Geometry Beyond 3D

Noah Snavely
Google Inc., Cornell University

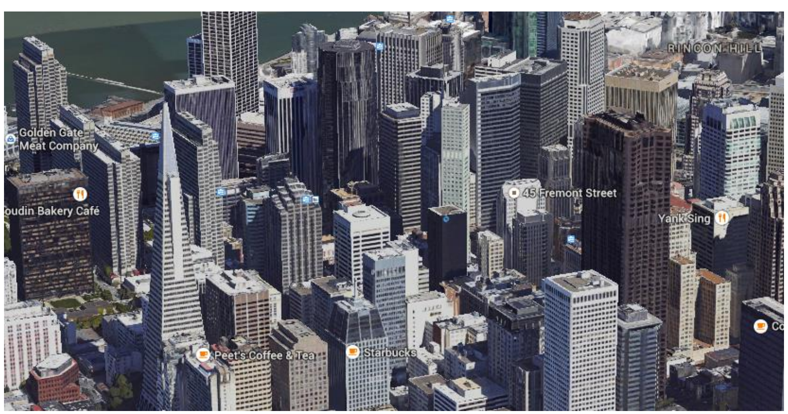
Bay Area Vision Meeting, 2014

Are we done with 3D modeling?

- Huge progress in the last 10 years

[Pollefeys et al. IJCV04]

[Snavely et al. SIGGRAPH06]

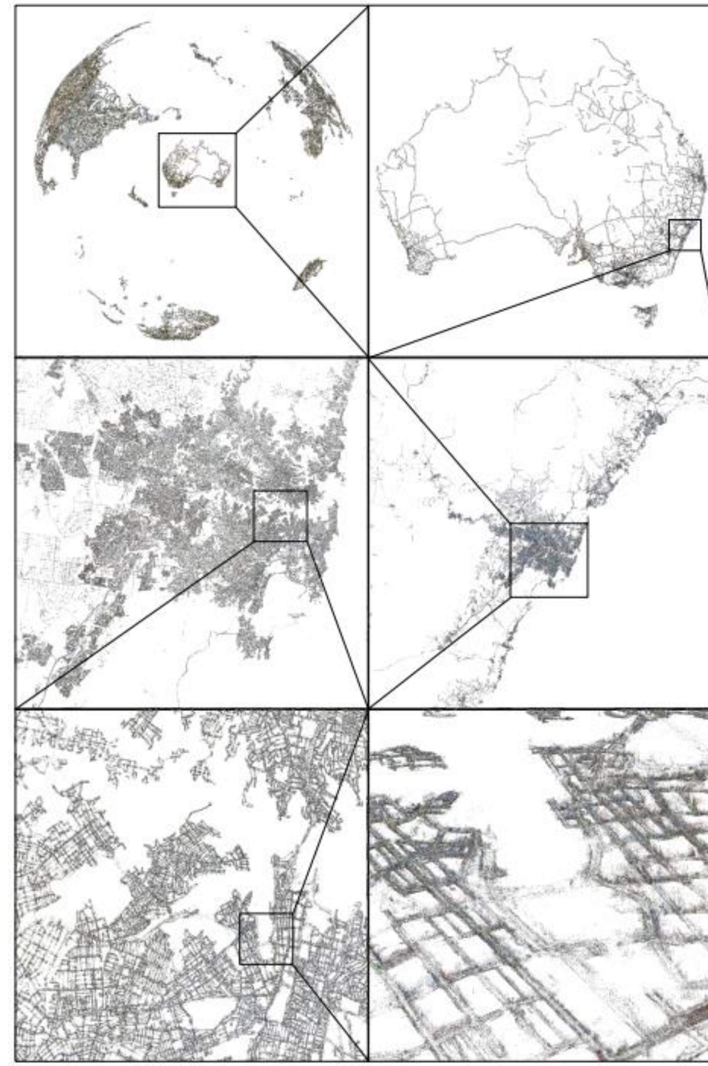


Aerial models

[Zhou & Koltun , SIGGRAPH14]

Are we done with 3D modeling?

[Agarwal et al. ICCV 2009]



[Klingner et al., ICCV 2013]

Are we done with 3D modeling?

- Not until we have a fully **realistic**, **editable**, **semantically meaningful** model of the entire world
- **Realistic** = correct geometry, materials, lighting; high-resolution; dynamic

- In other words, a model you can feed into your holodeck

See also the Visual Turing Test [Shan et al., 3DV 2013]

Times Square

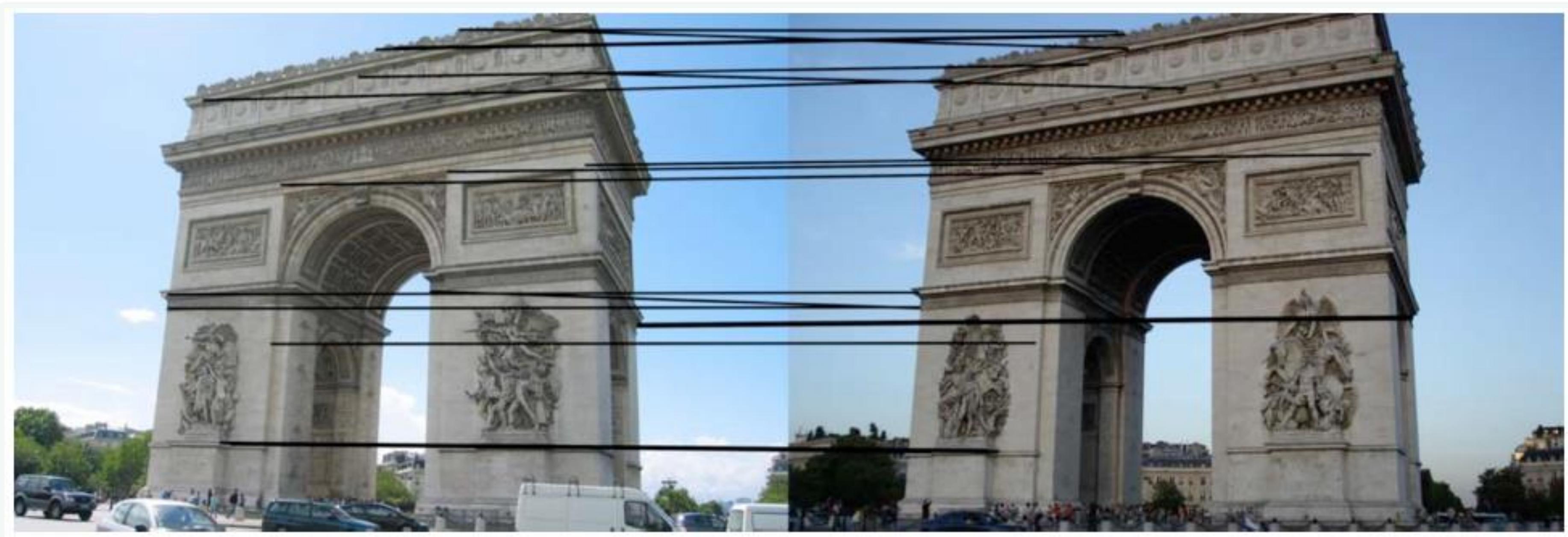
What are the key challenges?

- Scale – we have made great progress here
- **Robustness**
- **Time**
- **Materials**
- **Semantics / grounding**

- My own biased view

Robustness

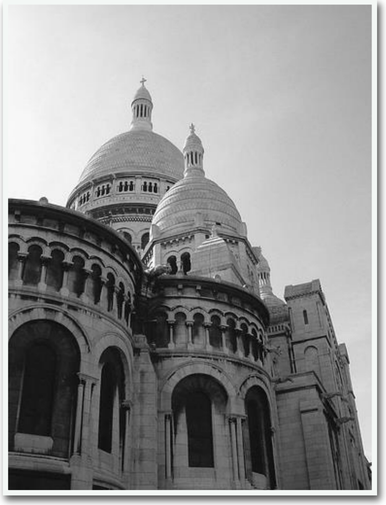
Are two things the same?

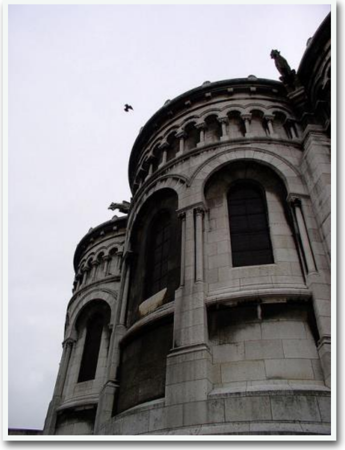


- How do we know what we are looking at is the same or different?

