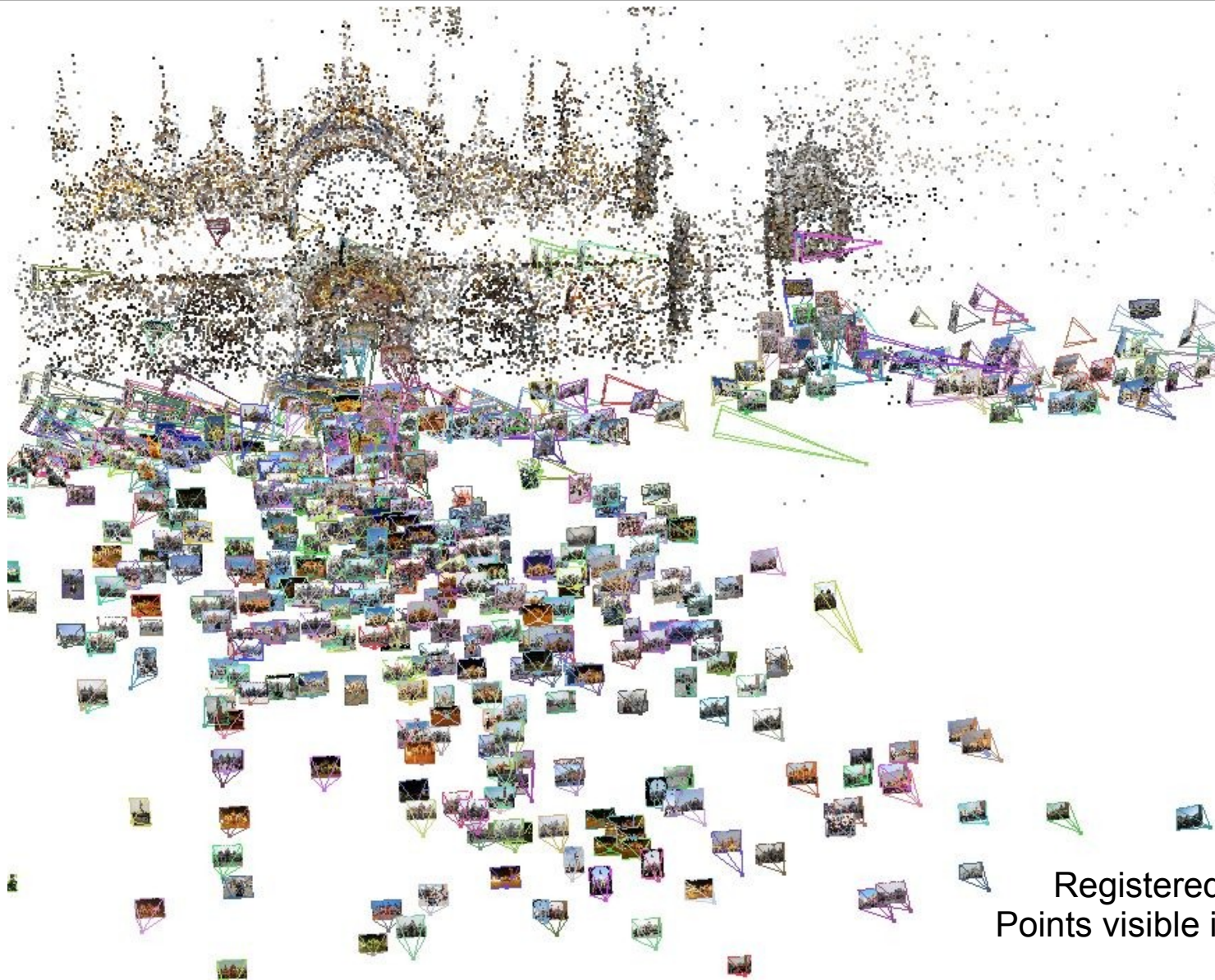
# San Marco dataset

43557 images originally, 198 iconics



# San Marco dataset

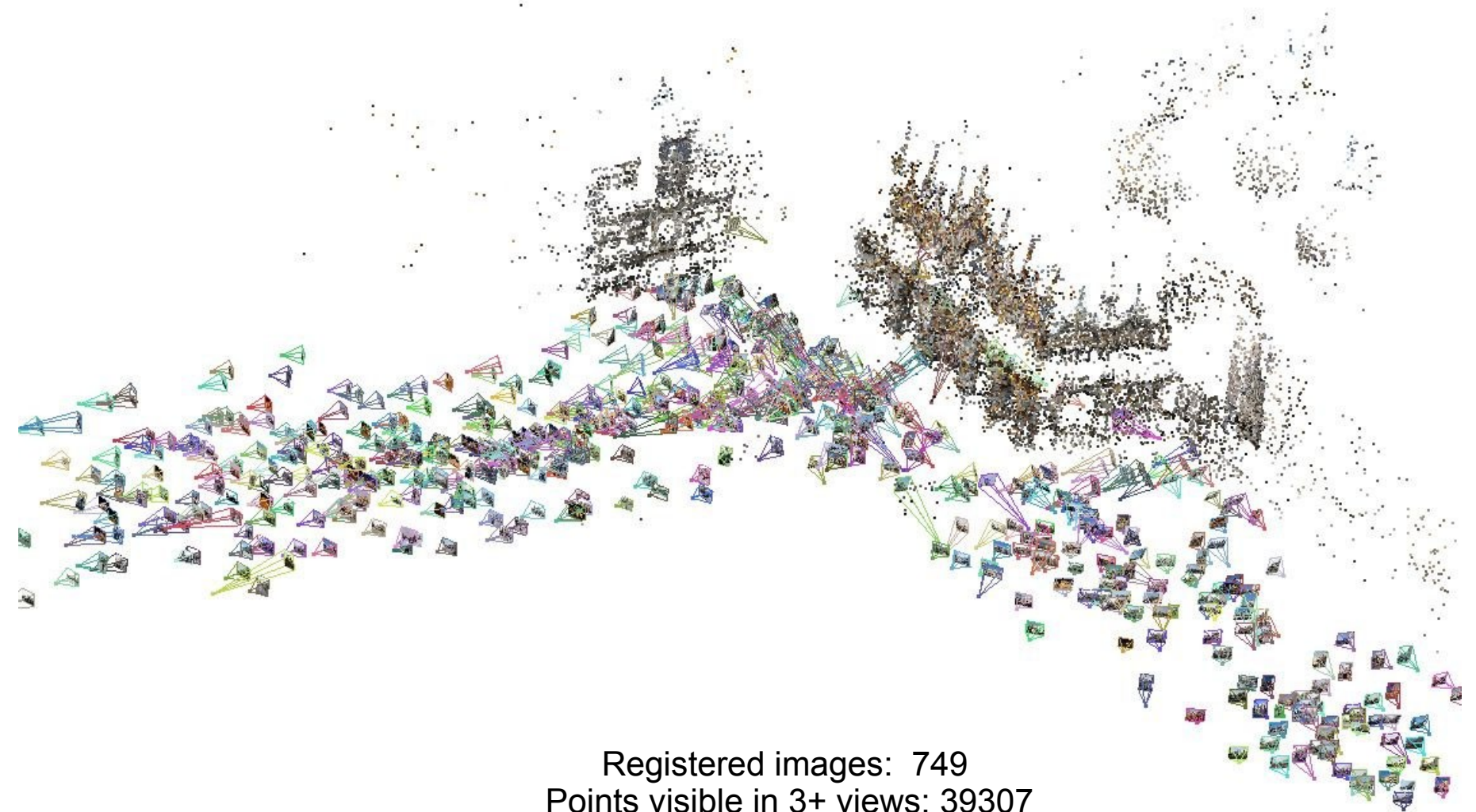
43557 images originally, 198 iconics



Registered images: 749  
Points visible in 3+ views: 39307

# San Marco dataset

43557 images originally, 198 iconics



# San Marco dataset

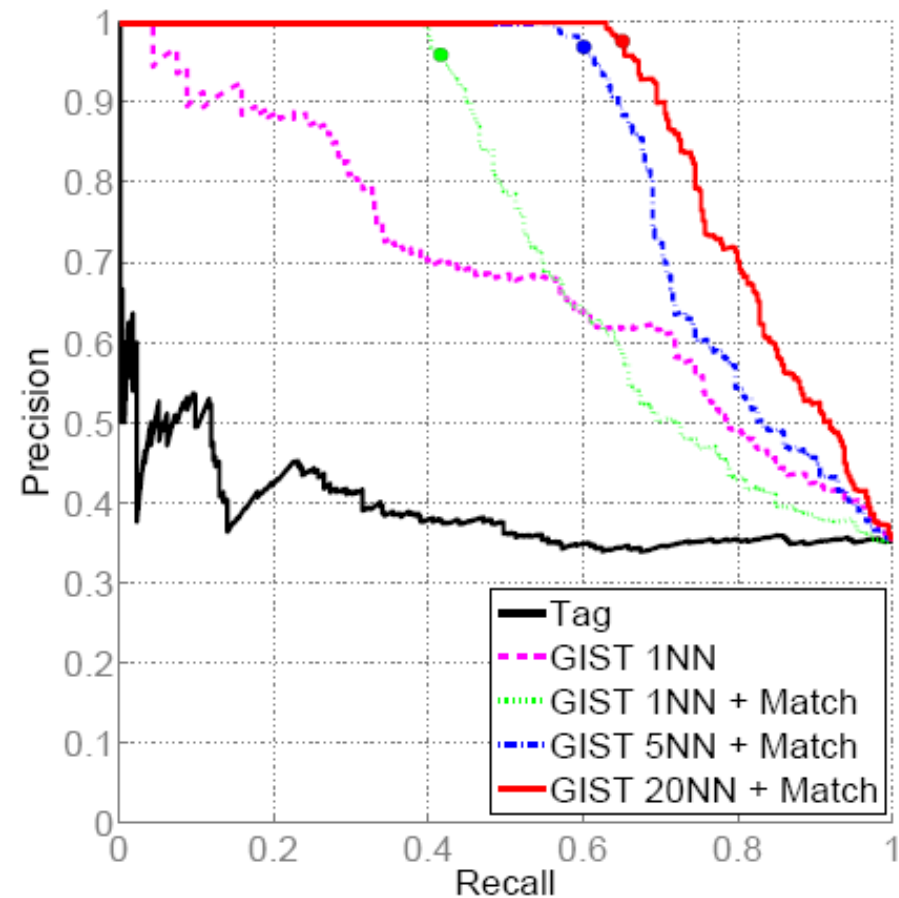
