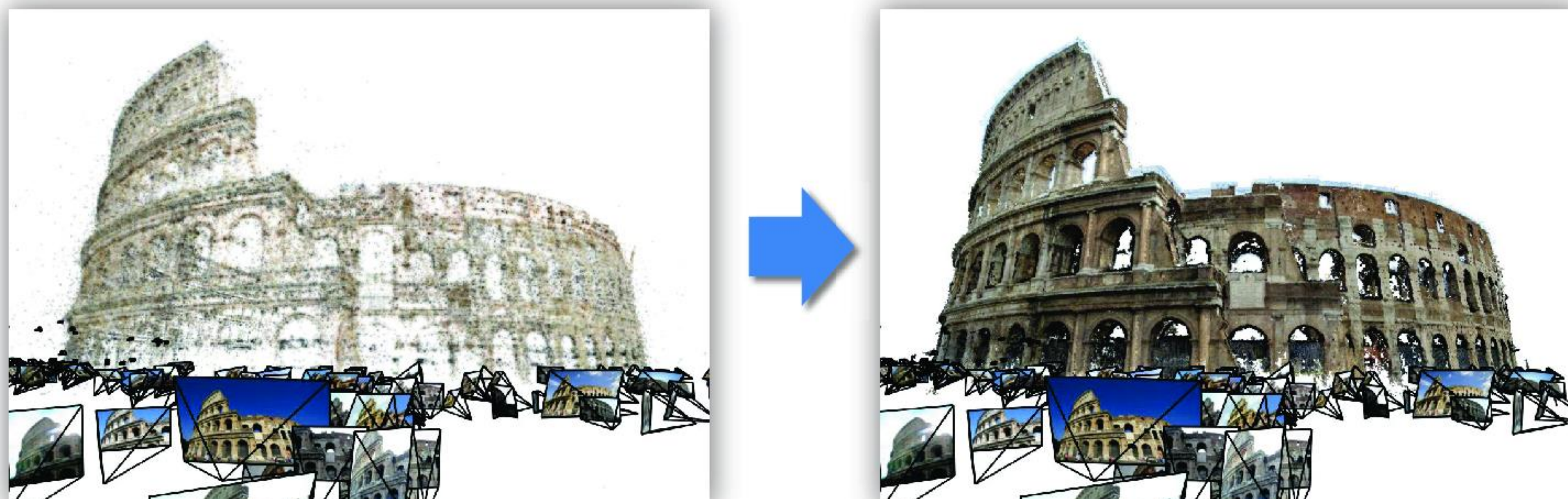
CS5670: Computer Vision

Noah Snavely

Multi-view stereo



Announcements

- Project 4 due April 22, 2019, by 11:59pm
- Final quiz in-class next Wednesday, April 24
- In-class final exam, May 6

Recommended Reading

- Szeliski Chapter 11.6
- *Multi-View Stereo: A Tutorial*, Furukawa and Hernandez, 2015
  - http://carlos-hernandez.org/papers/fnt_mvs_2015.pdf

Multi-view Stereo

Problem formulation: given several images of the same object or scene, compute a representation of its 3D shape

Stereo

Multi-view stereo

Multi-view Stereo

Point Grey's Bumblebee XB3

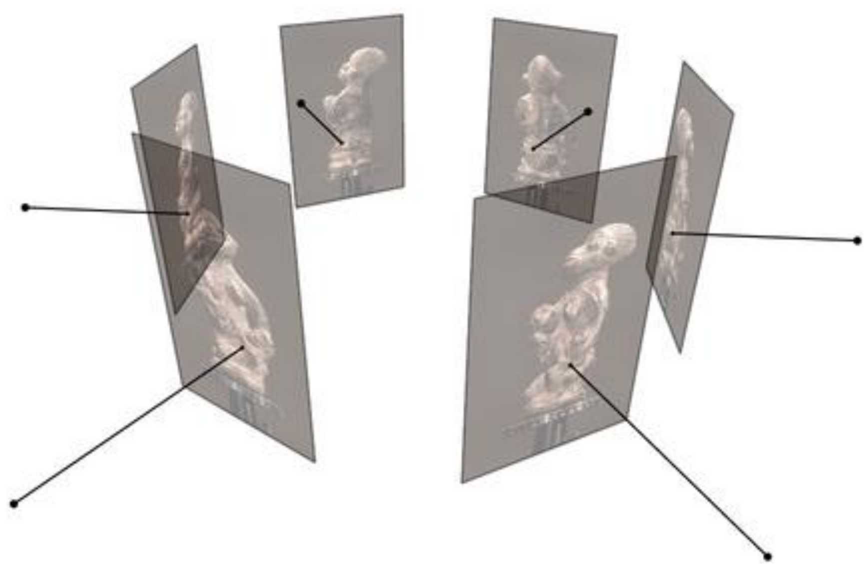
Point Grey's ProFusion 25

CMU's 3D Room

Multi-view Stereo

Input: calibrated images from several viewpoints (known intrinsics and extrinsics)