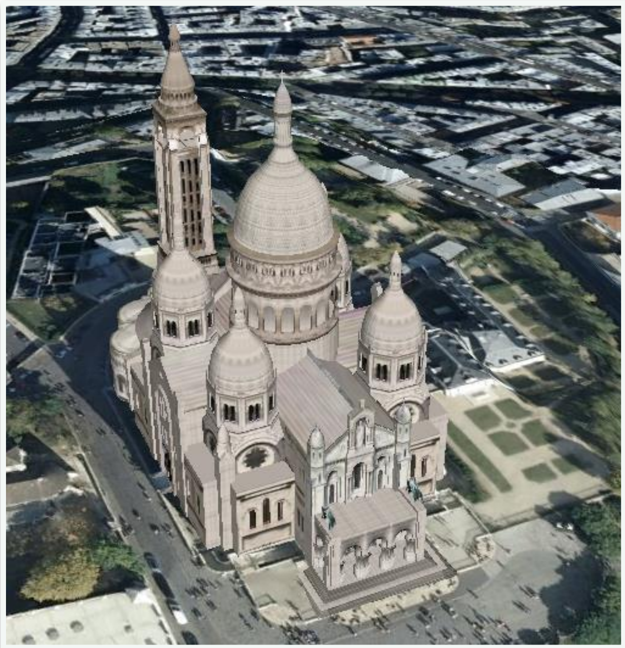
Structural similarities break SfM

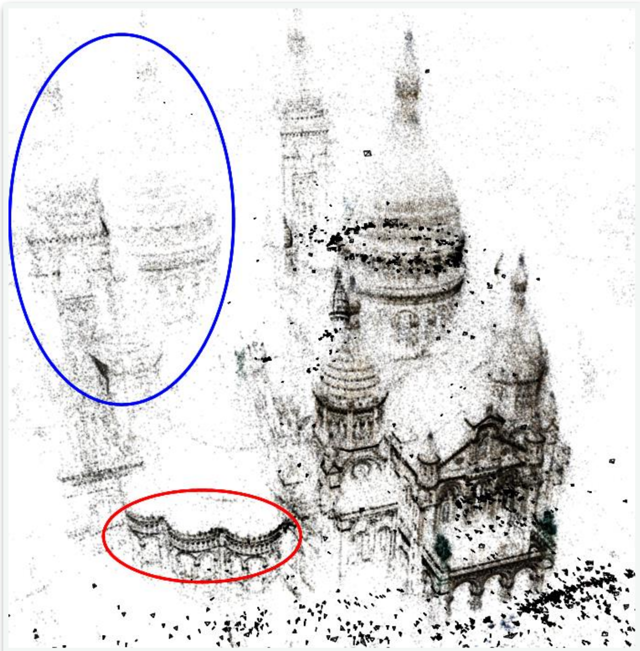


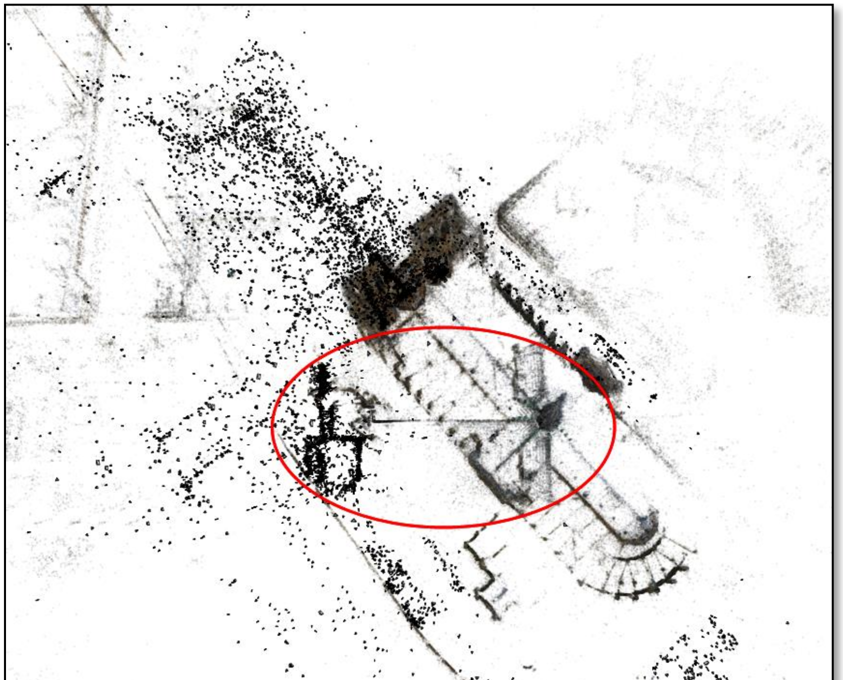


Structural similarities break SfM

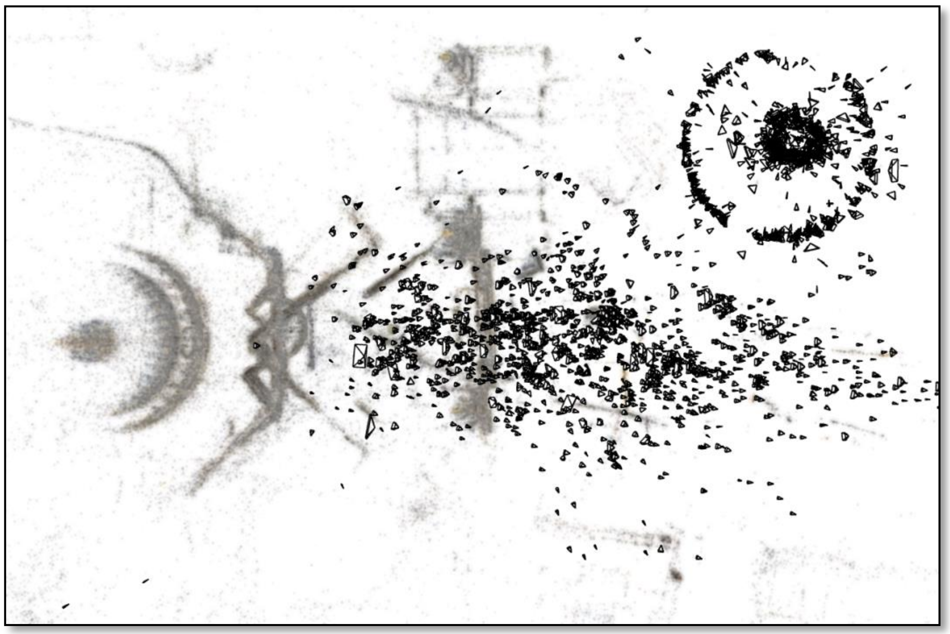




Other examples

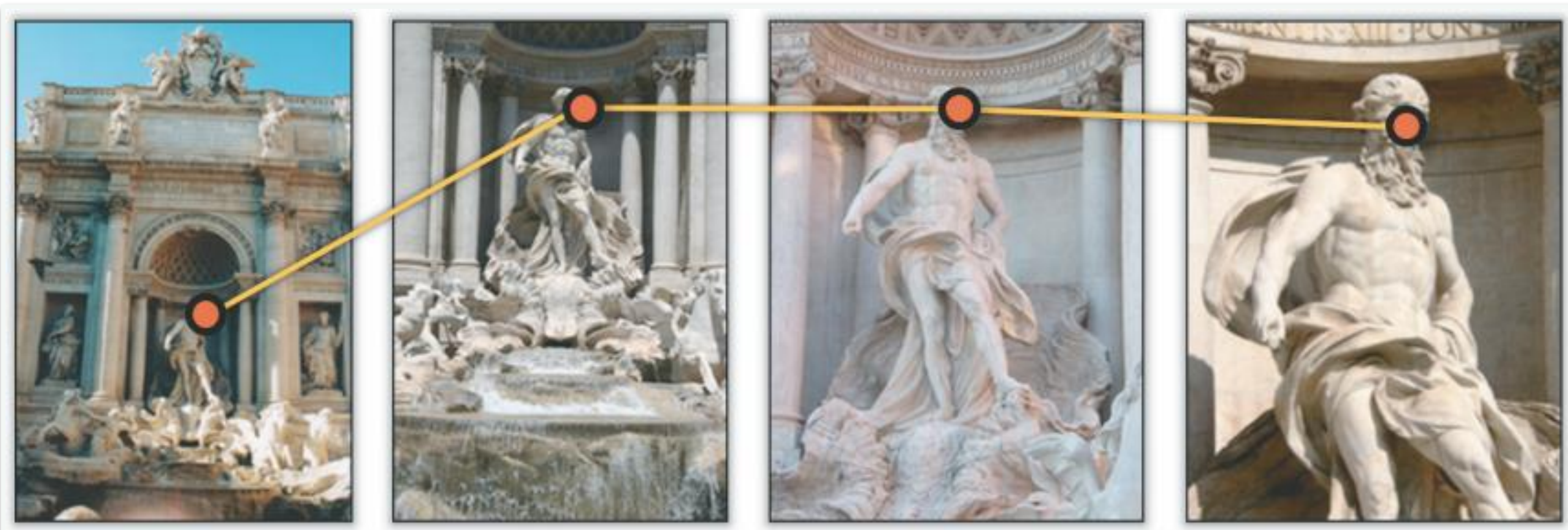


Notre Dame Cathedral

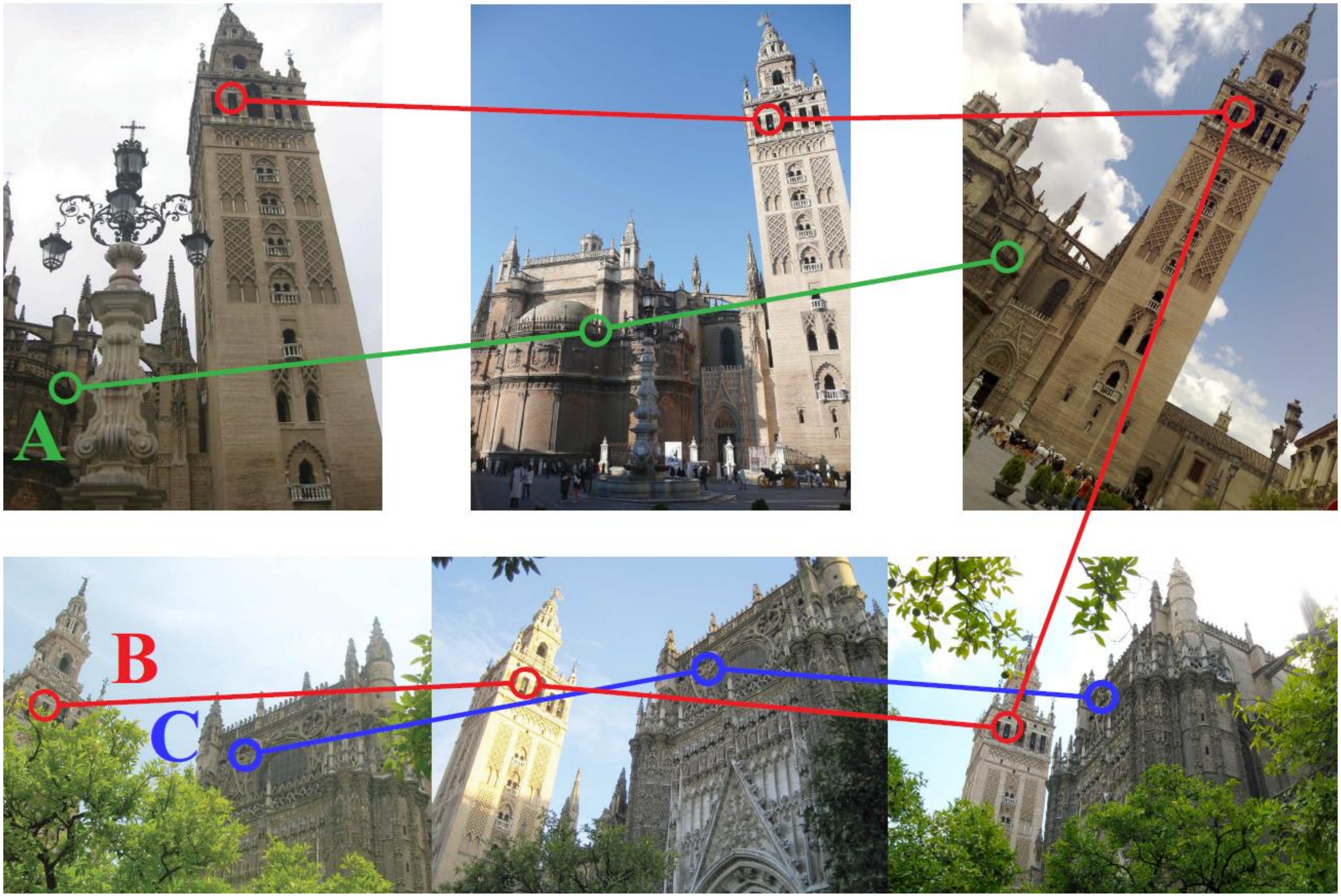


St. Paul's Cathedral

Tracks should contain one 3D point



Tracks can conflate distinct points



SfM Disambiguation

- Most methods reason about inconsistencies across many images
- Inconsistencies in
  - Loops of pairwise geometries
  - Visibility
  - Sequencing
  - Global geometry

[Zach et al., CVPR 2008], [Zach et al., CVPR 2010], [Roberts et al., CVPR 2011], [Jiang et al., CVPR 2012]

