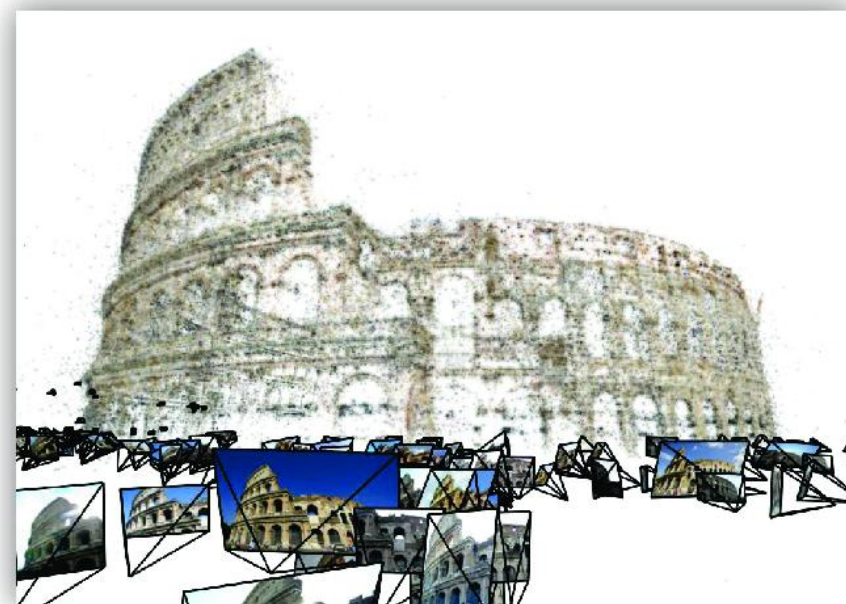


# CS6670: Computer Vision

Noah Snavely

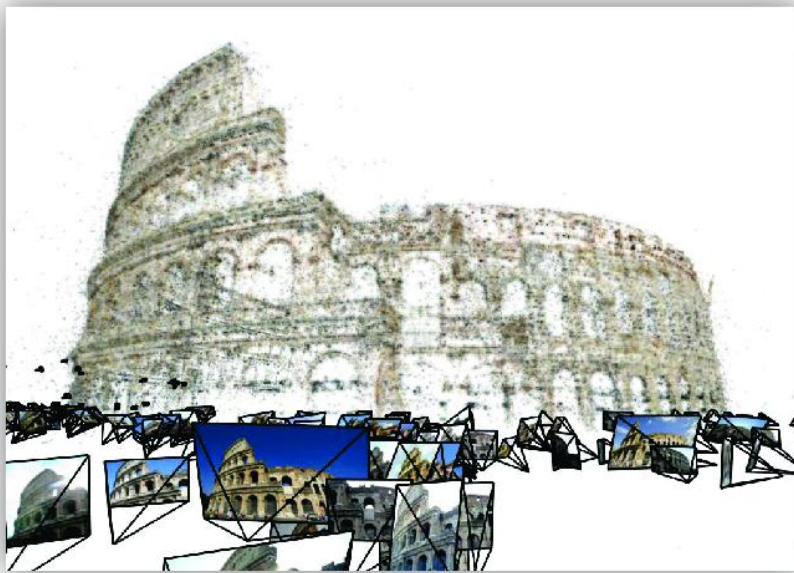
## Lecture 23: Structure from motion



# CS6670: Computer Vision

Noah Snavely

## Lecture 24: Multi-view stereo



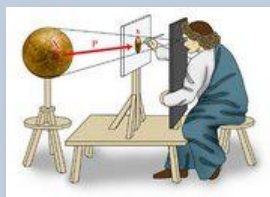
# Readings

- Szeliski, Chapter 11.6

# Final project proposals

- Great job on the proposals!
- You can go ahead and get started – I will contact a few groups with additional feedback today

# Computer Vision



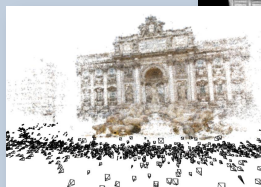
Single view



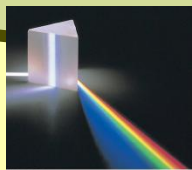
Two view



Multi-view



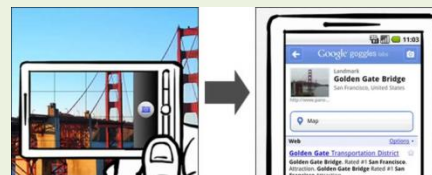
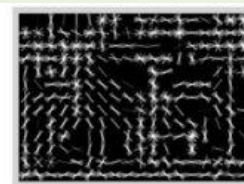
Geometry



Light and optics



Computational photography



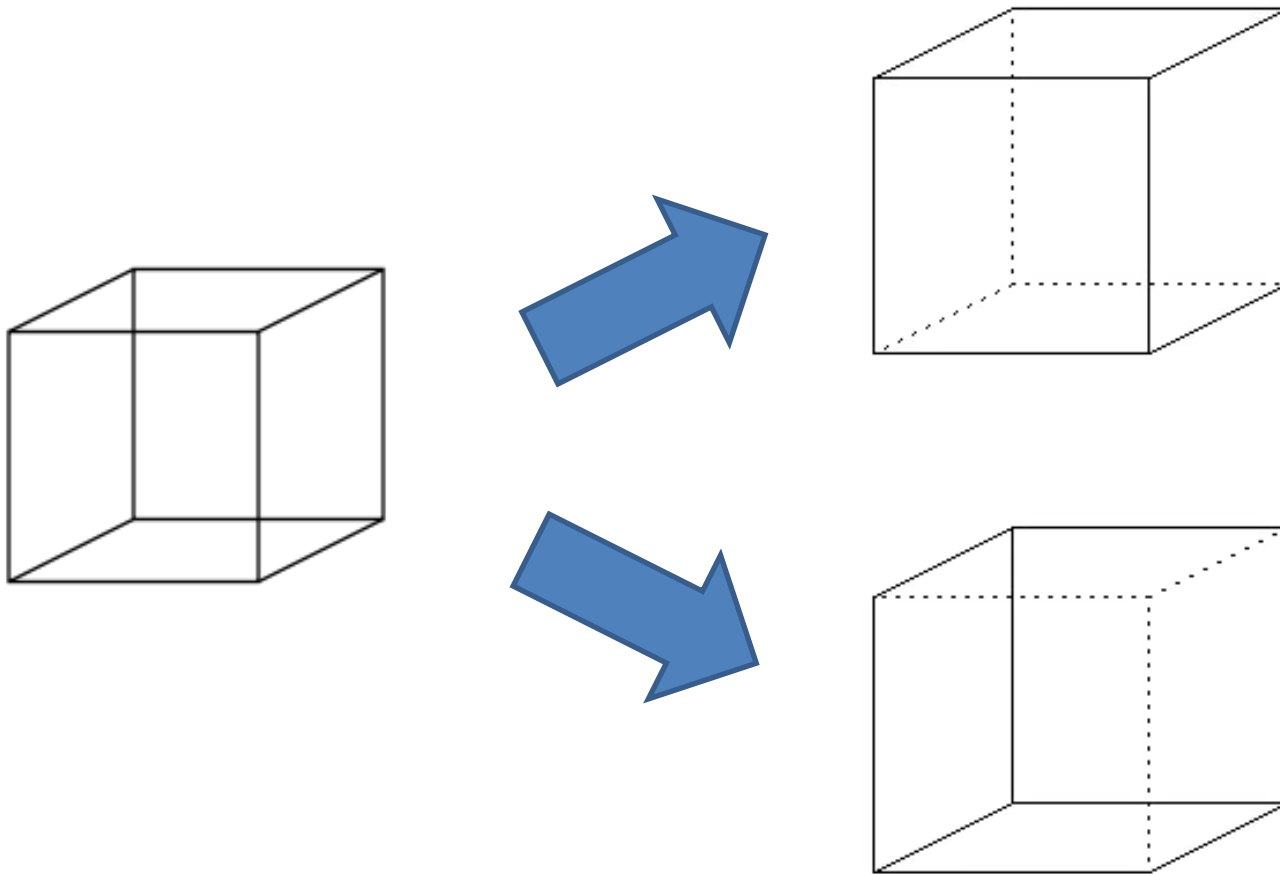
Recognition

# Libration



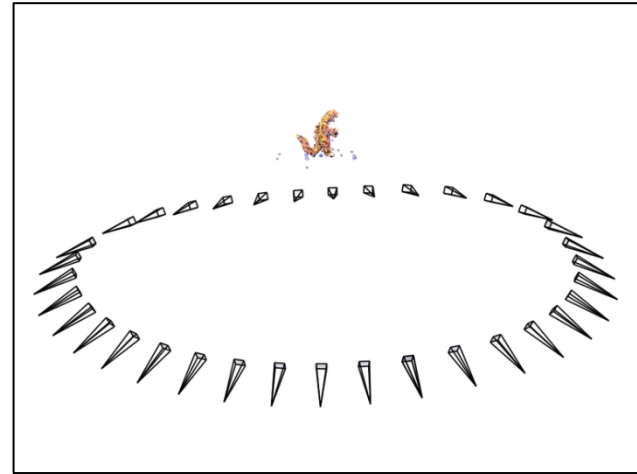
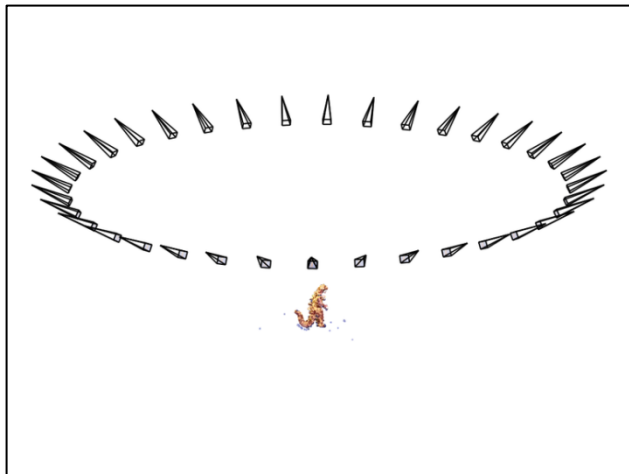
# Why SFM might fail...