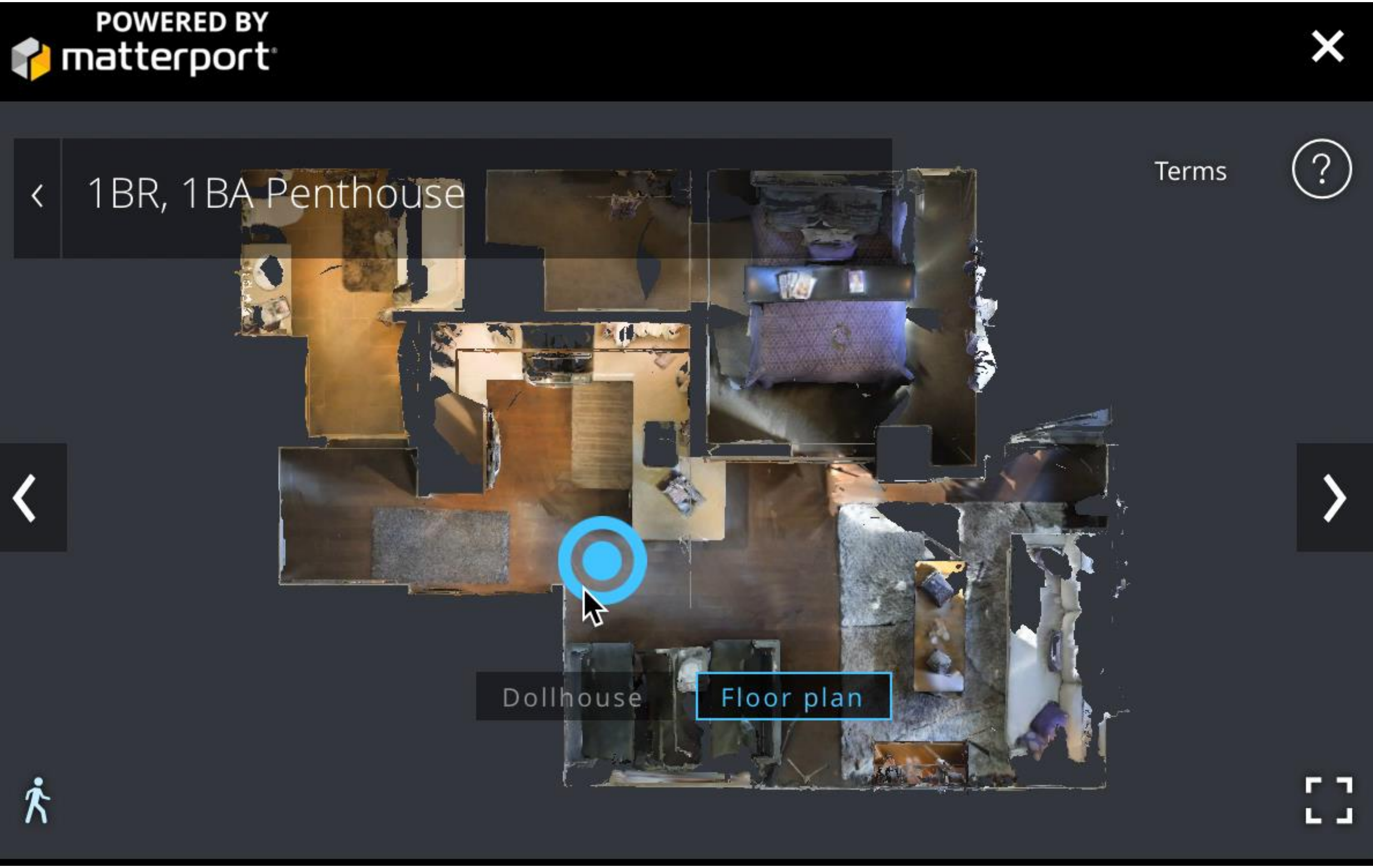
Output: 3D object model



Figures by Carlos Hernandez

Applications


Data SIO, NOAA, U.S. Navy, NGA, GEBCO
Google earth


POWERED BY
matterport
1BR, 1BA Penthouse
Terms
Dollhouse
Floor plan

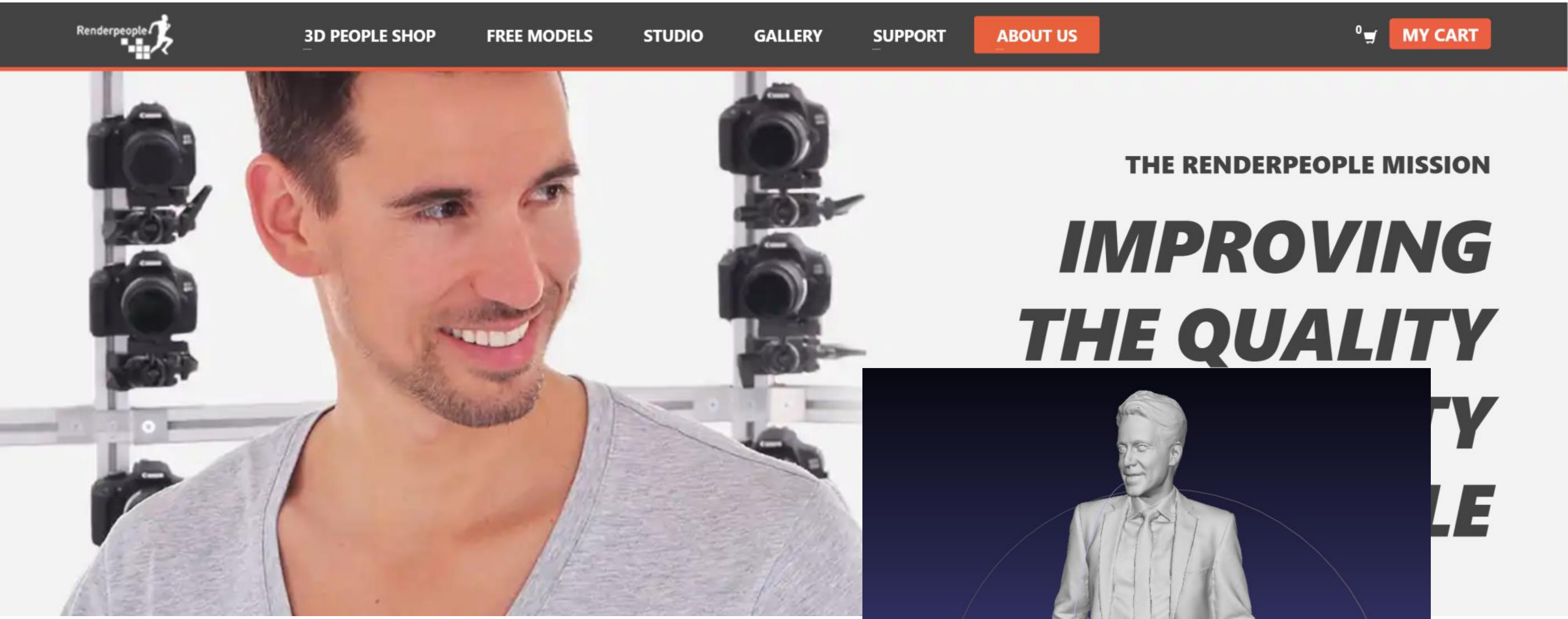
Whistle in the Form of Female Figure *600 AD - 900 AD*