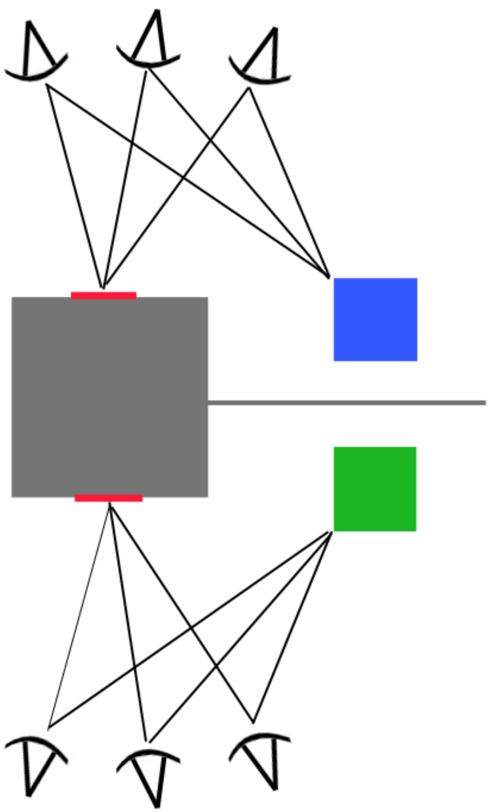
SfM Disambiguation in the Large

- We wanted a solution that was
  - As simple as possible
  - Scalable to huge image collections

- Intuition: visibility of points is (often) transitive

[Wilson & Snavely, Network Principles for SfM. ICCV 2013]

Graph topology is a cue for ambiguities

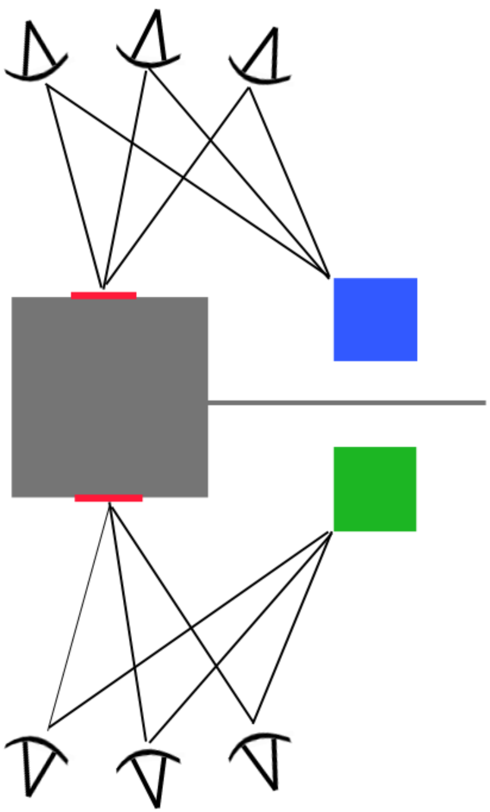


Schematic of a scene with an ambiguous feature (in red)

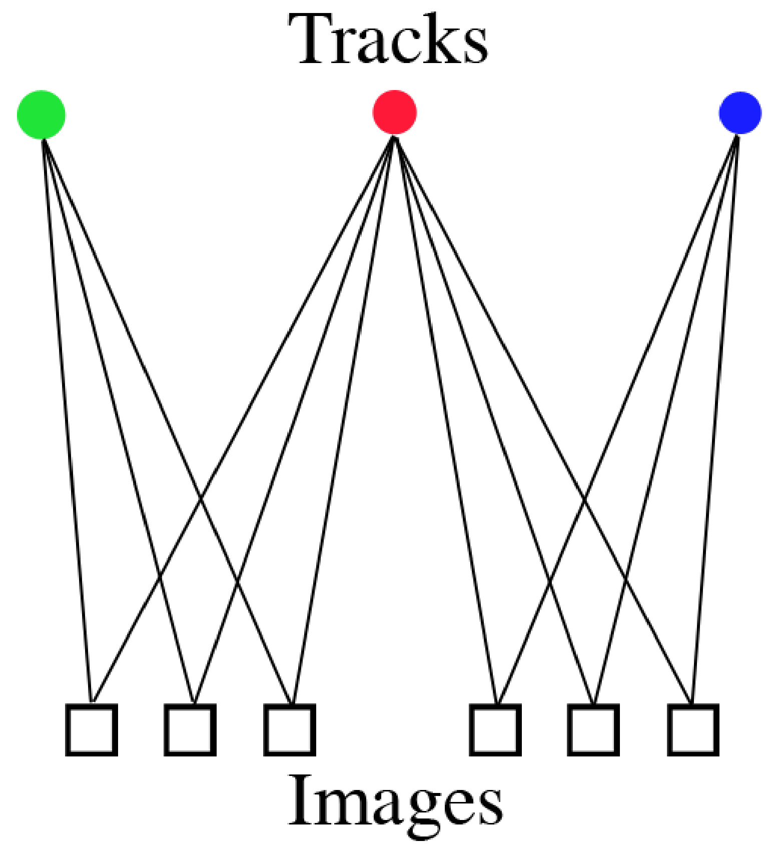
Note that the two sides of the scene have different background (blue and green)

[Wilson & Snavely, Network Principles for SfM. ICCV 2013]

Graph topology is a cue for ambiguities

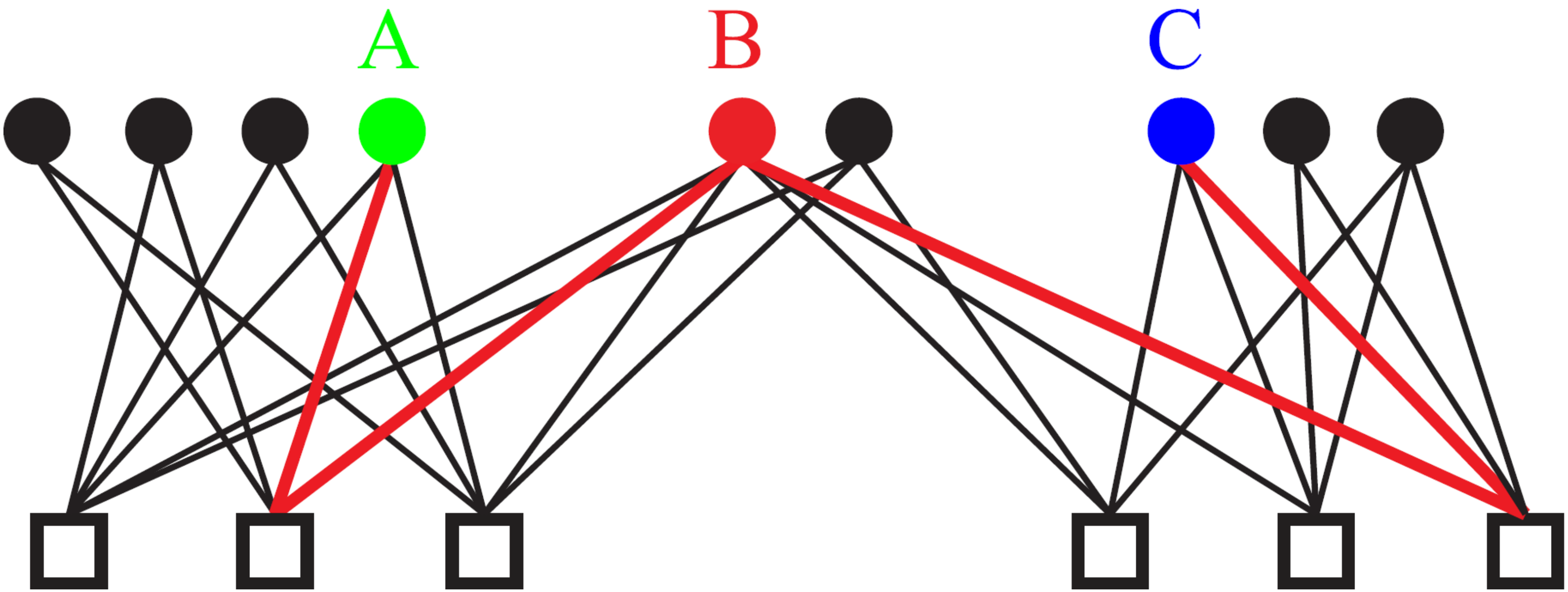


This structure can be seen in the *visibility graph*



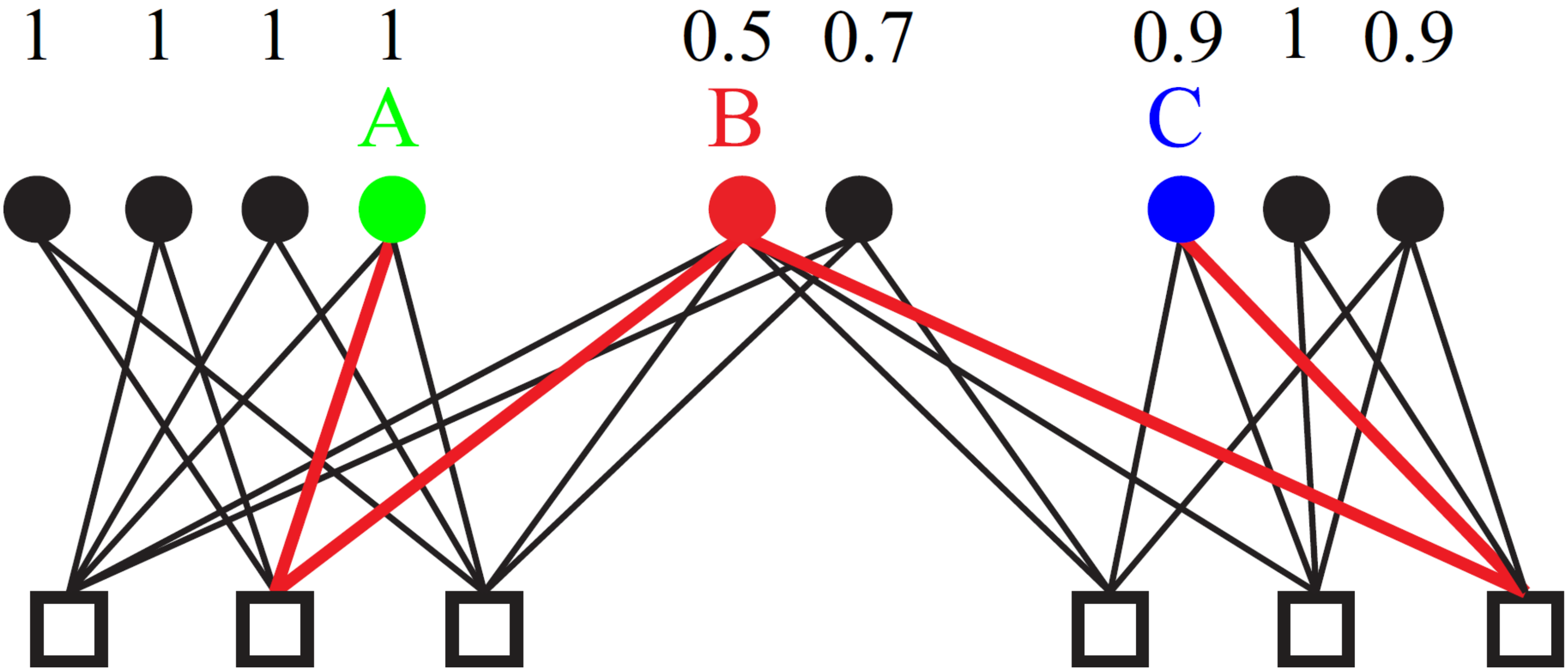
[Wilson & Snavely, Network Principles for SfM. ICCV 2013]

Larger example



Bad tracks have more than one cluster of context. Measure this with the <u>bipartite local clustering coefficient</u>.