- Necker reversal



# SfM – Failure cases

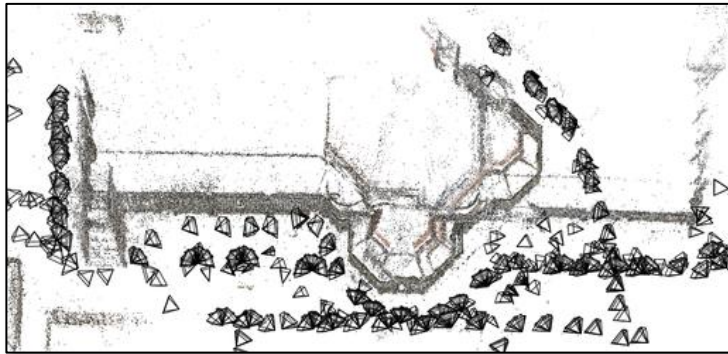
- Necker reversal





# Structure from Motion – Failure cases

- Repetitive structures



# Multi-view stereo



Stereo



Multi-view stereo

# Multi-view Stereo



[Point Grey](#)'s Bumblebee XB3



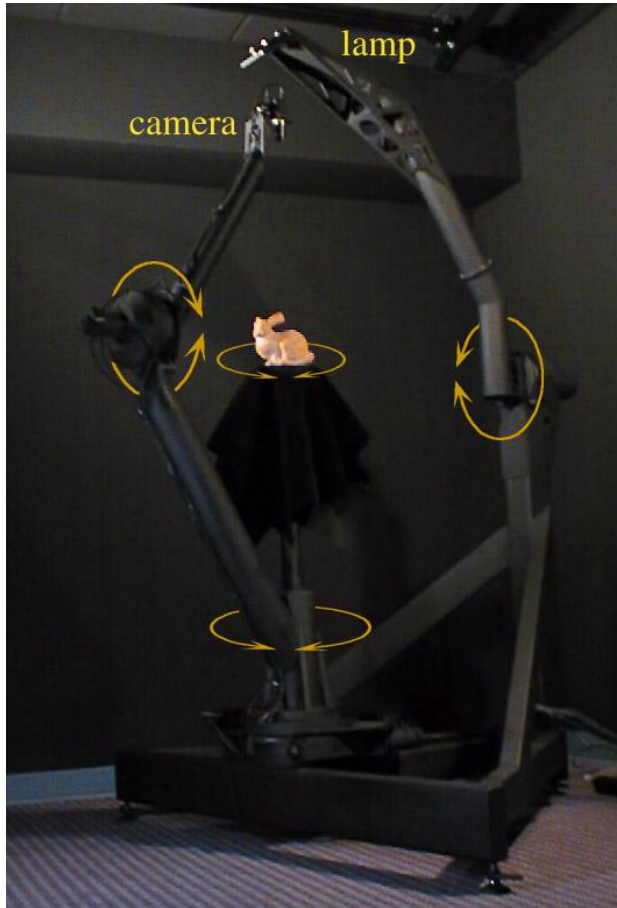
[Point Grey](#)'s ProFusion 25



CMU's [3D Room](#)



# Multi-view Stereo



flickr LOVES YOU™

Home The Tour Sign Up Explore

## Search













Photos Groups People

statue of liberty

Full text Tags only

✓ We found **80,865** results matching **statue** and **of** and **liberty**.

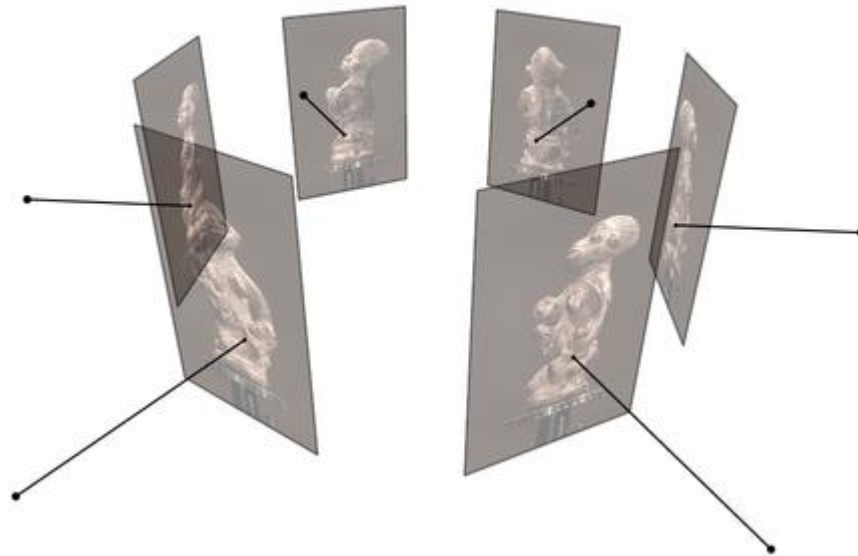
View: Most relevant • Most recent • Most interesting Show: Details • Thumbnails

 From <a href="#">mbell1975</a>	 From <a href="#">sbcreate11</a>	 From <a href="#">Marion Doss</a>	 From <a href="#">Barry Wright</a>
 From <a href="#">phileole</a>	 From <a href="#">aimk</a>	 From <a href="#">sbcreate11</a>	 From <a href="#">sbcreate11</a>
 From <a href="#">sjgardiner</a>	 From <a href="#">sjgardiner</a>	 From <a href="#">elesa.ah</a>	 From <a href="#">nicoatridge</a>

# Multi-view Stereo

Input: calibrated images from several viewpoints

Output: 3D object model

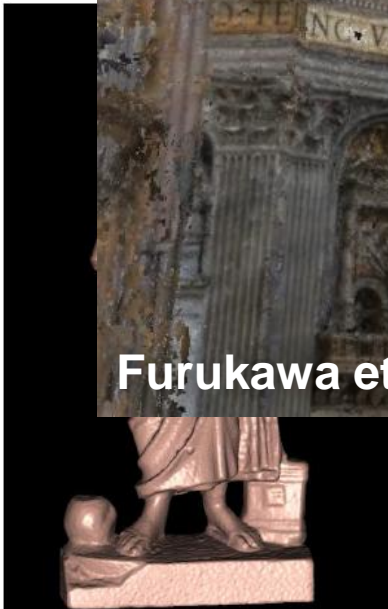


Figures by Carlos Hernandez



Furukawa et al., 2010

Faugeras, Keriven  
**1998**



Hernandez, Schmitt  
**2004**



Pons, Keriven, Faugeras  
**2005**

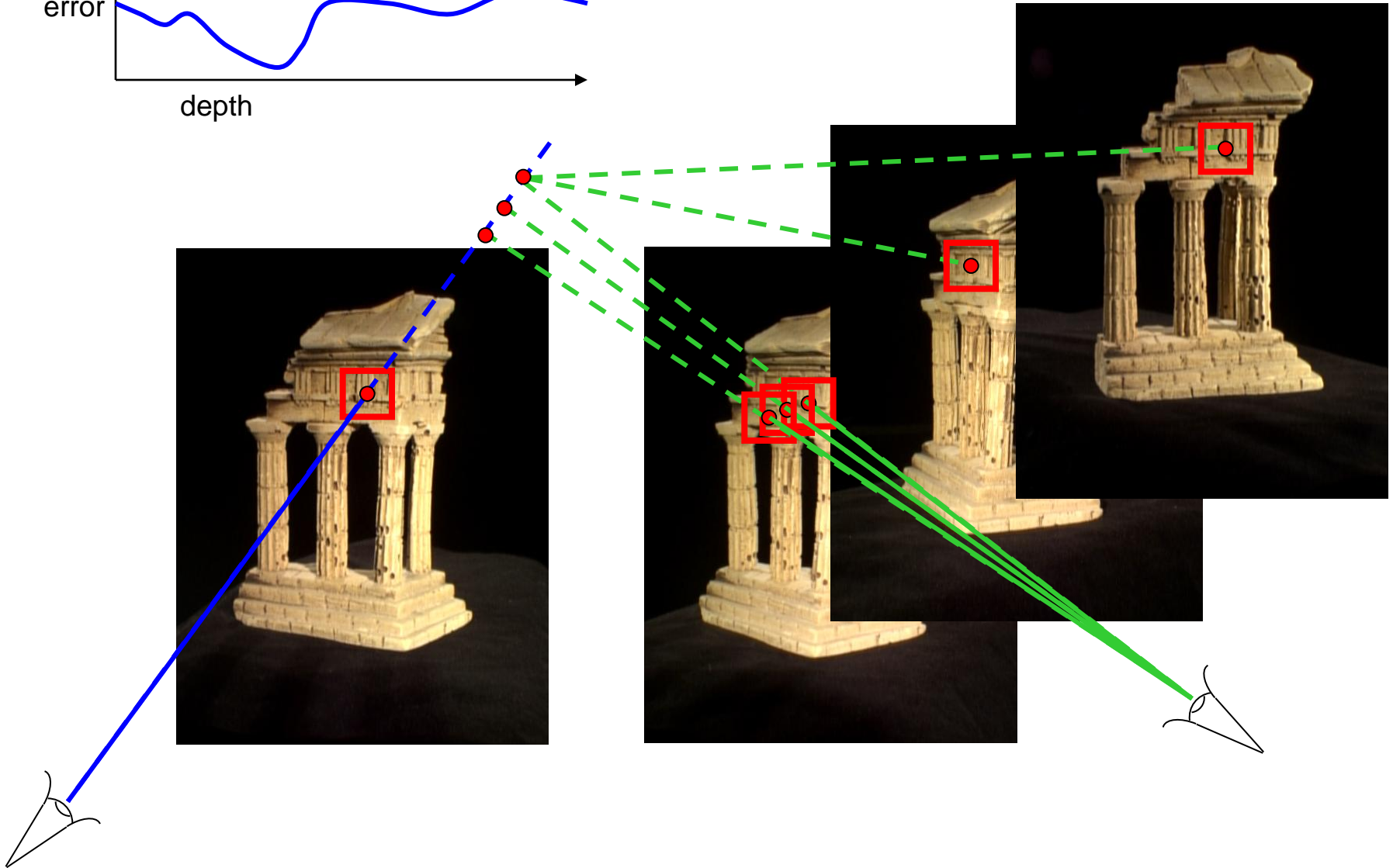
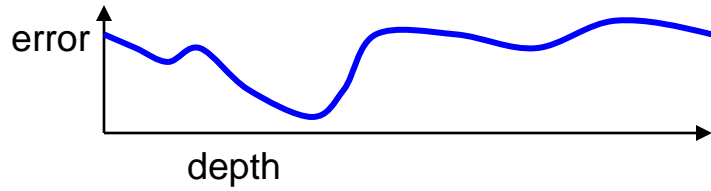