43557 images originally, 198 iconics



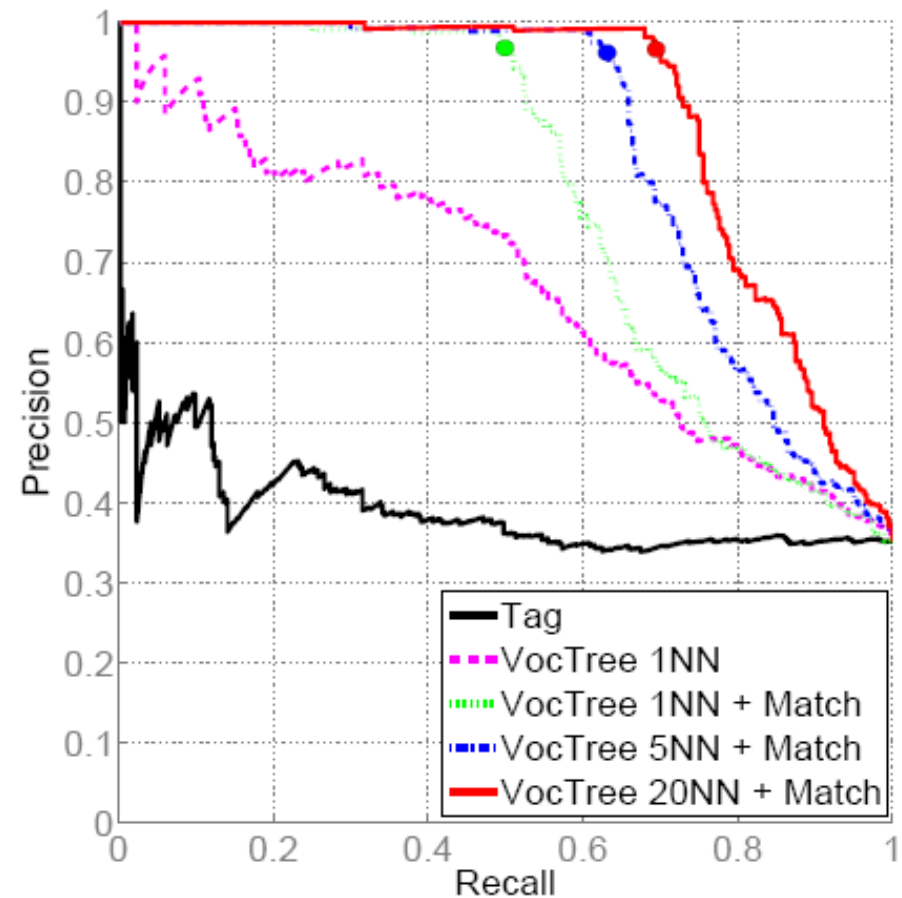
A different component

# San Marco dataset: Recognition

San Marco Testing - GIST



San Marco Testing - VocTree



# Key points

- “Universal” scene representation: useful for summarization, browsing, reconstruction, recognition
- Exploit redundancy of Internet photo collections
- Combine 2D appearance-based descriptors with 3D geometry to capture the relationships between images

# A "hierarchy" of relationships

2D appearance



3D structure



Semantics



# Future work

- Increase recall for modeling and recognition
- Make reconstruction and model merging more robust
- Explore semantic aspects of landmark representation