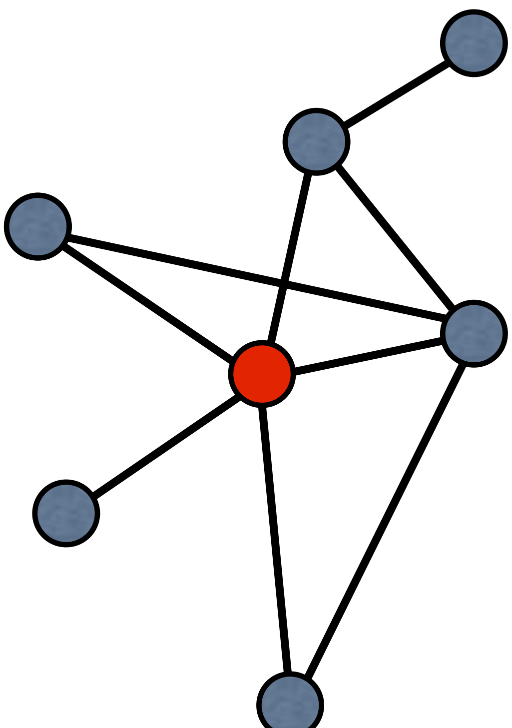
Larger example

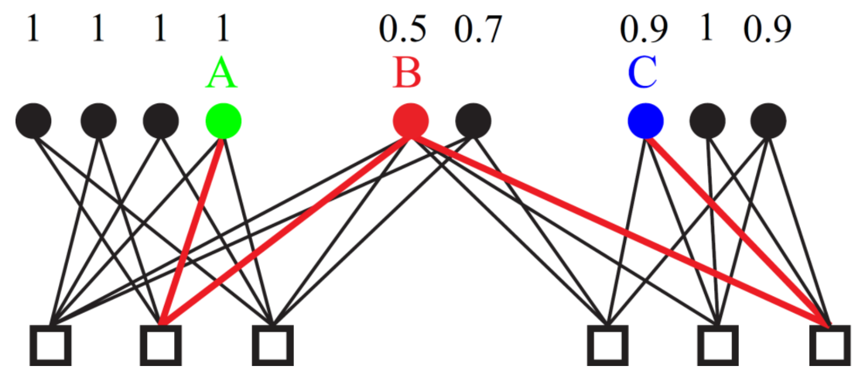


Bad tracks have more than one cluster of context. Measure this with the <u>bipartite local clustering coefficient</u>.

blcc is analagous to the local clustering coefficient

$$lcc(\text{red node}) = \frac{\text{closed triangles}}{\text{possible triangles}} = \frac{3}{10}$$

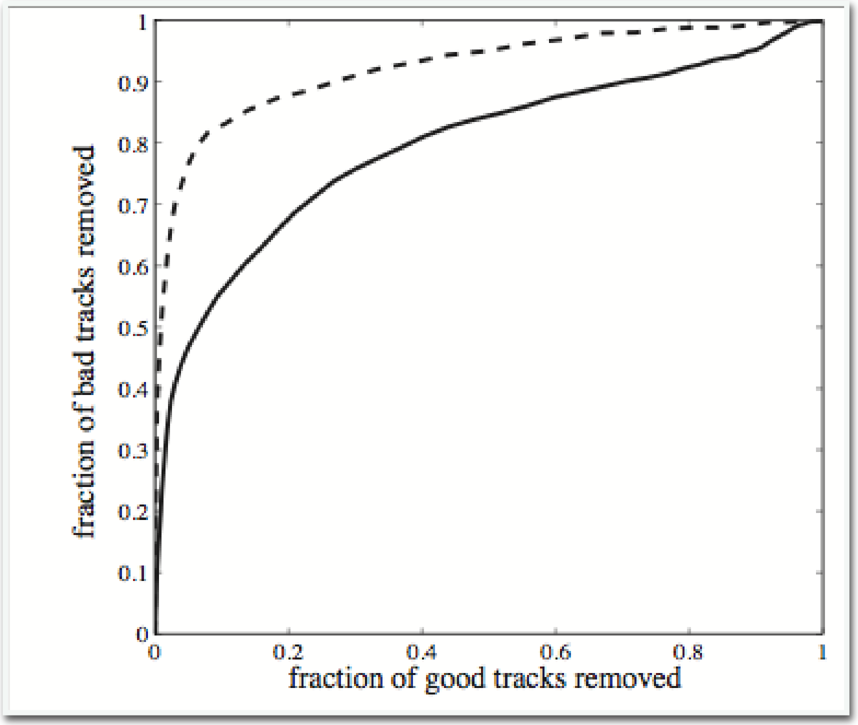




$$blcc(\text{red node}) = \frac{\text{closed 6-paths}}{\text{possible 6-paths}}$$

Filtering by *blcc* removes bad tracks

ROC curve for classifying bad tracks

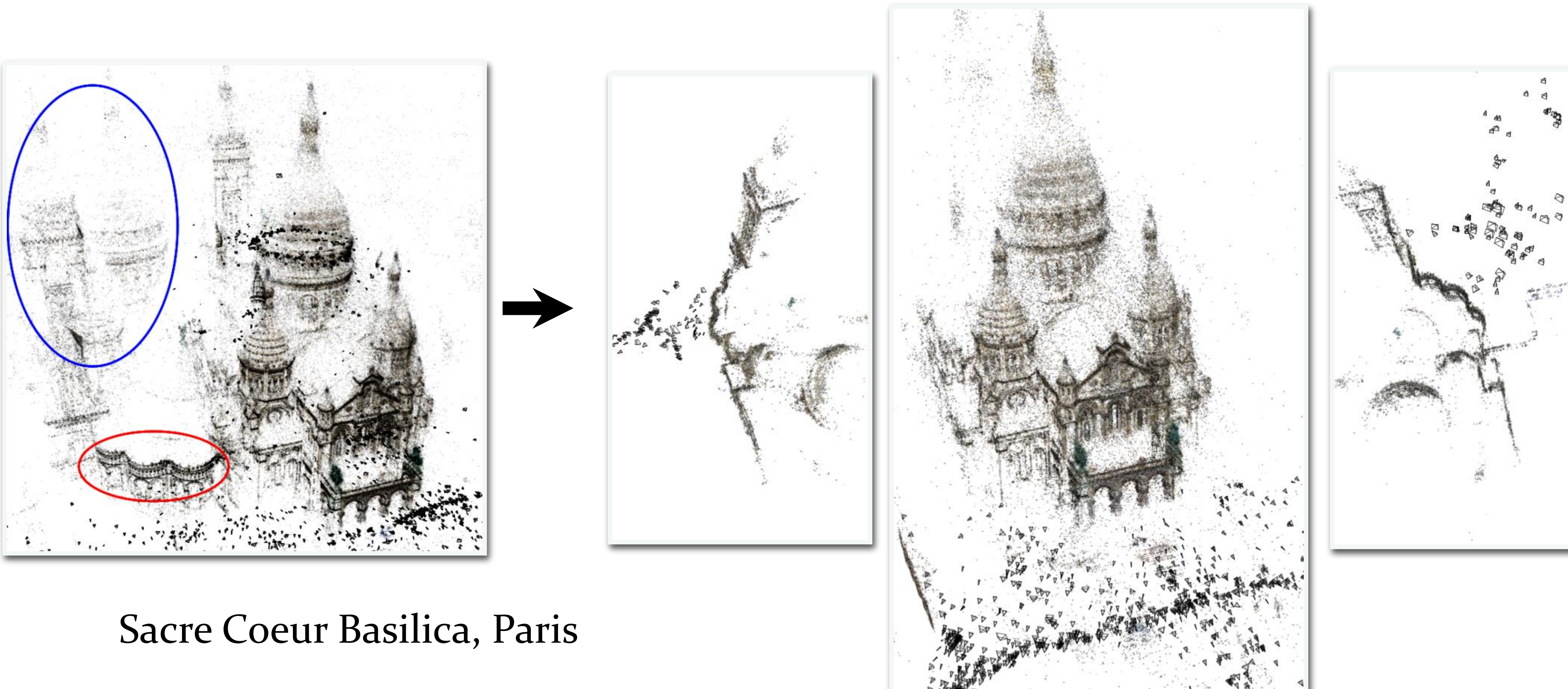


Solid line: thresholding tracks on blcc.
Dotted line: same, but on a more uniform subgraph.

Algorithm:

1. Compute a covering subgraph
2. Compute blcc for each track
3. Remove tracks lower than a threshold
   Use lowest threshold that separates the graph into a user-predetermined number of components.
4. Reconstruct each component separately
5. Rigidly merge components if possible