Furukawa, Ponce  
**2006**



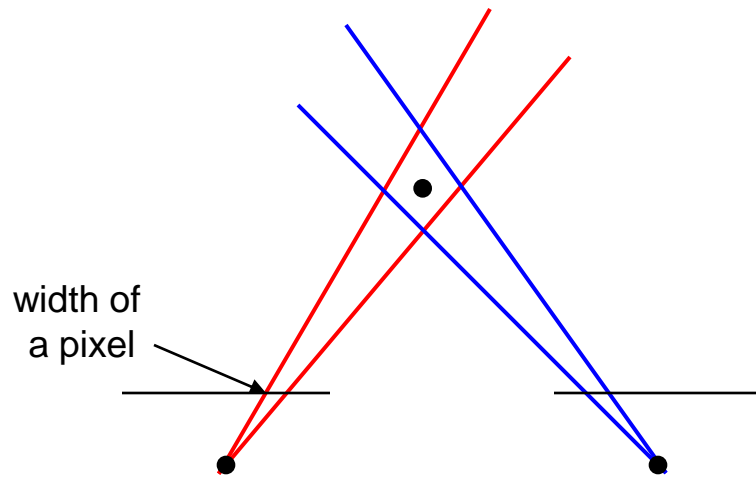
Goesele et al.  
**2007**

# Stereo: another view

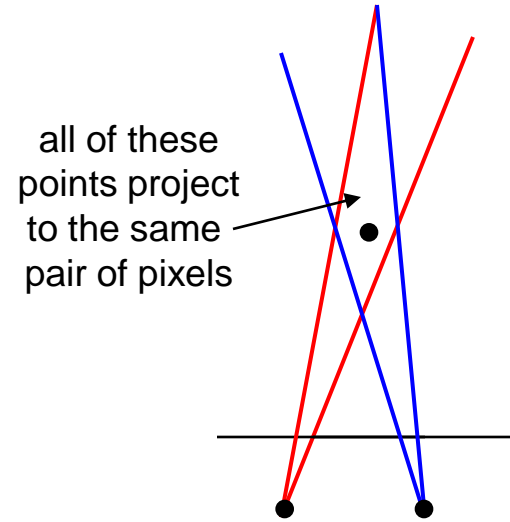




# Choosing the stereo baseline



**Large Baseline**



**Small Baseline**

What's the optimal baseline?

- Too small: large depth error
- Too large: difficult search problem



# The Effect of Baseline on Depth Estimation

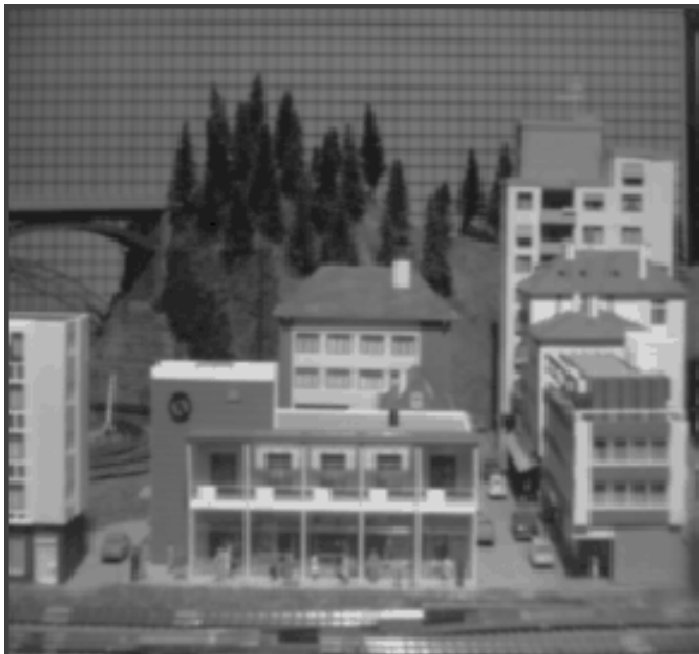
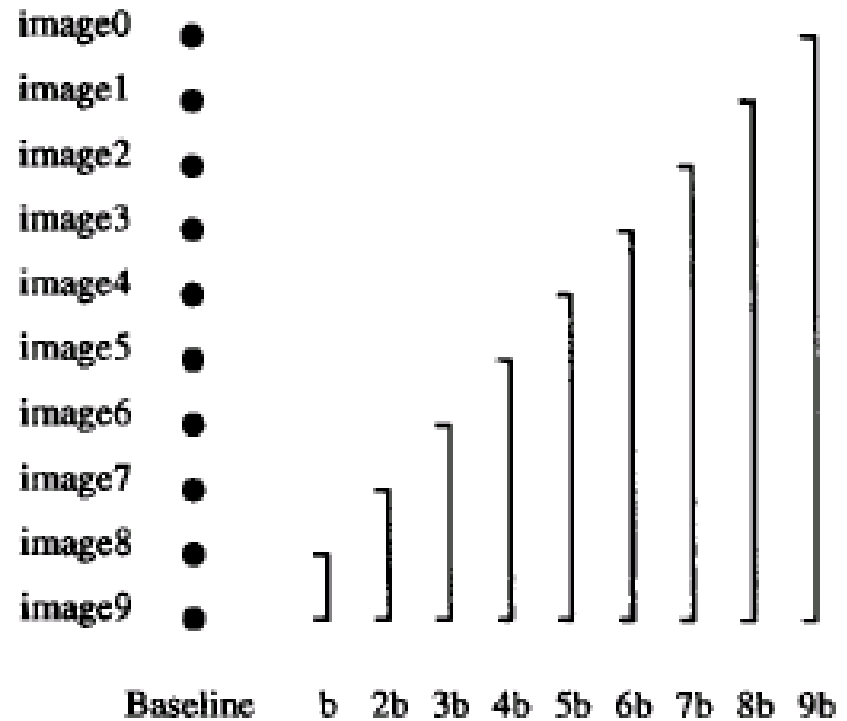
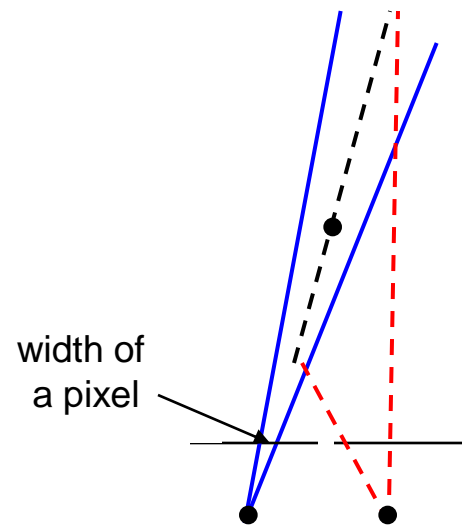
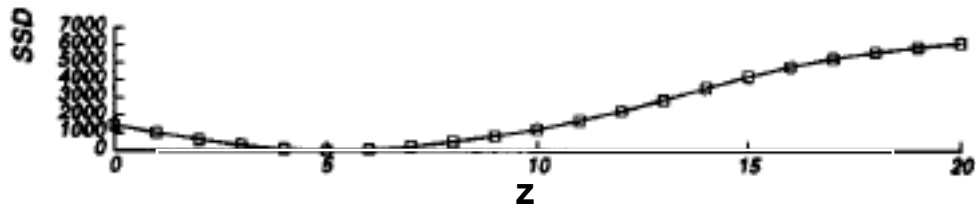
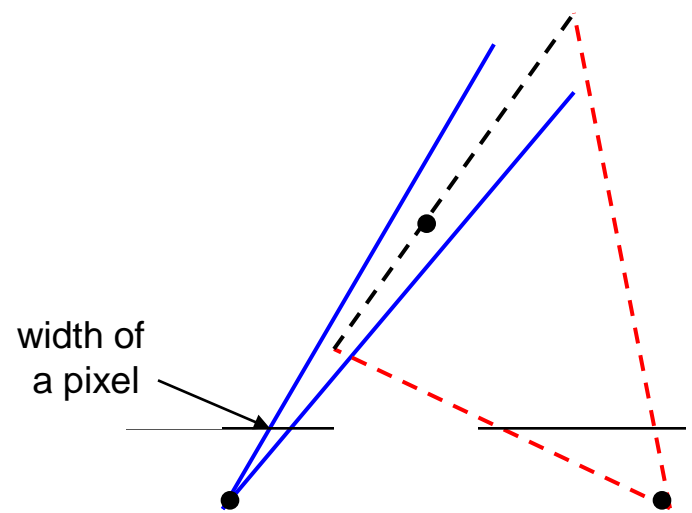
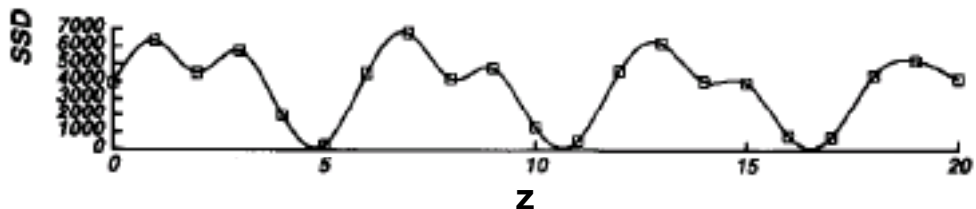