Details Los Angeles County Museum of Art

LACMA Los Angeles County Museum of Art | Sculpture | Mexico

Share Compare Saved 0 Discover

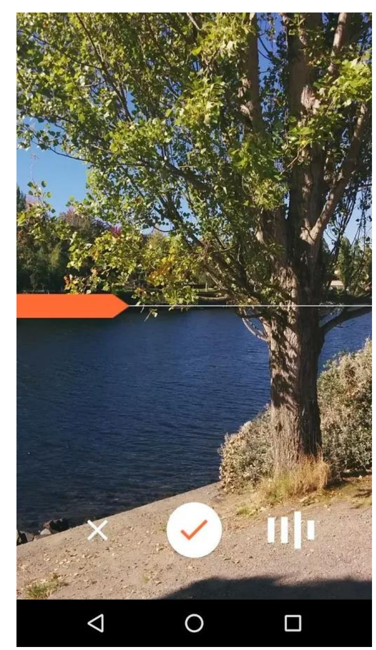
Google



https://renderpeople.com/about-us/

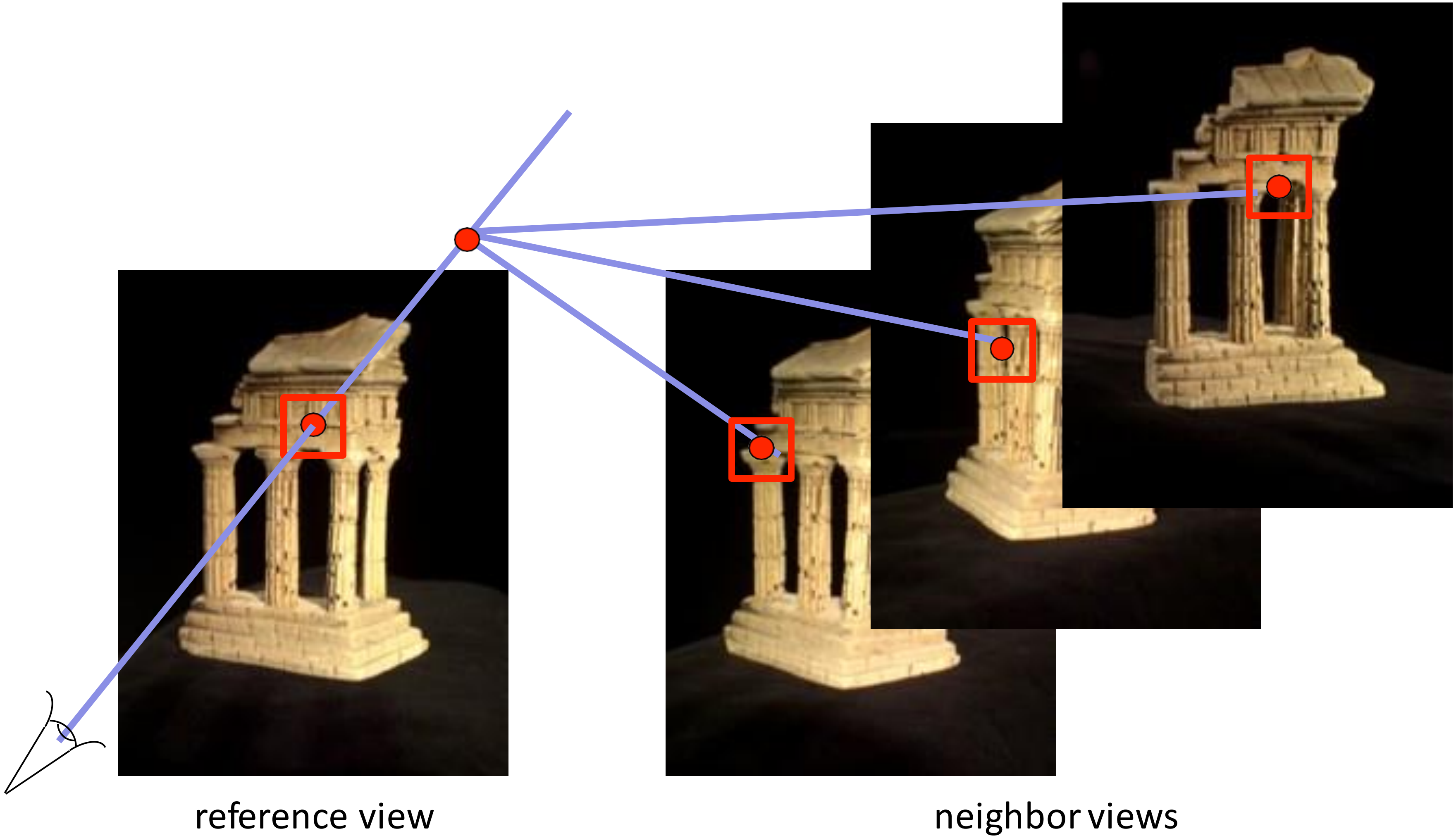
Google

JUMP



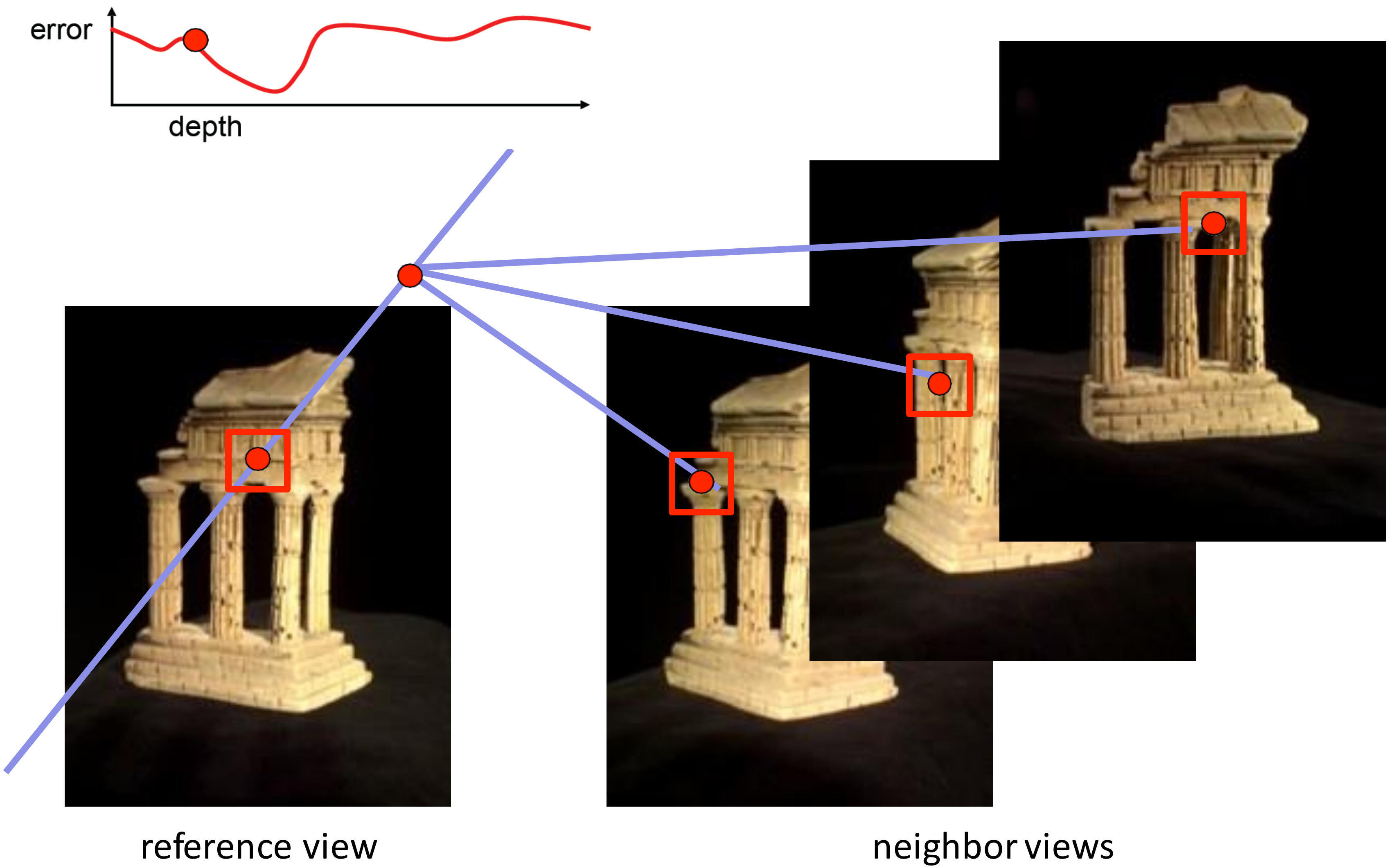
https://code.facebook.com/posts/1755691291326688/introducing-facebook-surround-360-an-open-high-quality-3d-360-video-capture-system?hc_location=ufi