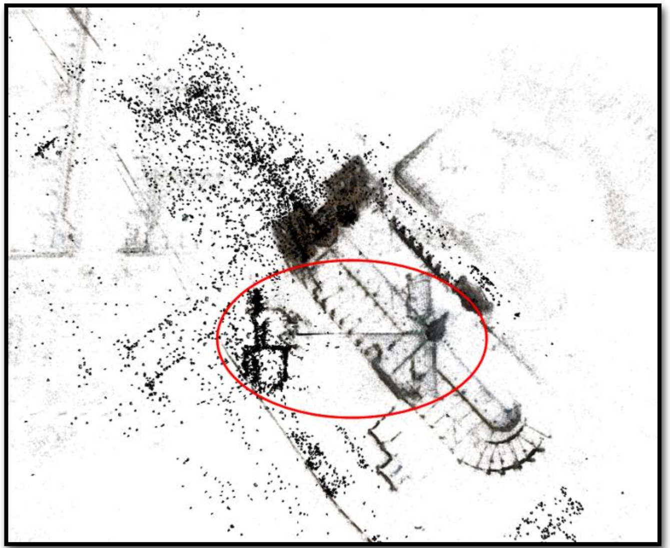
Disambiguation results



Sacre Coeur Basilica, Paris

Disambiguation results

Before

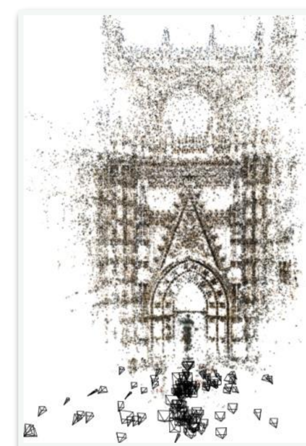


After

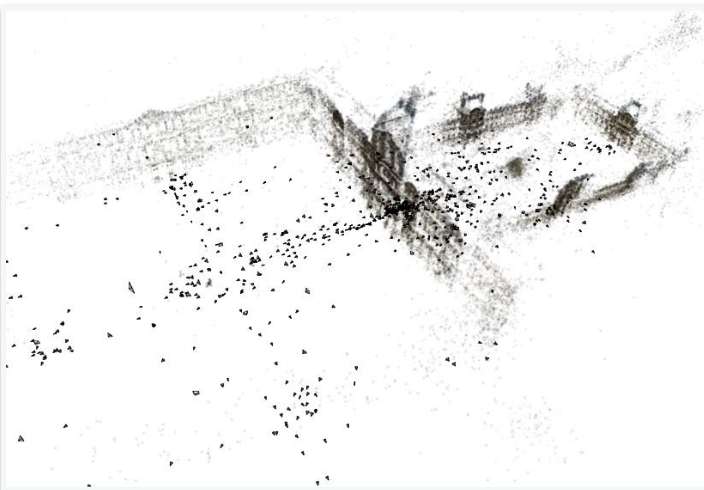
Notre Dame Cathedral, Paris

Disambiguation results

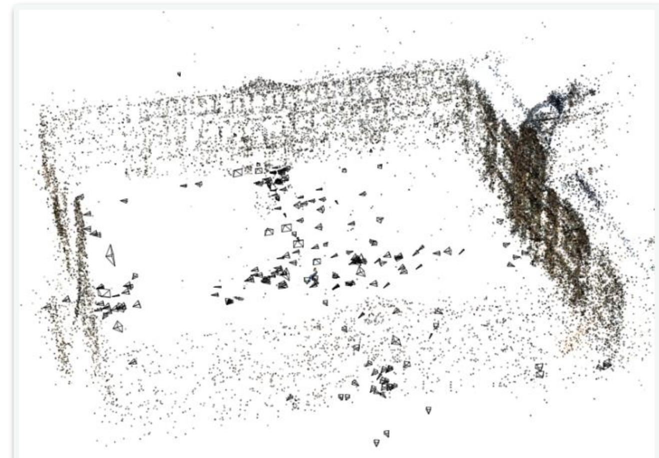
Seville Cathedral



Disambiguation results



Outside the Louvre, Paris



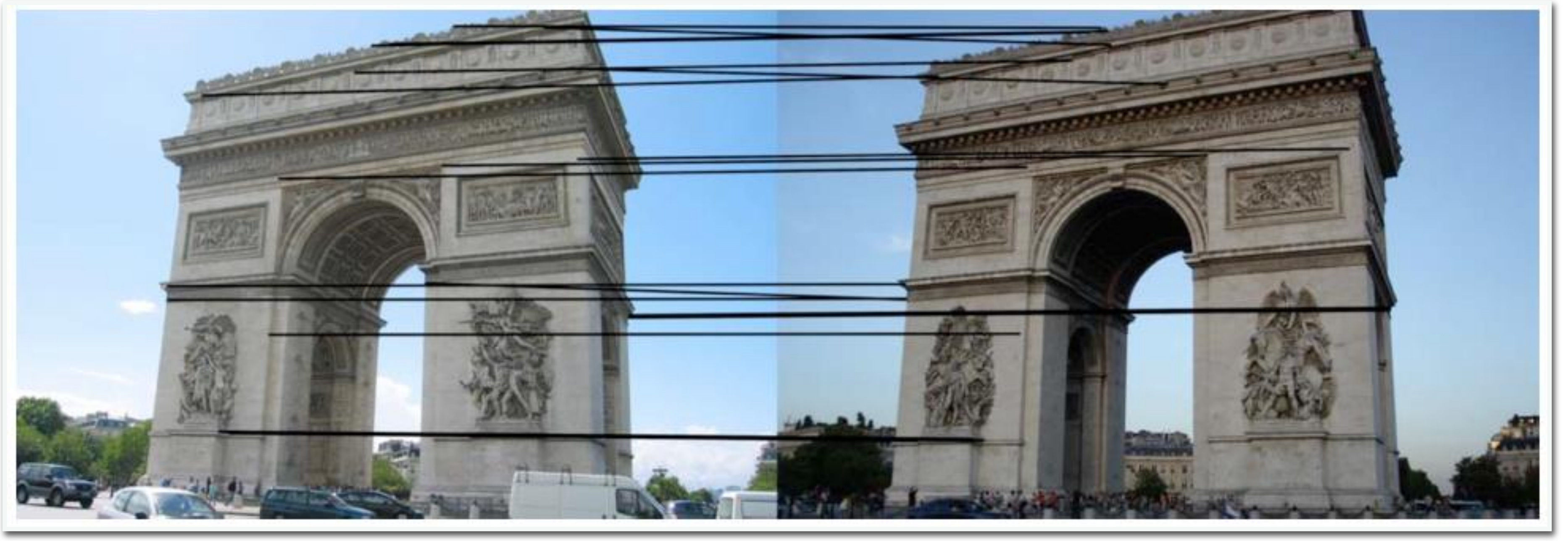
Network Principles for SfM

+ Extremely fast method
+ Based on simple local reasoning
+ Very simple to implement

- Can sometimes oversegment models
- Theoretical guarantees?

See also [Heinly et al. ECCV 2014]

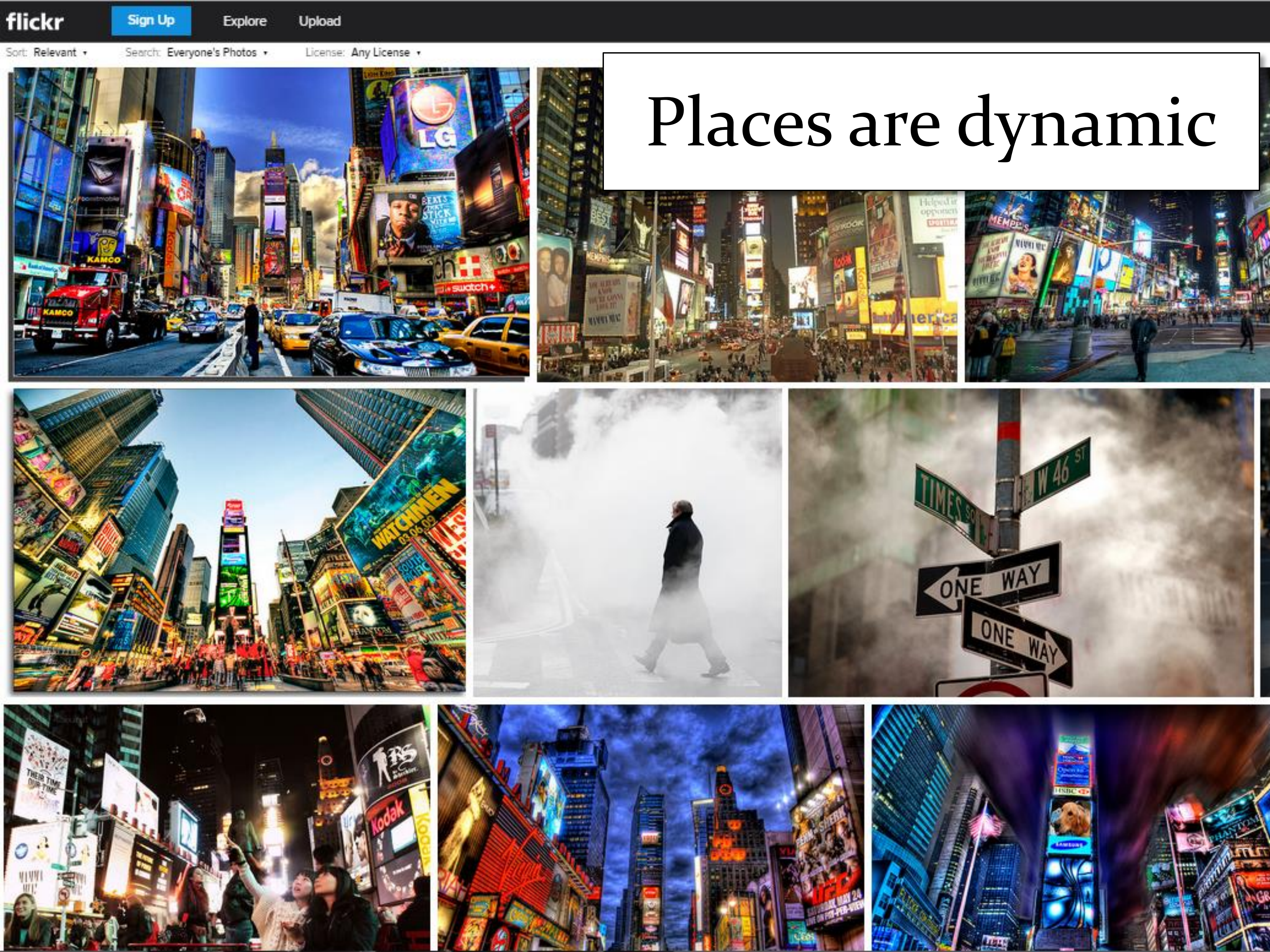
Feature matching as recognition



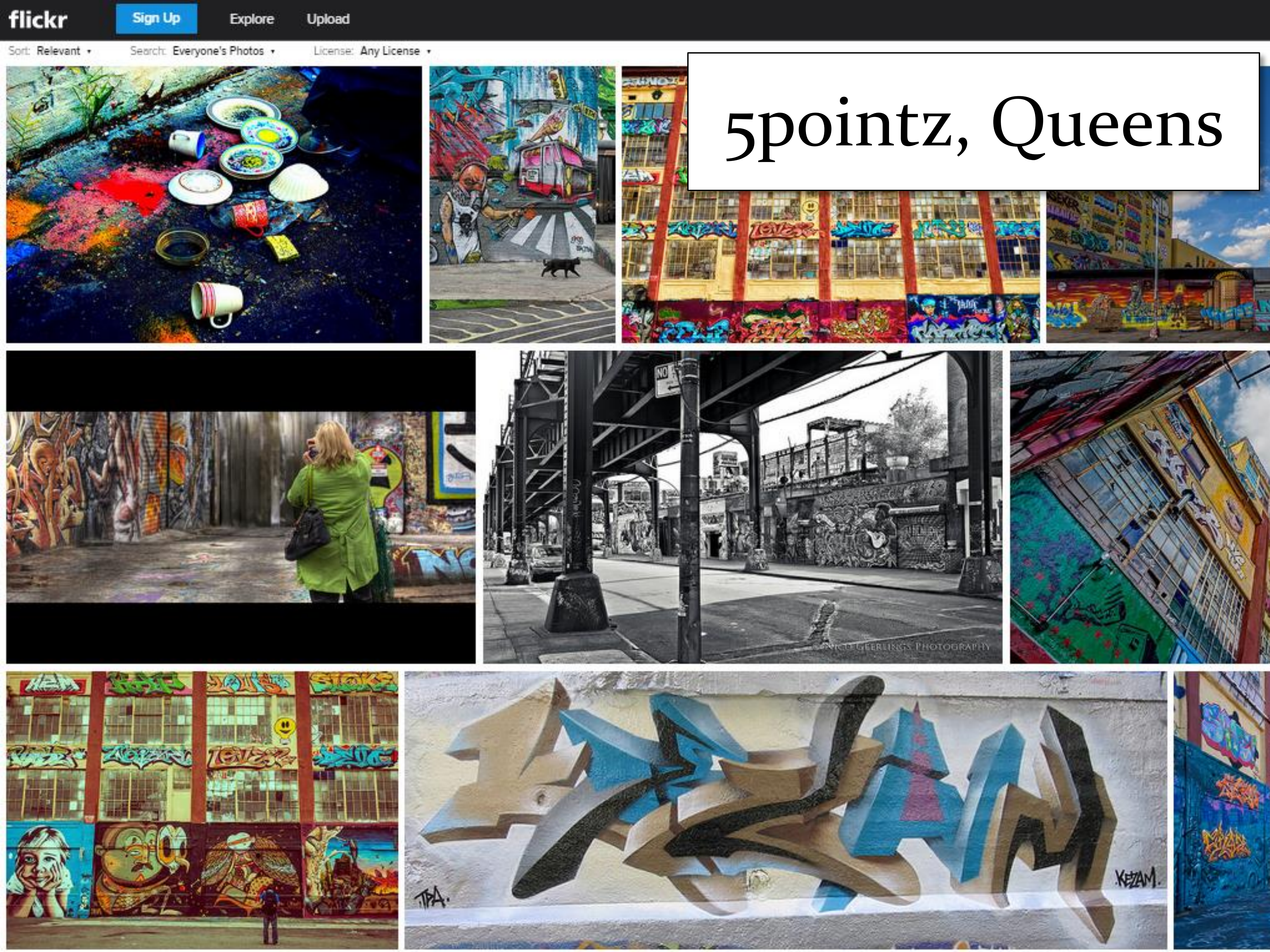
- Can’t we just solve this problem using appearance alone?
- Better features or image metrics?