Figure 2: An example scene. The grid pattern in the background has ambiguity of matching.





pixel matching score



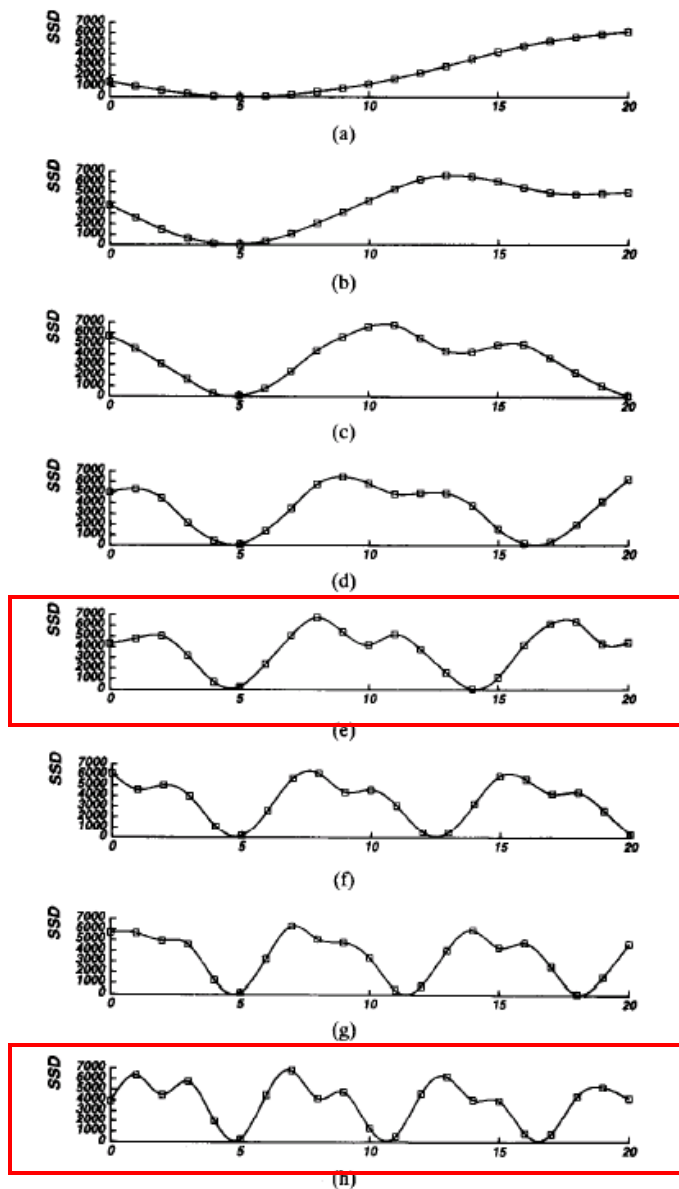


Fig. 5. SSD values versus inverse distance: (a)  $B = b$ ; (b)  $B = 2b$ ; (c)  $B = 3b$ ; (d)  $B = 4b$ ; (e)  $B = 5b$ ; (f)  $B = 6b$ ; (g)  $B = 7b$ ; (h)  $B = 8b$ . The horizontal axis is normalized such that  $8bF = 1$ .

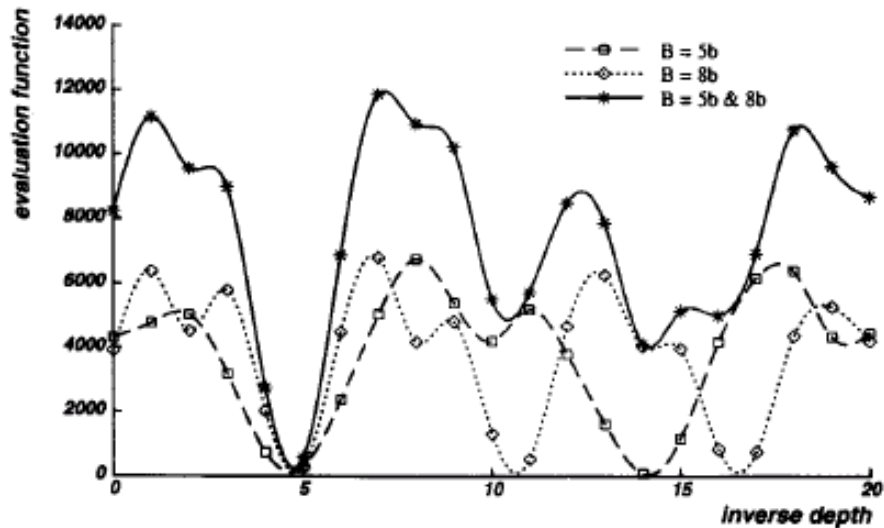


Fig. 6. Combining two stereo pairs with different baselines.

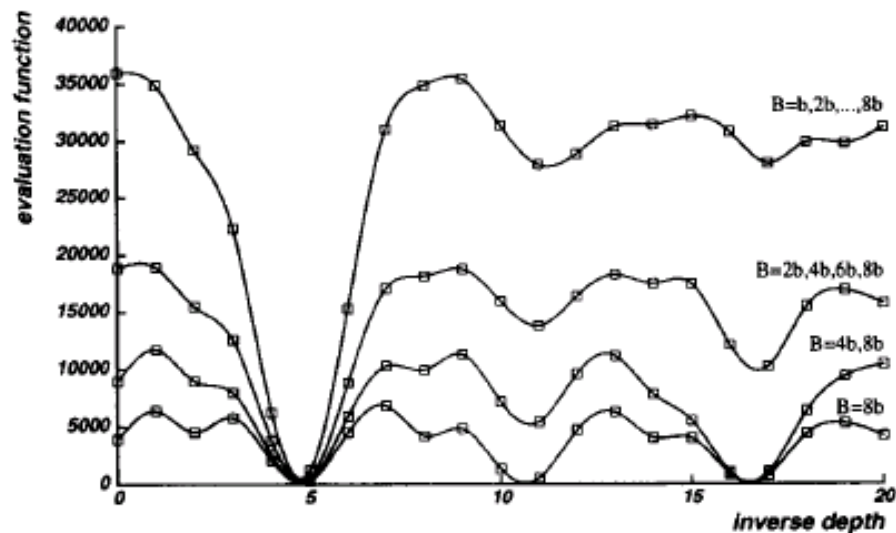


Fig. 7. Combining multiple baseline stereo pairs.