Multi-view stereo: Basic idea



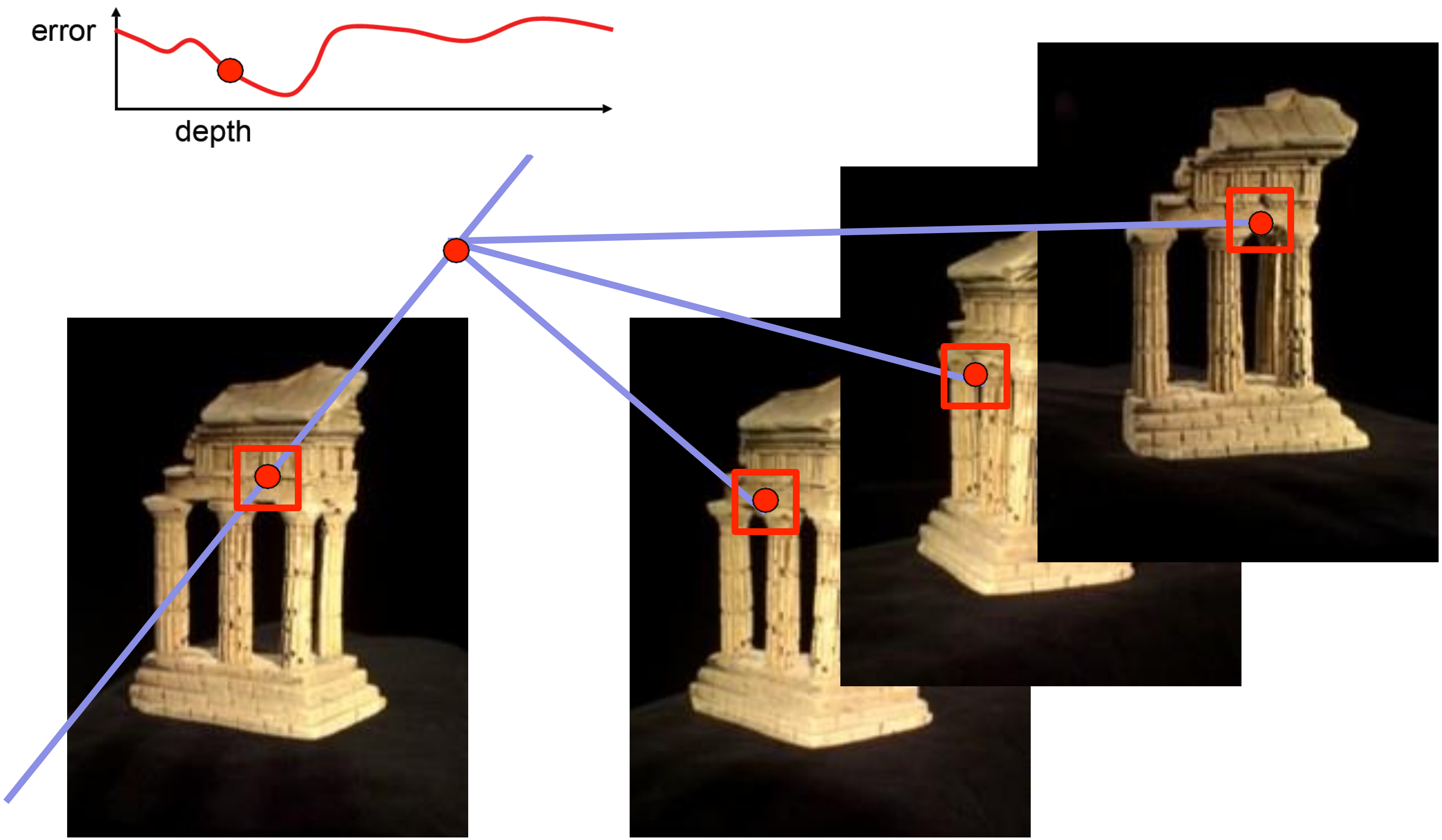
Source: Y. Furukawa

Multi-view stereo: Basic idea



Source: Y. Furukawa

Multi-view stereo: Basic idea

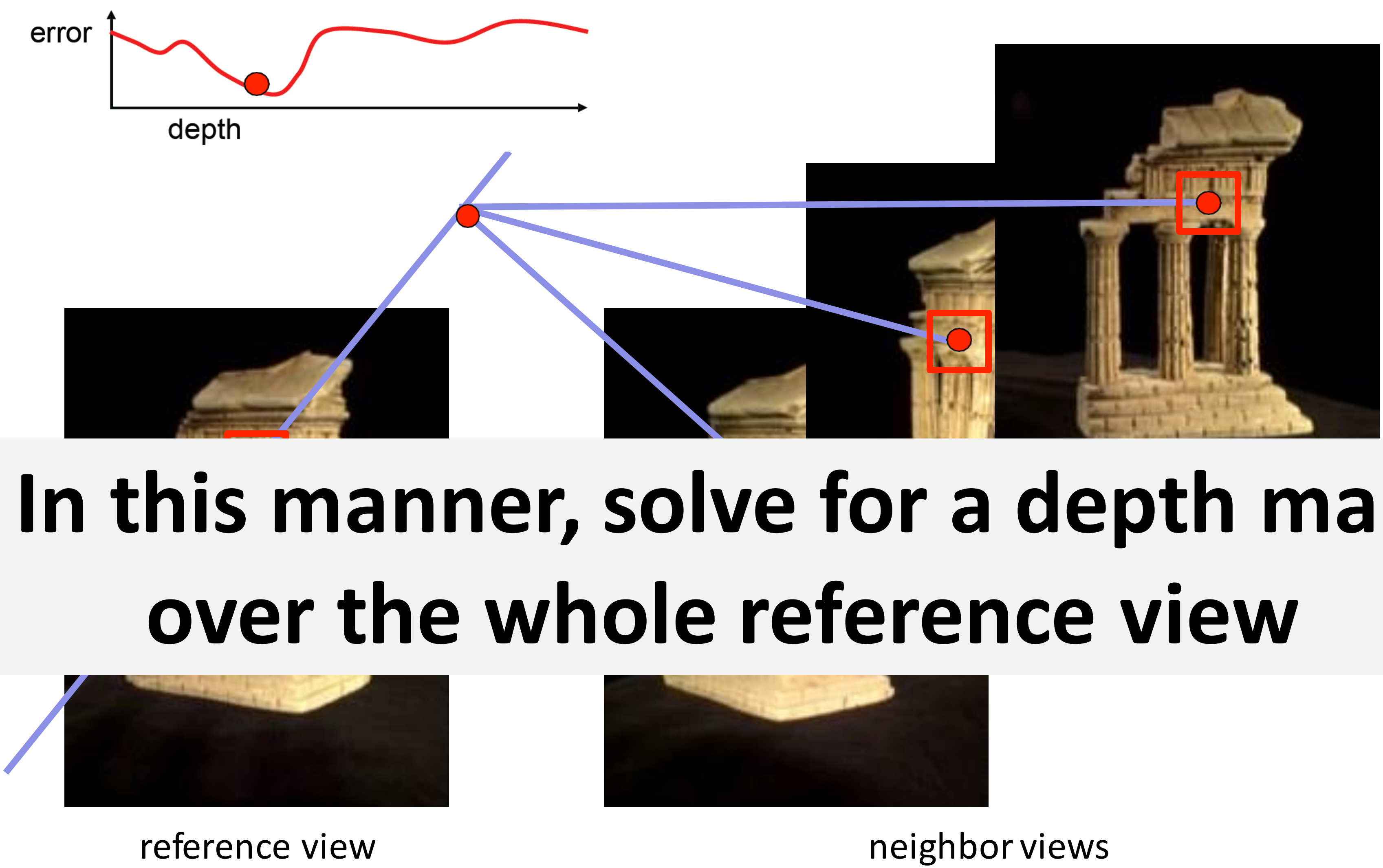


reference view

neighbor views

Source: Y. Furukawa

Multi-view stereo: Basic idea



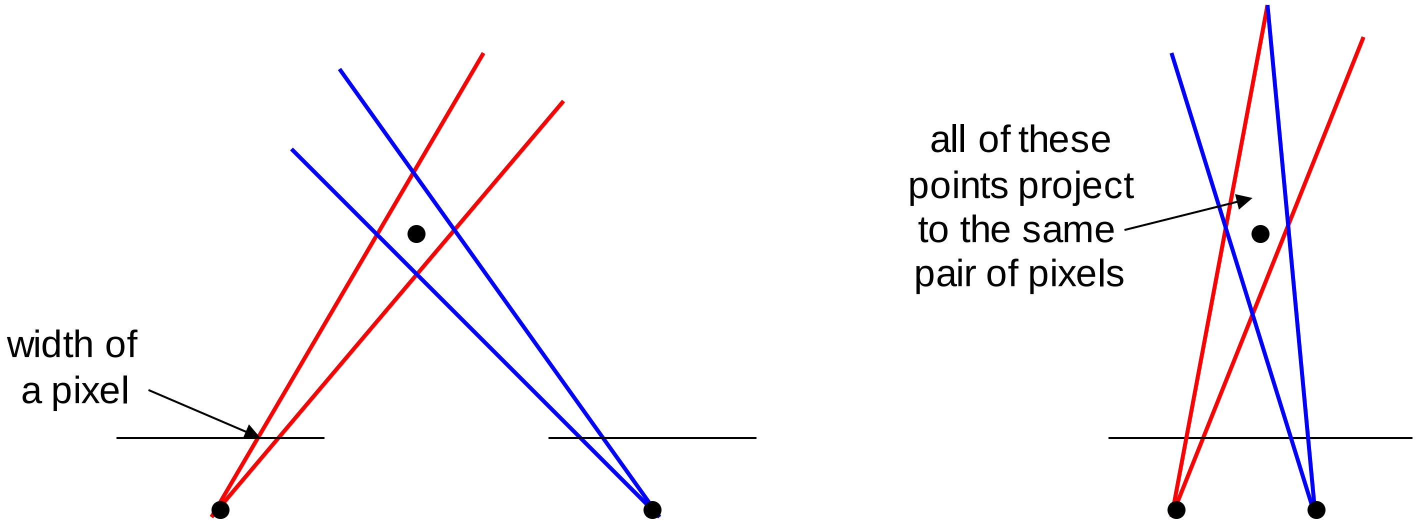
In this manner, solve for a depth map over the whole reference view

Source: Y. Furukawa

Multi-view stereo: advantages

- Can match windows using more than 1 neighbor, giving a **stronger match signal**
- If you have lots of potential neighbors, can **choose the best subset** of neighbors to match per reference image
- Can reconstruct a depth map for each reference frame, and the merge into a **complete 3D model**