Time

Places are dynamic



5pointz, Queens

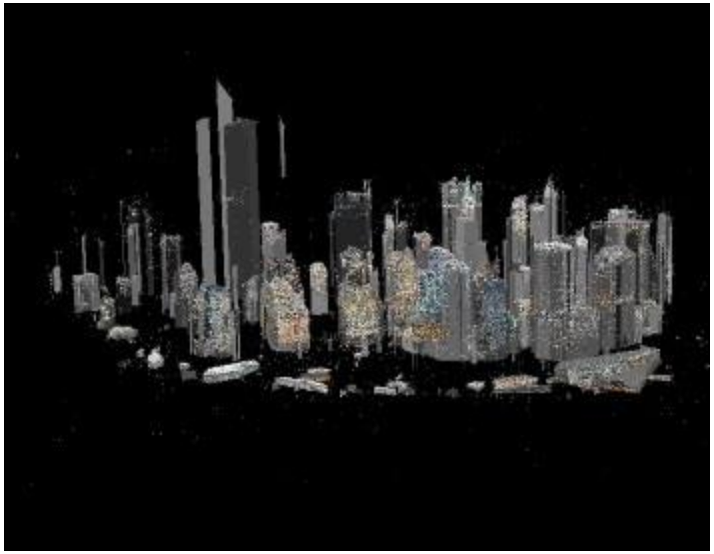


5pointz

How do we model these time-varying scenes?

[Graffiti Archaeology, Cassidy Curtis]

4D Cities



[Frank Dellaert, Grant Schindler, et al.]

Scene Chronology

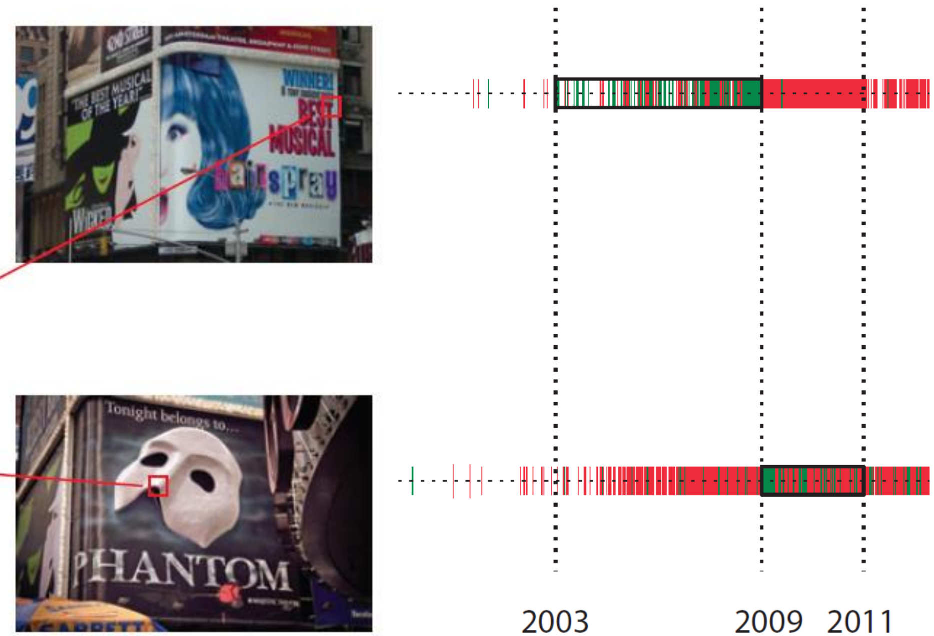
Step 1: Download photos from Flickr

Step 2: Reconstruct a single 3D model with all times mixed up together

Step 3: Recover the *chronology* of the scene

Kevin Matzen and Noah Snavely, **Scene Chronology**, ECCV 2014
Best Paper Award Winner

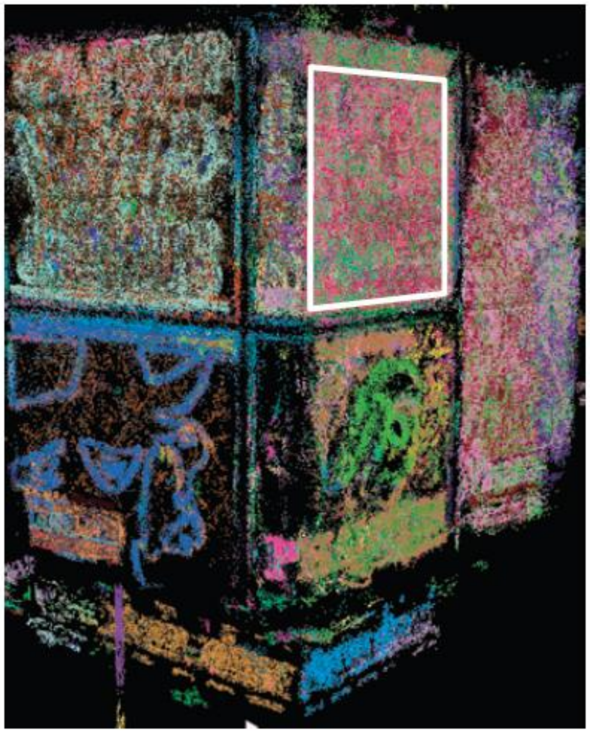
Single 3D Model (from ~100,000 images)



Per-Point Time Observations

CATHERINE ZETA-JONES
ANGELA LANSBURY
A LITTLE NIGHT MUSIC
SEX AND THE CITY
CARRIE ON MAY 27
SEXANDTHECITYMOVIE.COM
HBO
BANANA SHPEEL
CIRQUE DU SOLEIL
LIMITED ENGAGEMENT
COME HOME TO YANKEE STADIUM.
THE CROWD GOES WILD
JERSEY BOYS
BEST MUSICAL
LOVE OUT
HAIR
BEST MUSICAL REVIVAL · 2009 TONY AWARD
Tonight belongs to...
PHANTOM
WINNER! 7 TONY AWARDS
BEST MUSICAL REVIVAL
RODGERS & HAMMERSTEIN'S
SOUTH PACIFIC
LINCOLN CENTER
PALACE
WEST SIDE STORY
TICKETMASTER.COM / 212-307-4100
CHICAGO
CALIENTE · INCONTOURNABLE
THE INTERNATIONAL
TELECHARGE.COM/CHICAGO
CHICAGOTHEMUSICAL.COM
FUEL
THE CHAN
WEST SIDE STORY
GUEST SUITES
"IT WILL BLOW YOU AWAY."
"SO EXCITING IT MAKES YOU ACHE WITH PLEASURE."
MONDAY TO SATURDAY 10 A.M. - 8 P.M....
DoubleTree Guest Suites
sbarro
GIFTS

Catherine Zeta-Jones
Angela Lansbury
A Little Night Music
SEX AND THE
CARRIE ON MAY 27
SEXANDTHECITYMOVIE.COM
HBO
CIRQUE DU SOLEIL
BANANA SHPEEL
LIMITED ENGAGEMENT
COME HOME TO YANKEE STADIUM
JERSEY BOYS
BEST MUSICAL
HAIR
BEST MUSICAL REVIVAL · 2009 TONY AWARD
Tonight belongs to...
PHANTOM
WICKED
QUARTET
SOUTH PACIFIC
WINNER! 7 TONY AWARDS
BEST MUSICAL REVIVAL
RODGERS & HAMMERSTEIN'S
LINCOLN CENTER
PALACE
WEST SIDE STORY
TICKETMASTER.COM / 212-307-4100
CHICAGO
CALIENTE · INCONTOURNABLE · 熱的
THE INTERNATIONAL
TELECHARGE.COM/CHICAGO
CHICAGOTHEMUSICAL.COM
FUEL
THE CHAN
GUEST SUITES
"SO EXCITING IT MAKES YOU ACHE WITH PLEASURE"
"IT WILL BLOW YOU AWAY."
MONDAY TO SATURDAY 10 A.M. - 8 P.M....
DoubleTree Guest Suites
sbarro
GIFTS

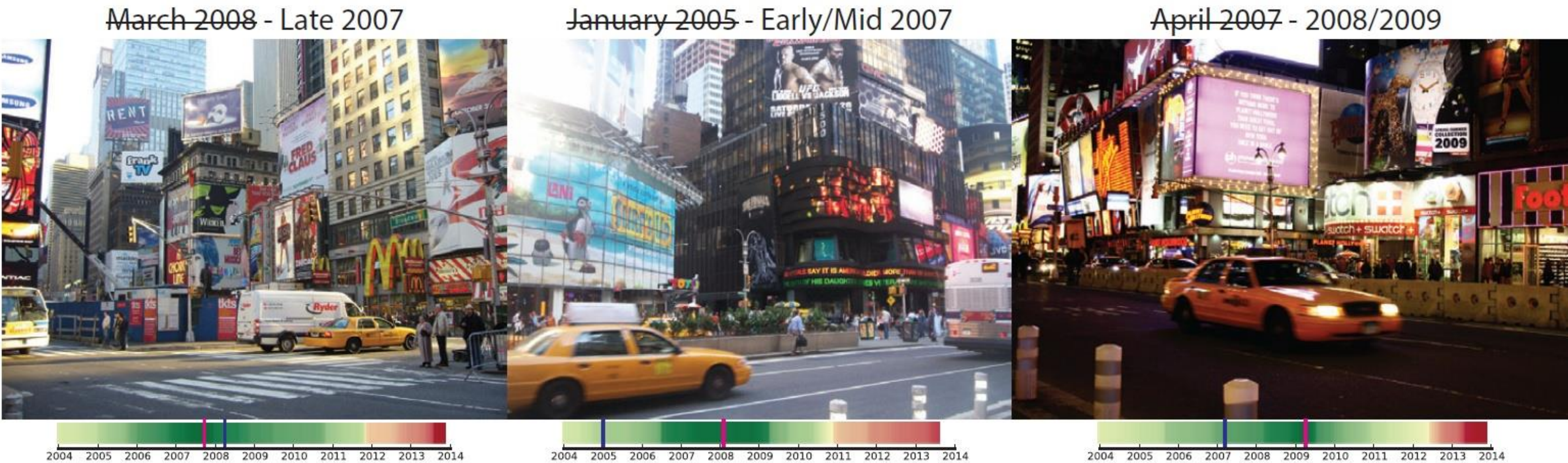


Space-Time Point Clustering

Exploded View across Time

5Pointz - Queens, New York City

Re-time-stamping



Blue: original timestamp
Red: our predicted timestamp

Physics

Weather

People

Times Square, 1922

Eisenstadt, 1945

Materials


Ceiling
Wall
Sofa
Leather
Plastic
Chair
Fabric
Mat
Fur
Wood
Floor
living room

OpenSurfaces

Sean Bell, Paul Upchurch, Noah Snavely, Kavita Bala
Cornell University

Statistics

	Good	All
Labeled Scenes	25,357	91,876
Whitebalanced Photos	17,839	24,771
Segmentations	70,005	103,513
Planar Segmentations	36,482	70,005
Named Materials	56,625	68,761
Named Objects	31,697	42,203
Rectified Textures	16,829	22,219