# Multibaseline Stereo

## Basic Approach

- Choose a reference view
- Use your favorite stereo algorithm BUT
  - replace two-view SSD with SSSD over all baselines

## Limitations



# Problem: *visibility*

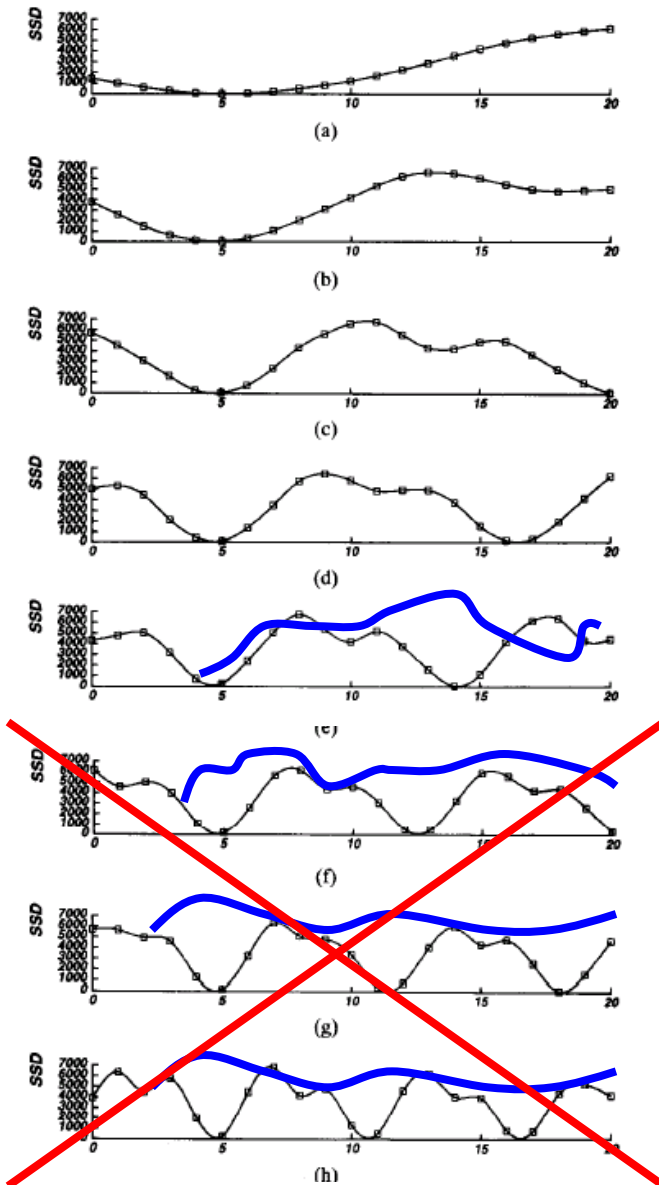


Fig. 5. SSD values versus inverse distance: (a)  $B = b$ ; (b)  $B = 2b$ ; (c)  $B = 3b$ ; (d)  $B = 4b$ ; (e)  $B = 5b$ ; (f)  $B = 6b$ ; (g)  $B = 7b$ ; (h)  $B = 8b$ . The horizontal axis is normalized such that  $8bF = 1$ .

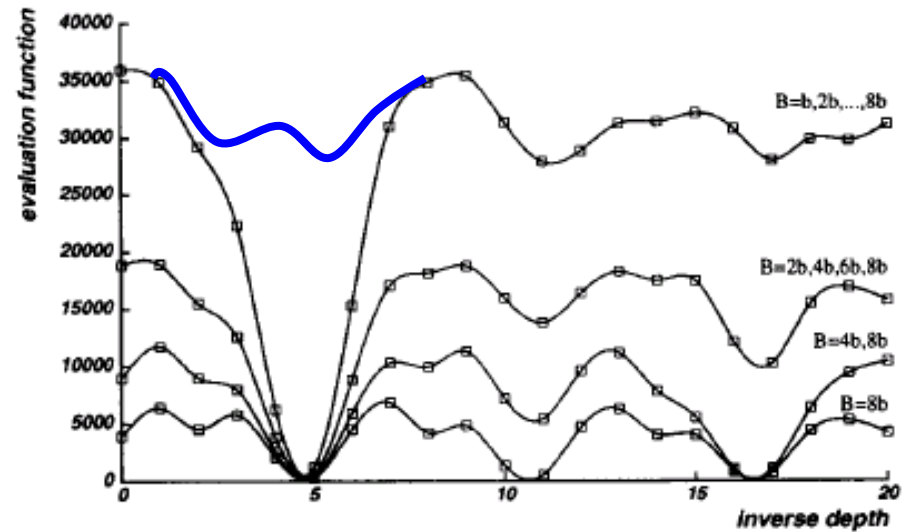


Fig. 7. Combining multiple baseline stereo pairs.

## Some Solutions

- Match only nearby photos [Narayanan 98]
- Use NCC instead of SSD, Ignore NCC values > threshold [Hernandez & Schmitt 03]

# Popular matching scores

- SSD (Sum Squared Distance)

$$\sum_{x,y} |W_1(x,y) - W_2(x,y)|^2$$

- NCC (Normalized Cross Correlation)

$$\frac{\sum_{x,y} (W_1(x,y) - \overline{W_1})(W_2(x,y) - \overline{W_2})}{\sigma_{W_1} \sigma_{W_2}}$$

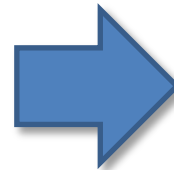
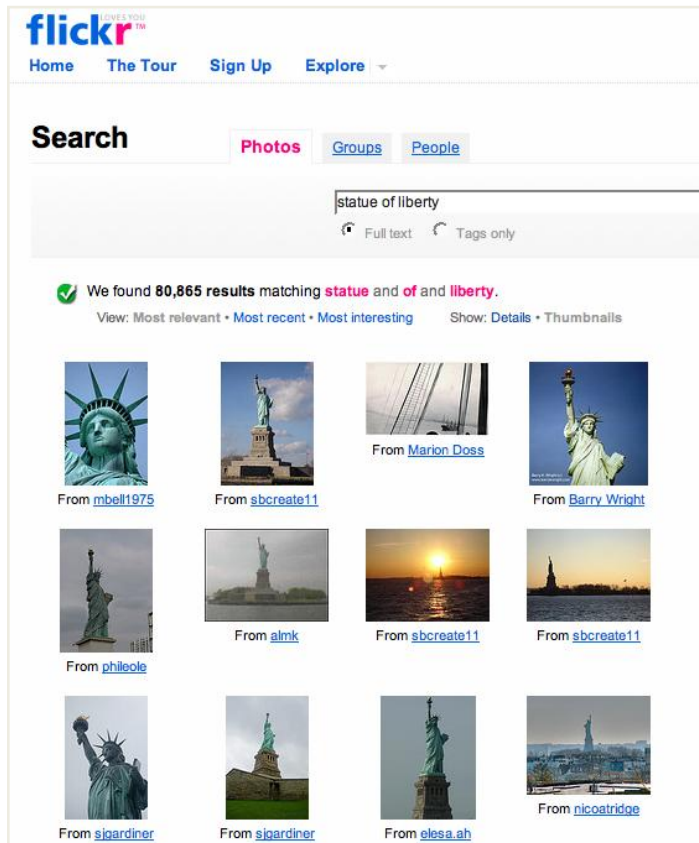
– where  $\overline{W_i} = \frac{1}{n} \sum_{x,y} W_i$        $\sigma_{W_i} = \sqrt{\frac{1}{n} \sum_{x,y} (W_i - \overline{W_i})^2}$

– what advantages might NCC have?

Questions?

# Multi-view stereo from Internet Collections

[\[Goesele, Snavely, Curless, Hoppe, Seitz, ICCV 2007\]](#)