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

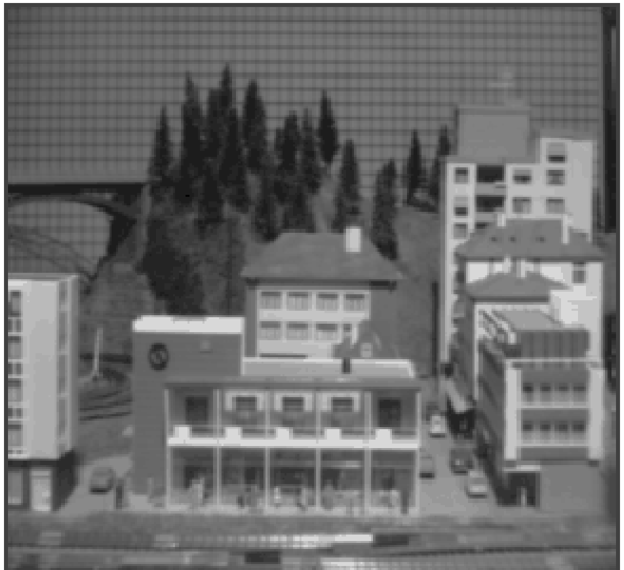
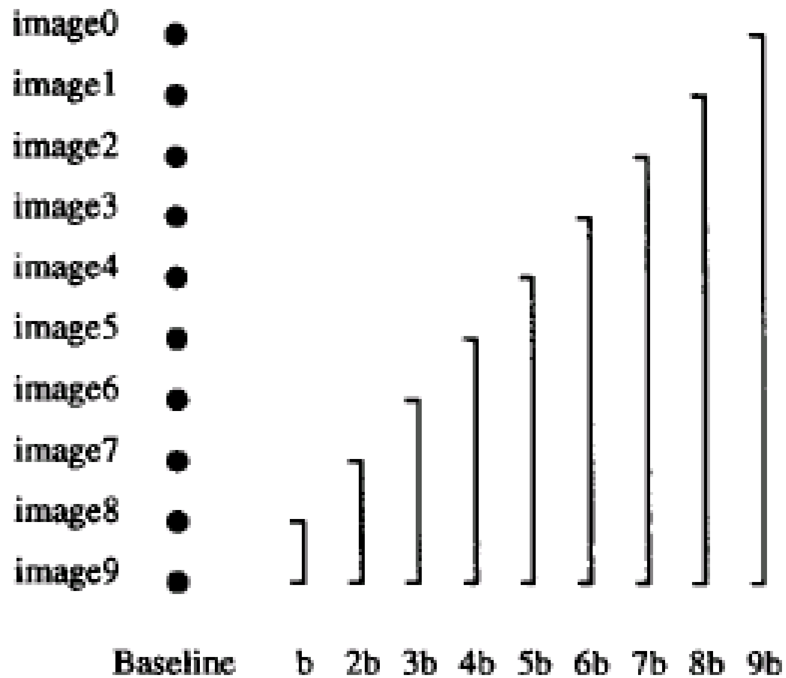
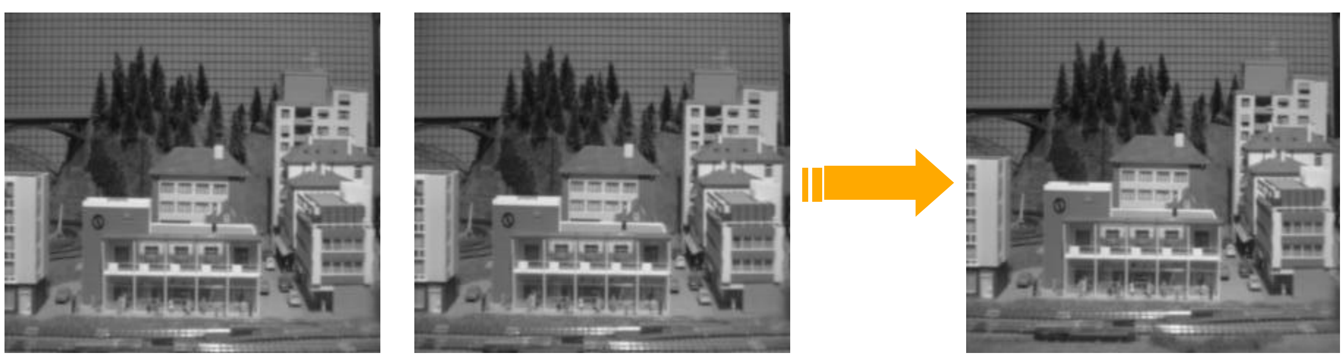


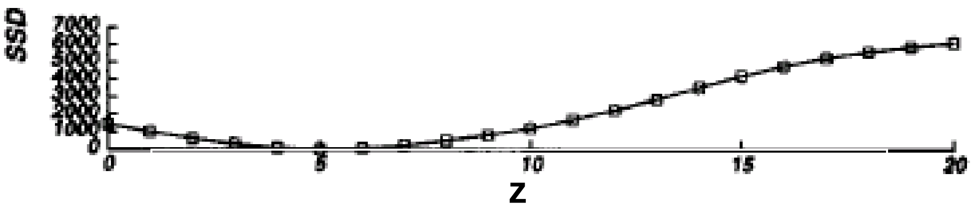
Figure 2: An example scene. The grid pattern in the background has ambiguity of matching.

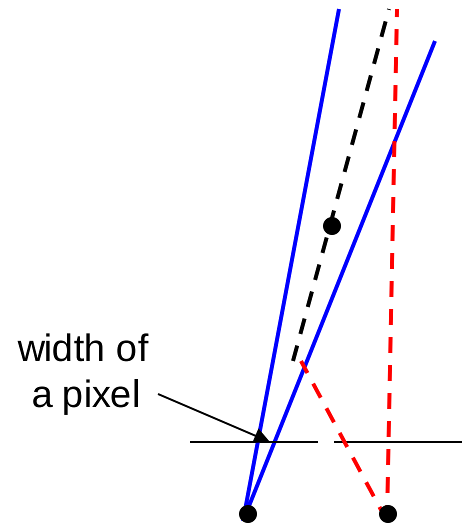




I_1 I_2 I_{10}

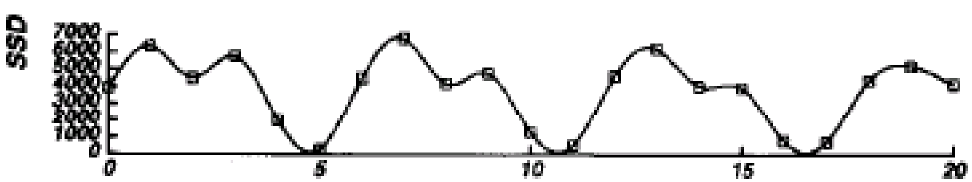
Multiple-baseline stereo

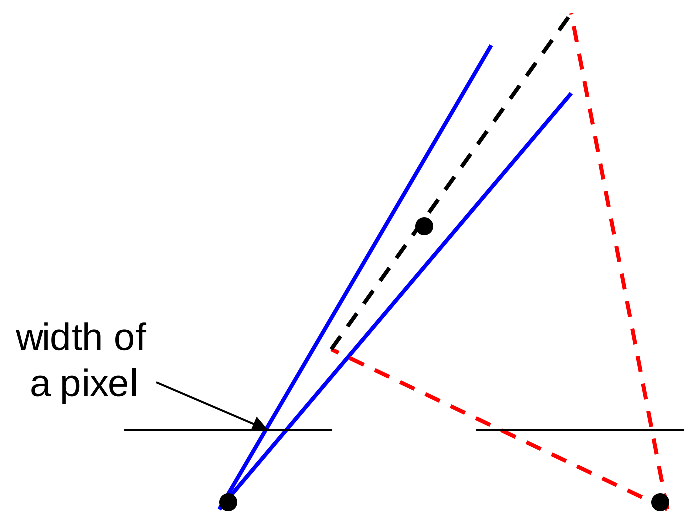




pixel matching score

- For short baselines, estimated depth will be less precise due to narrow triangulation





- For larger baselines, must search larger area in second image

M. Okutomi and T. Kanade, "A Multiple-Baseline Stereo System," IEEE Trans. on Pattern Analysis and Machine Intelligence, 15(4):353-363 (1993).