Materials

Reflectances

Textures

Sean Bell, Paul Upchurch, Noah Snavely, Kavita Bala, SIGGRAPH 2013
http://opensurfaces.cs.cornell.edu/

Query

Results: wood floors in kitchens, sorted by diffuse color

Query

Results: fabric sofas in living rooms, sorted by diffuse color similarity

Sean Bell, Kavita Bala, Noah Snavely, SIGGRAPH 2014, http://intrinsic.cs.cornell.edu

Semantics / Grounding

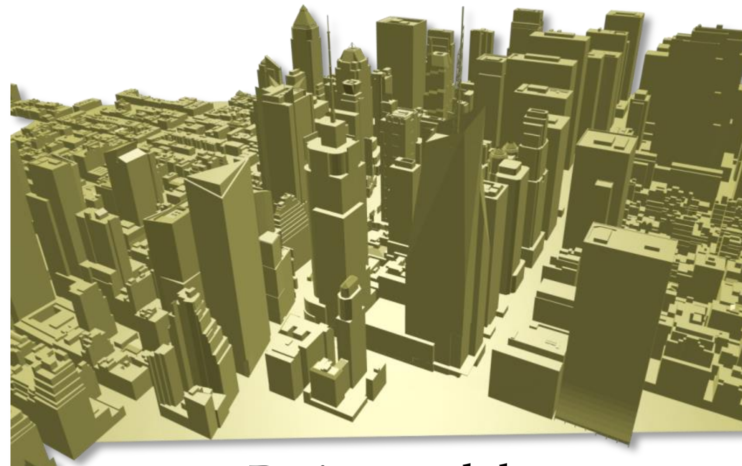
Every image tells a story…

José Luis Murillo


Vivienne Gucwa

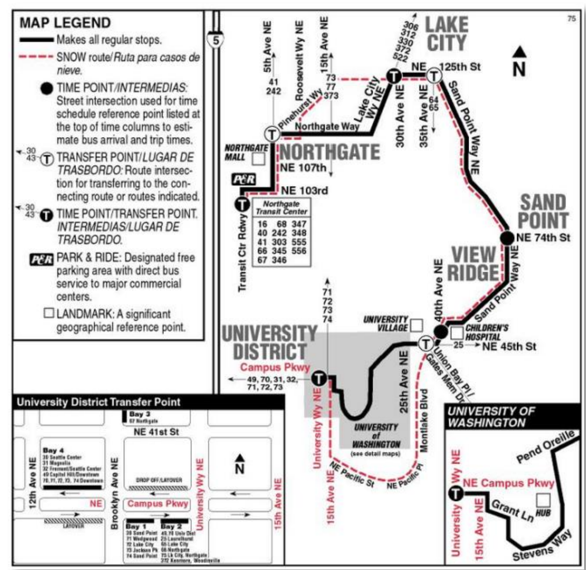
Grounding vision in the world

OpenStreetMap



3D city models

Weather data



Bus schedules

NYC OpenData

Search & Browse Datasets and Views

Name

1. **311 Service Requests from 2010 to Present** Social Services 311, 311 service requests, 2010, 2011, 2012, ...
All 311 Service Requests from 2010 to present. This information is automatically updated daily.
2. **Electric Consumption by ZIP Code - 2010** Environment electricity, energy, environment, planning, power, ...
2010 electricity consumption in kWh and GJ, by ZIP code, building type, and utility company.
3. **Zip Codes Map** Social Services geographic, location, map, cartography, zip, code, ...
Polygon representing the boundary of the zip codes in the city.
4. **MTA Data** Transportation traffic, vehicles, route, schedules, clean web
Information pertaining to MTA (Metropolitan Transportation Authority of the State of New York) subways, buses, commuter rail,
5. **Restaurant Inspection Results** Health restaurant inspection results, ...
NYC restaurant inspection results
6. **Basic Description of Colleges and Universities** Education doitt gis, geographic, location, map, cartography, ...
Location of colleges and universities with basic descriptive information.
7. **SAT (College Board) 2010 School Level Results** Education lifelong learning
New York City school level College Board SAT results for the graduating seniors of 2010. Records contain 2010 College-bound
8. **Mapped In NY Companies** Business jobs, tech, jobs and economic mobility
Raw data which powers the Mapped In NY site at http://www.mappedinny.com/
9. **Filming Locations (Scenes from the City)** Business film, movie, scene, scenes from the city
List of filming locations mentioned in the book Scenes from the City
10. **2012 NYC Noise Complaints - Heat Map** Other 311, 311 service requests, 2010, 2011, ...
NYC Noise complaints for 2012

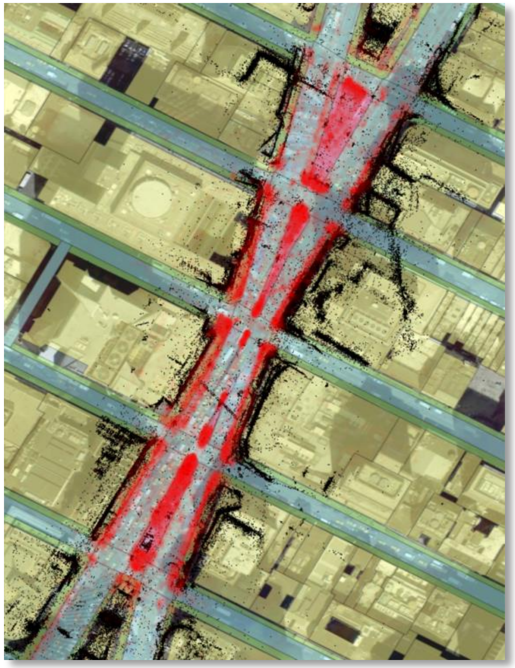
https://nycopendata.socrata.com (https://data.sfgov.org/, https://data.seattle.gov/, ...)

Grounding vision in the world

- Which direction is north?
- What is the shape of the buildings?
- What was the weather like?
- Where are streets?
- What is the #51 bus schedule in Rome?

Goal: Integrate images into this ecosystem of geographic data

First steps: NYC3DCars

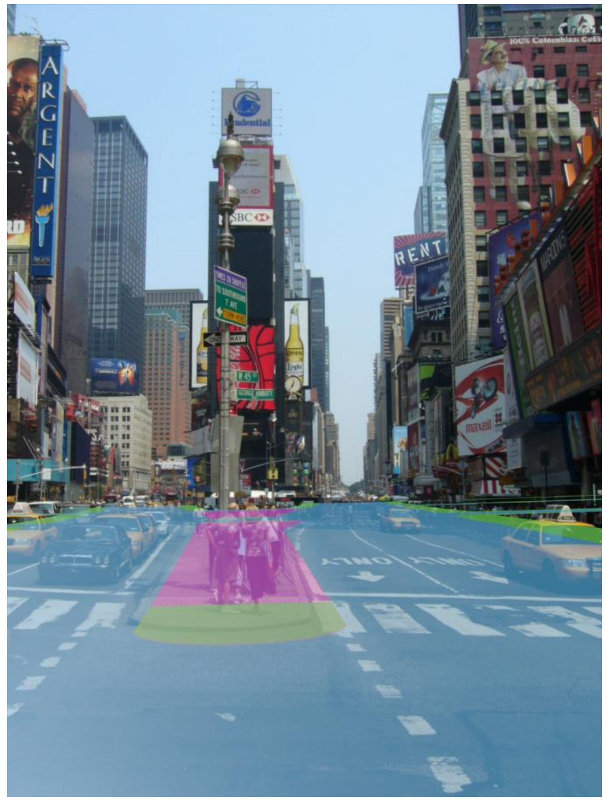


[Kevin Matzen and Noah Snavely, ICCV 2013

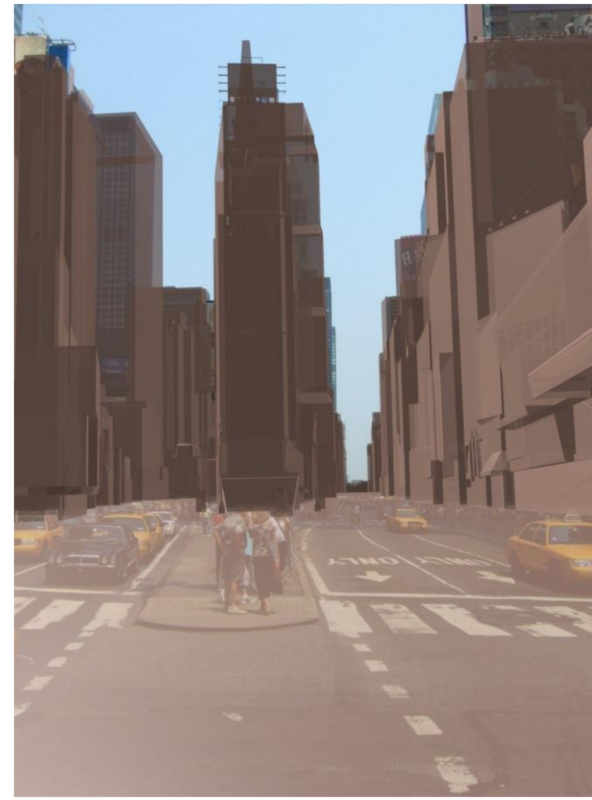
NYCOpenData Roadbeds

Vision grounded in the real world

Input photo



Overlayed GIS data
(roads / sidewalks / medians)