# Challenges

- appearance variation

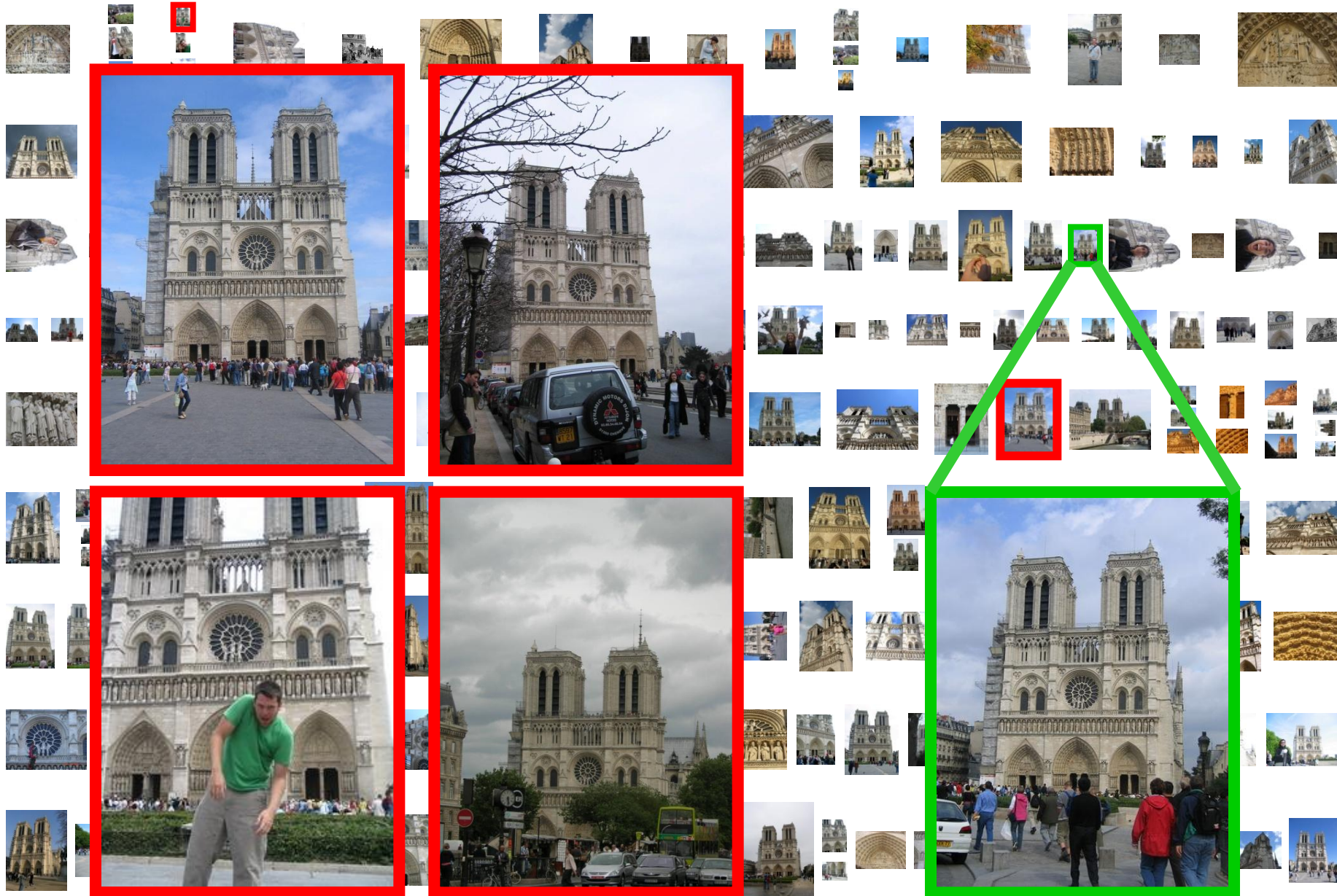


- resolution



- massive collections

82,754 results for photos matching **notre** and **dame** and **paris**





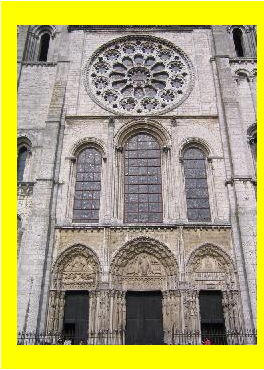


4 best neighboring views



reference view

- Automatically select neighboring views for each point in the image
- Desiderata: good matches AND good baselines



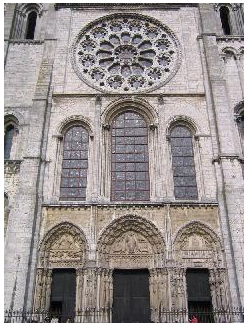
4 best neighboring views



reference view

- Automatically select neighboring views for each point in the image
- Desiderata: good matches AND good baselines





4 best neighboring views

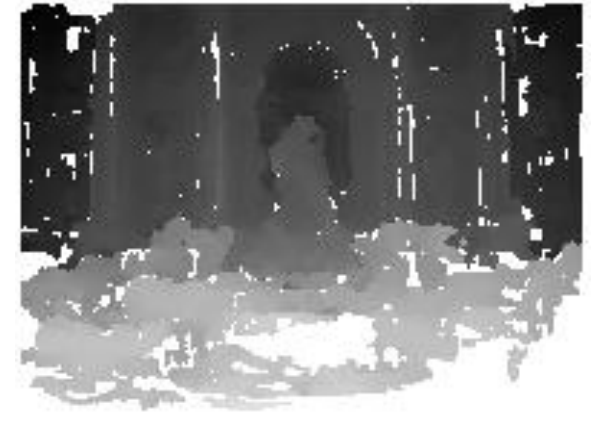
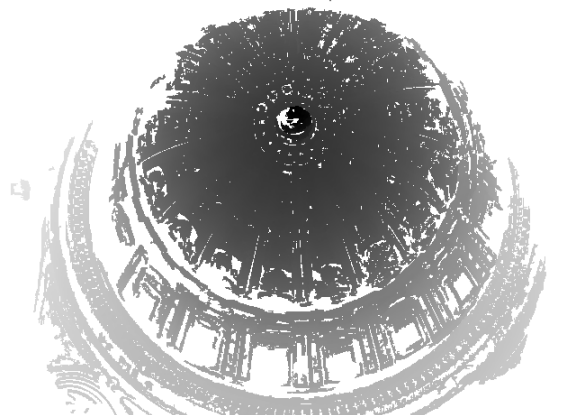
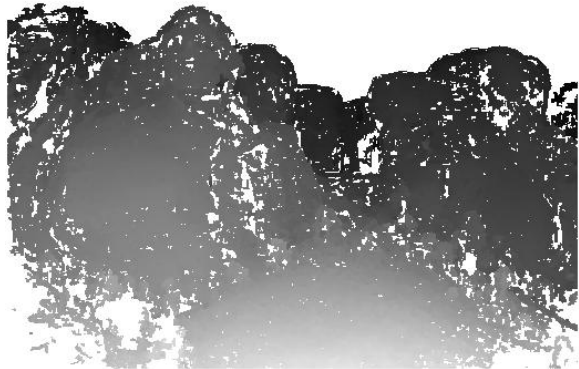


reference view

## Local view selection

- Automatically select neighboring views for each **point** in the image
- Desiderata: good matches AND good baselines