Multiple-baseline stereo

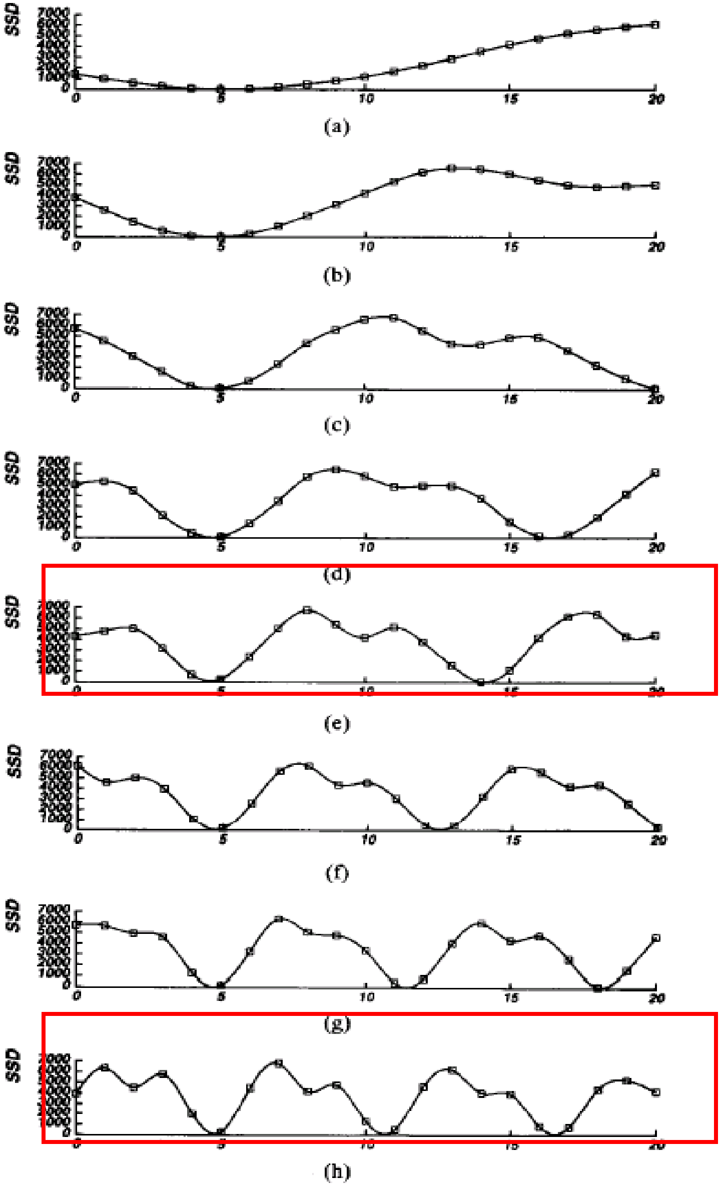


Fig. 5. SSD values versus inverse distance: (a) $B = b$; (b) $B = 2b$; (c) $B = 3b$; (d) $B = 4b$; (e) $B = 5b$; (f) $B = 6b$; (g) $B = 7b$; (h) $B = 8b$. The horizontal axis is normalized such that $8bF = 1$.

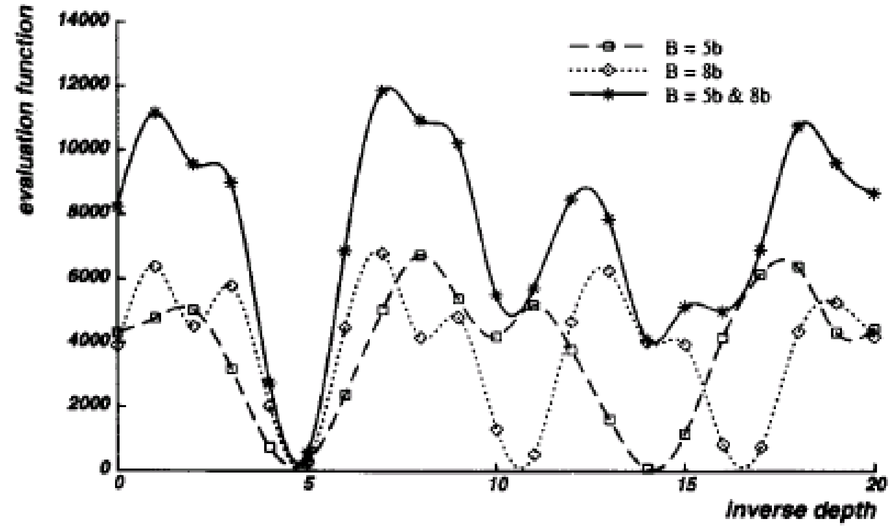


Fig. 6. Combining two stereo pairs with different baselines.

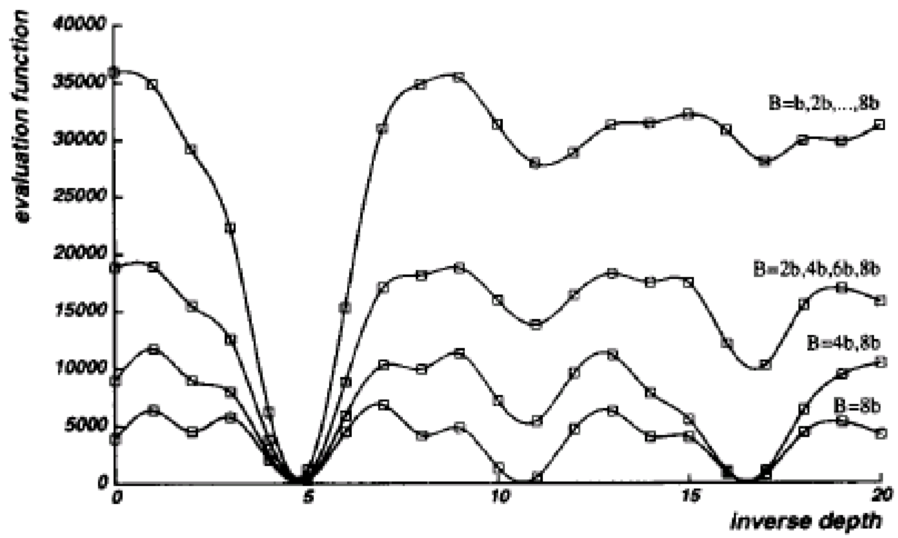
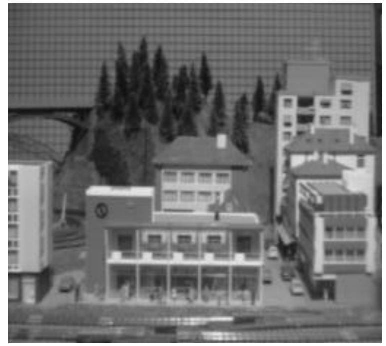


Fig. 7. Combining multiple baseline stereo pairs.