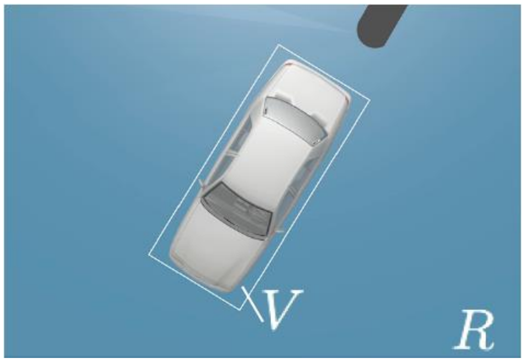
Overlayed Google Earth models

Annotated 3D Vehicles

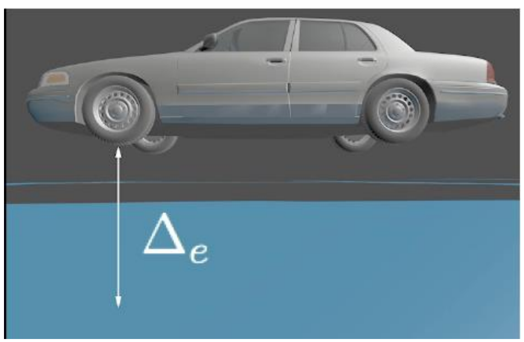
Video

3D Detection

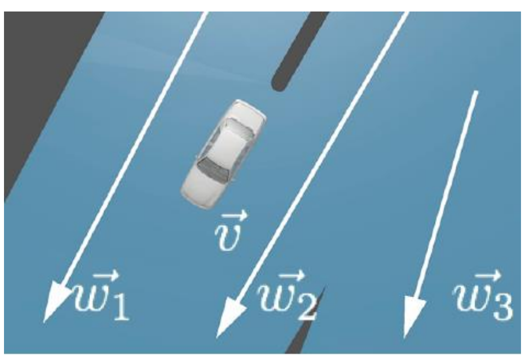
Appearance score



Ground coverage score

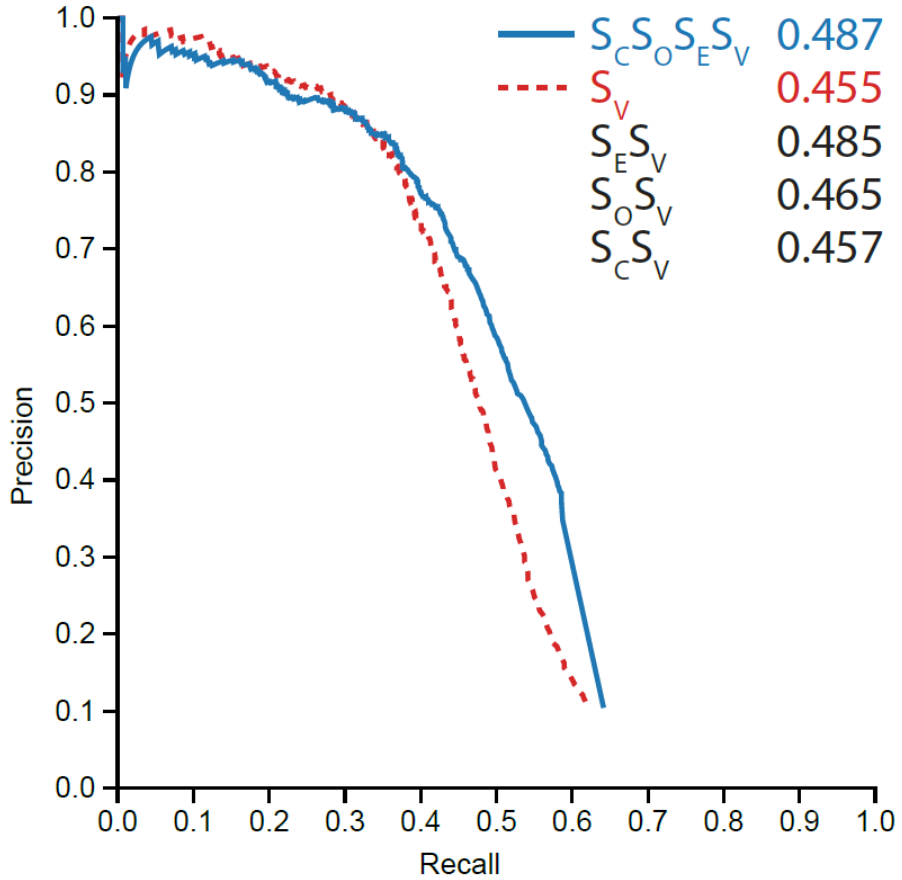


Elevation score

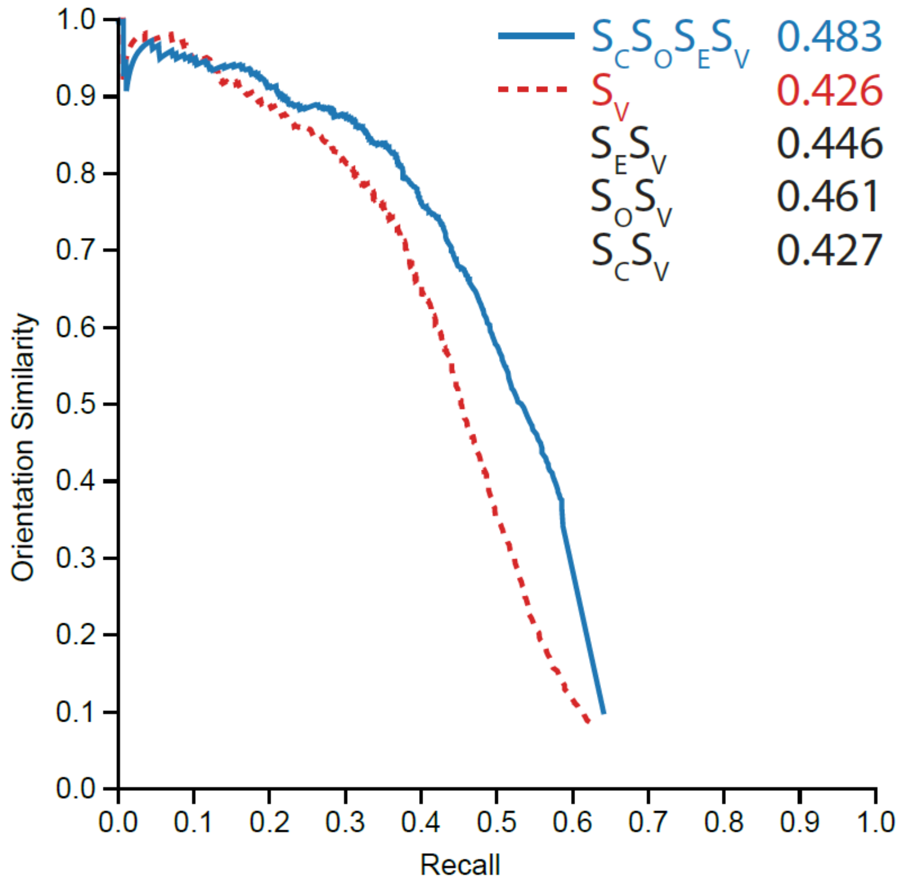


3D orientation score

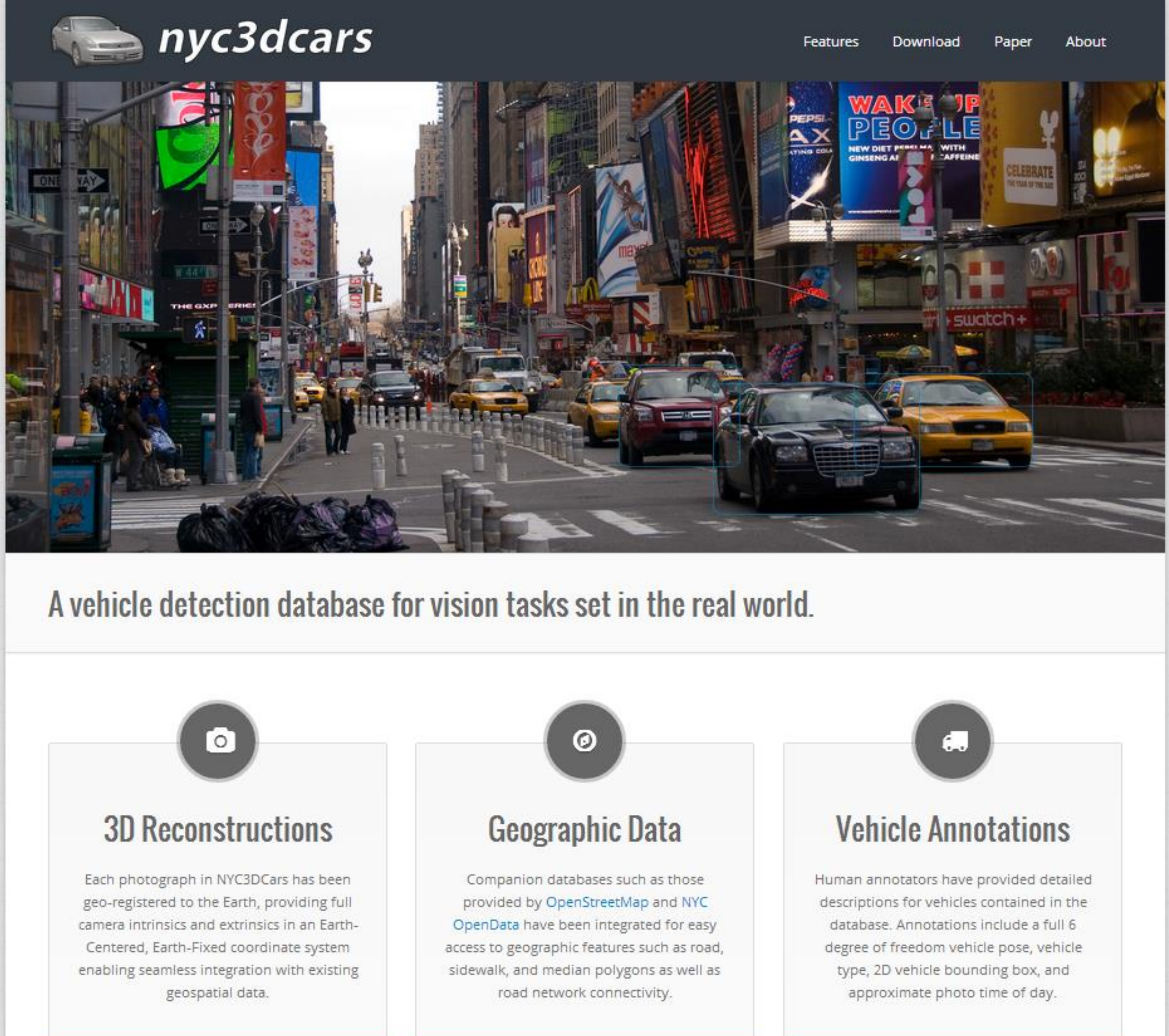
Results



Precision / Recall



Orientation similarity / Recall



http://nyc3d.cs.cornell.edu/

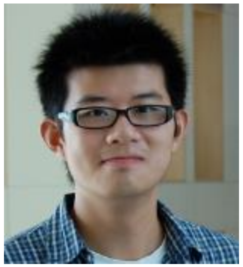
Summary

- Many interesting challenges in modeling the world
- Contributions from every area (cf. much wonderful recent work):
  - Scene understanding, object detection, material recognition, illumination modeling, …
  - Learning?

Acknowledgements

Students

Sean Bell



Song Cao

Daniel Hauagge

Kevin Matzen

Paul Upchurch

Chun-Po Wang

Scott Wehrwein

Kyle Wilson

Collaborators

Kavita Bala

Dan Huttenlocher

Yunpeng Li

Dave Crandall

- National Science Foundation
- Intel Center for Science and Technology – Visual Computing
- Amazon AWS for Education

Thank you!

More information at
http://www.cs.cornell.edu/~snavely/