# Results

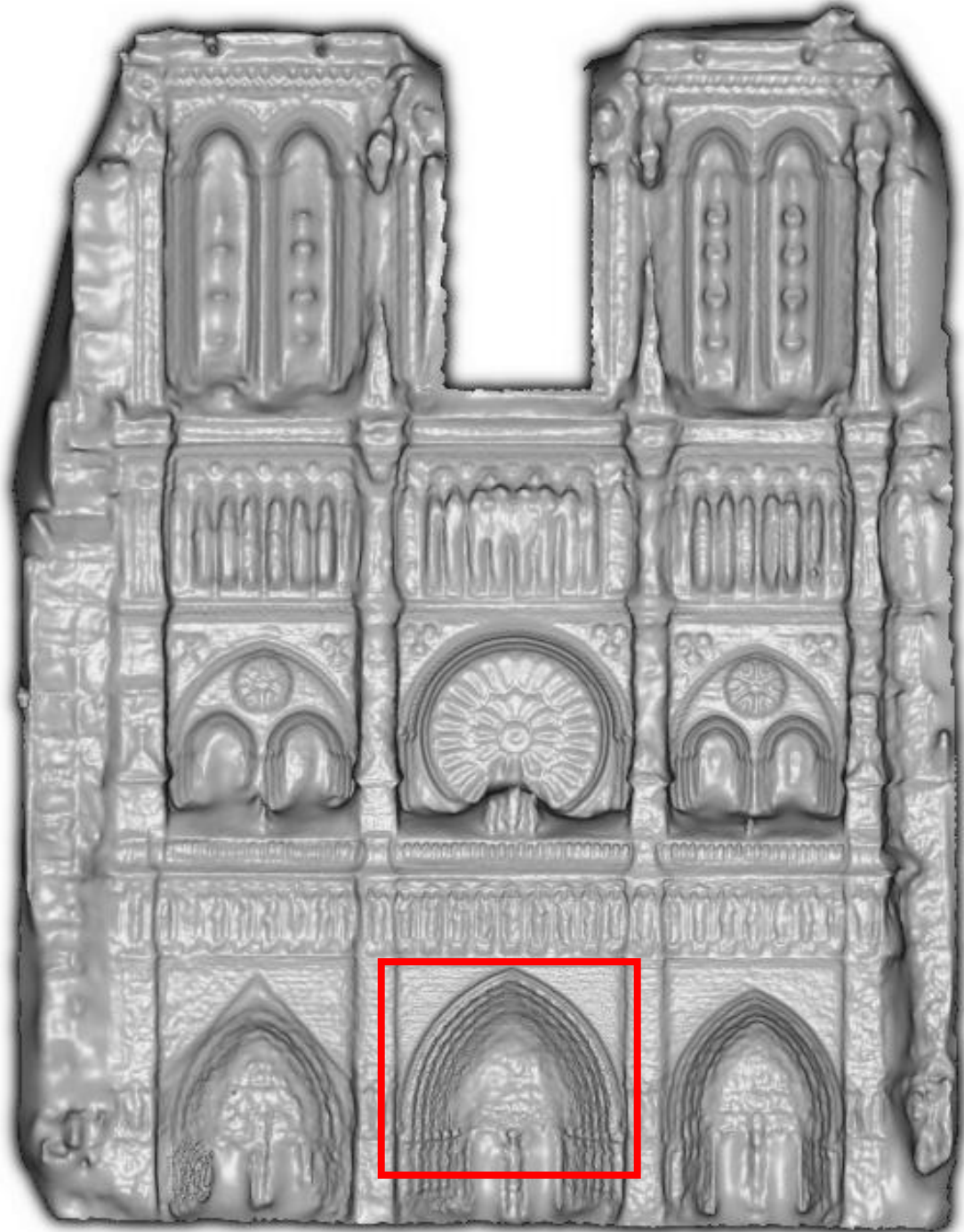


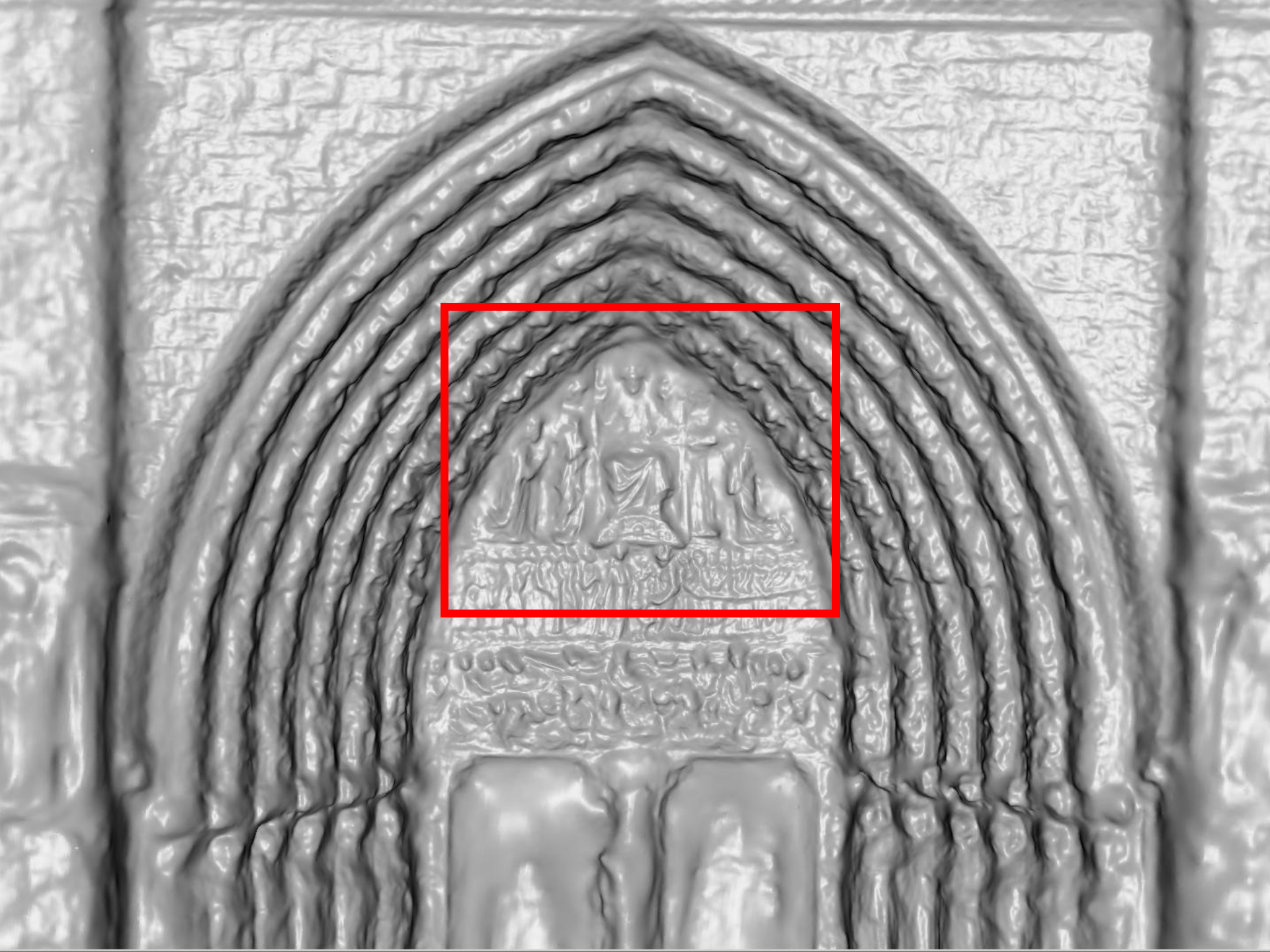


Notre Dame de Paris

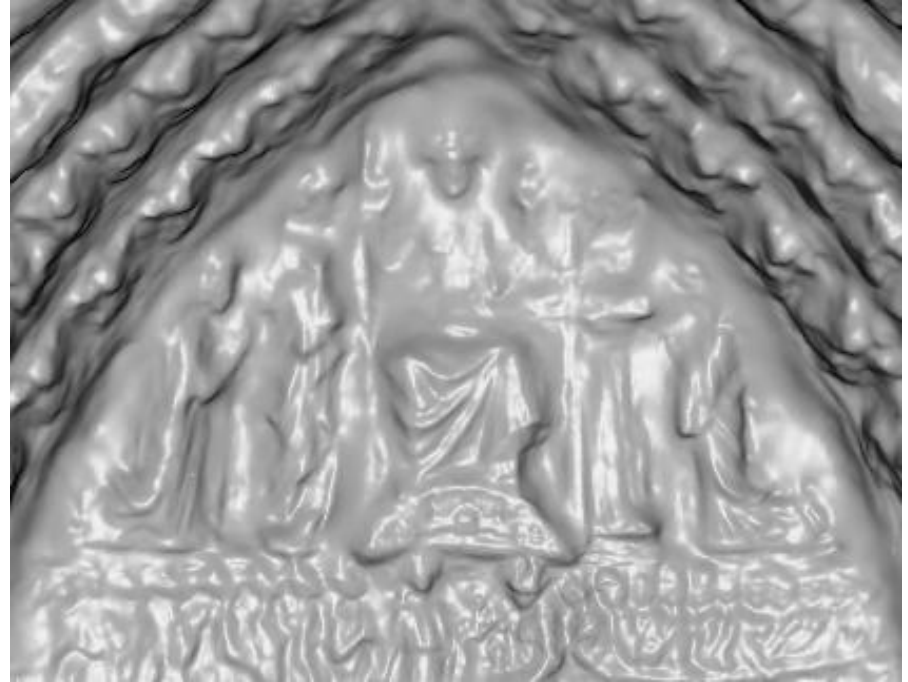
653 images

313 photographers

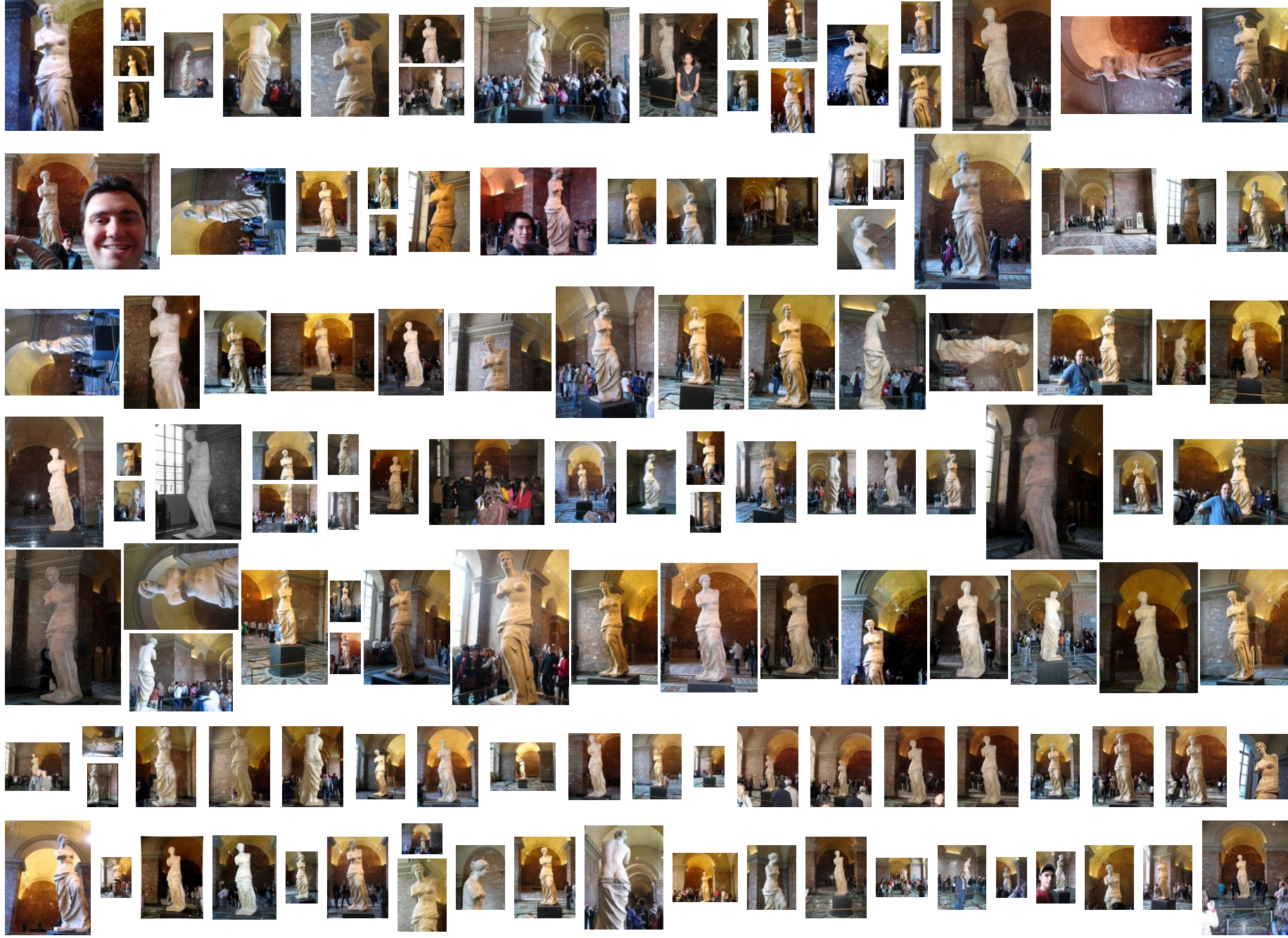








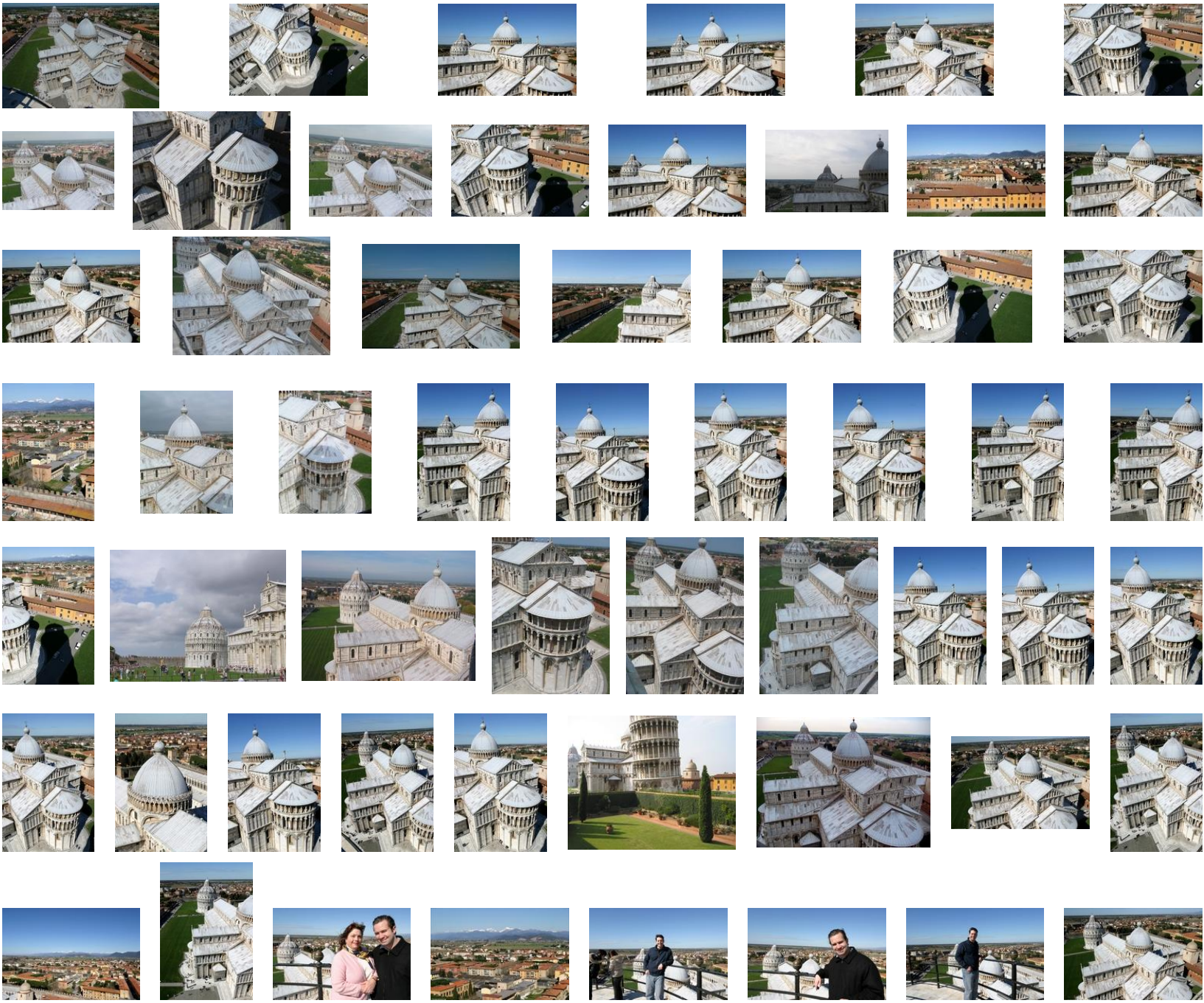


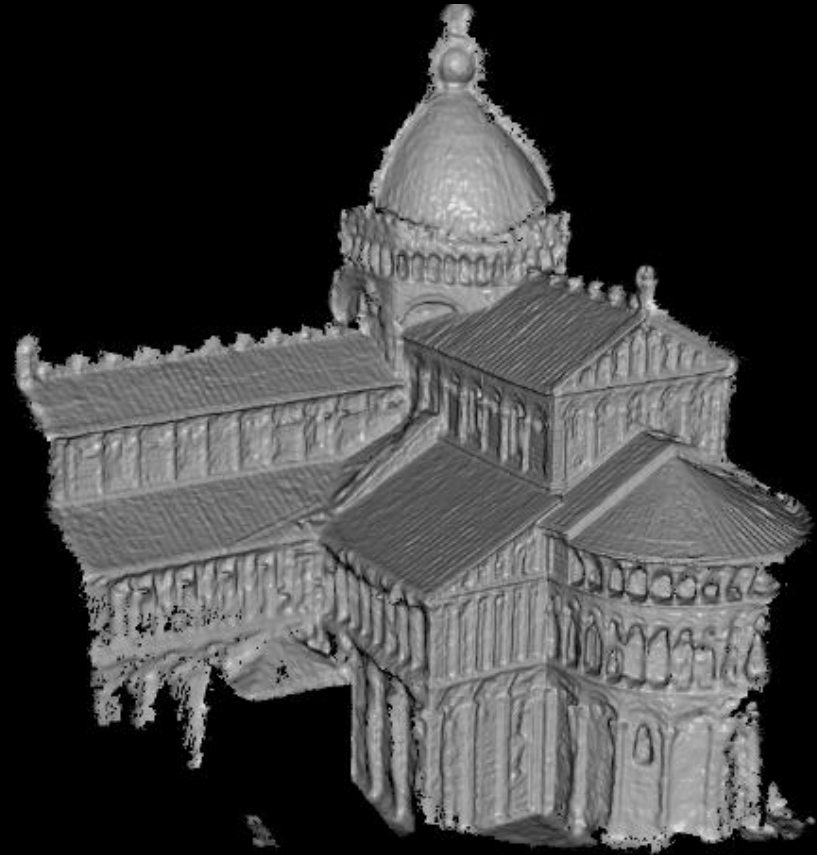




merged model of Venus de Milo

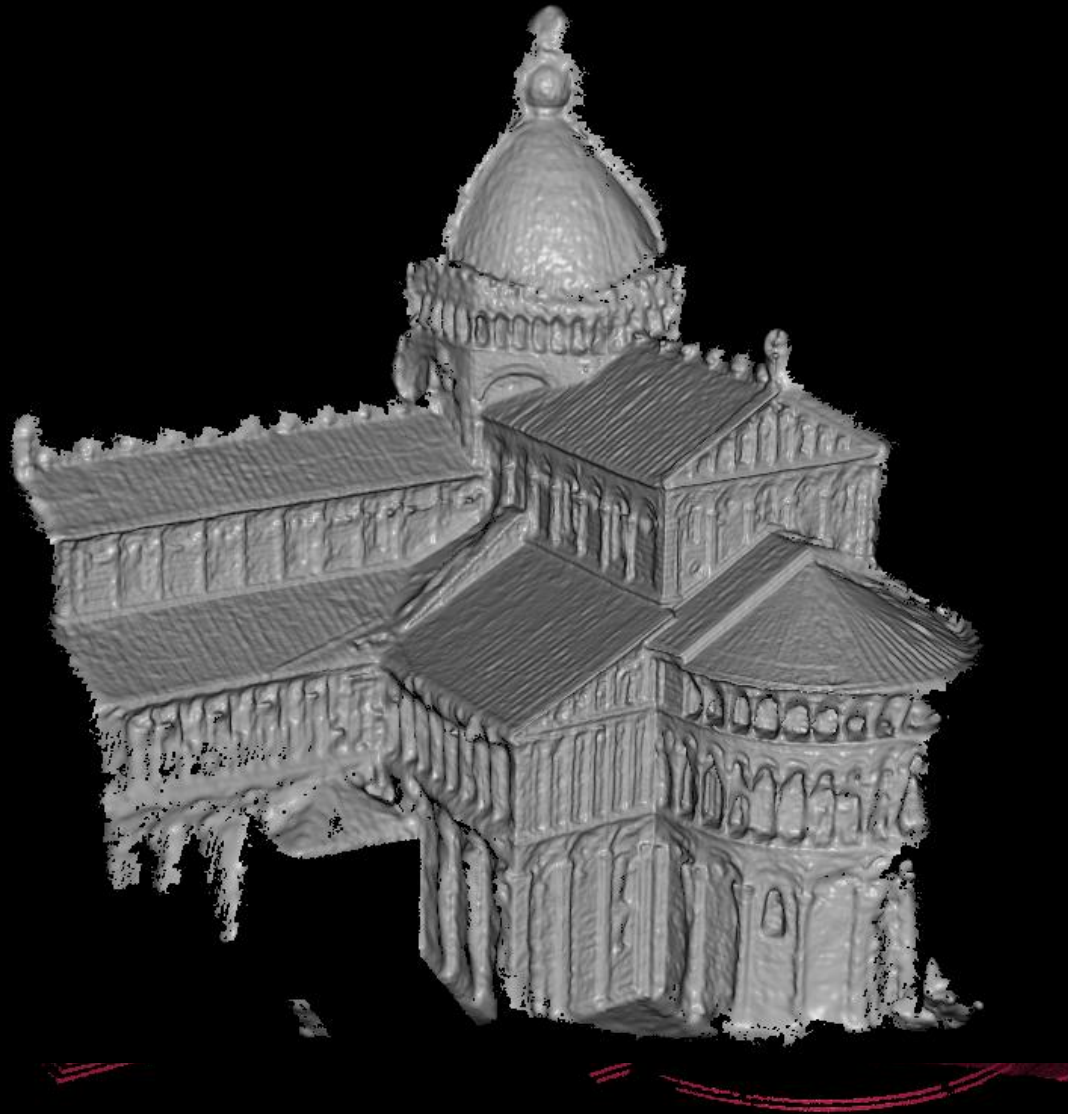






merged model of Pisa Cathedral





Accuracy compared to laser scanned model:  
90% of points within 0.25% of ground truth

# Demo