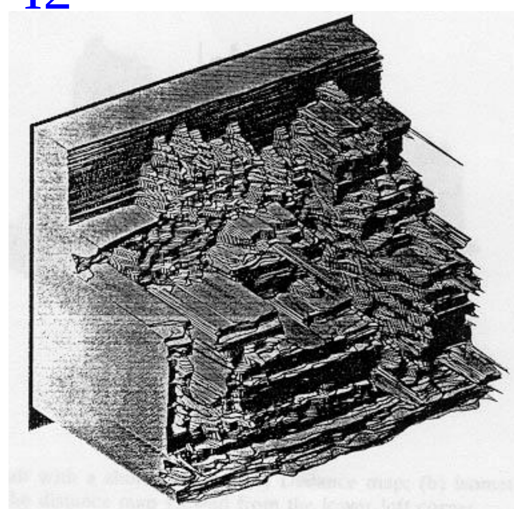
Multiple-baseline stereo results

I1



I2

I10



M. Okutomi and T. Kanade, *A Multiple-Baseline Stereo System,* IEEE Trans. on Pattern Analysis and Machine Intelligence, 15(4):353-363 (1993).

Multibaseline Stereo

Basic Approach

- Choose a reference view
- Use your favorite stereo algorithm BUT
  - replace two-view SSD with **SSSD** over all baselines
  - **SSSD**: the SSD values are computed first for each pair of stereo images, and then add all together from multiple stereo pairs.

Limitations

- Only gives a depth map (not an "object model")
- Won't work for widely distributed views.

Problem: *visibility*

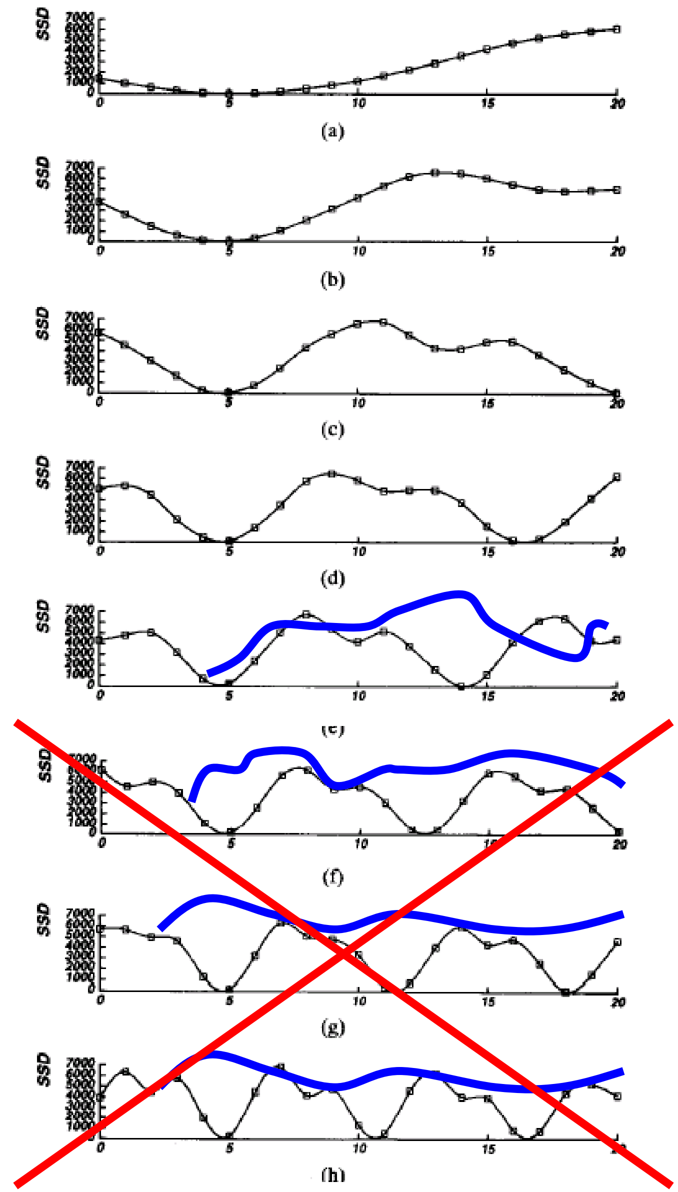


Fig. 5. SSD values versus inverse distance: (a) $B = b$; (b) $B = 2b$; (c) $B = 3b$; (d) $B = 4b$; (e) $B = 5b$; (f) $B = 6b$; (g) $B = 7b$; (h) $B = 8b$. The horizontal axis is normalized such that $8bF = 1$.

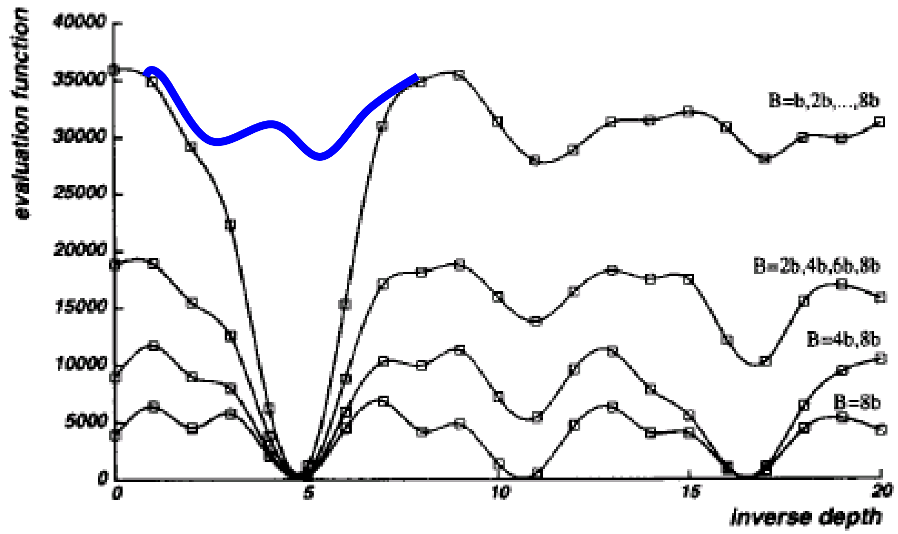


Fig. 7. Combining multiple baseline stereo pairs.

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)

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

- SAD (Sum of Absolute Difference) $\sum_{x,y} |W_1(x,y) - W_2(x,y)|$
- ZNCC (Zero-mean Normalized Cross Correlation)

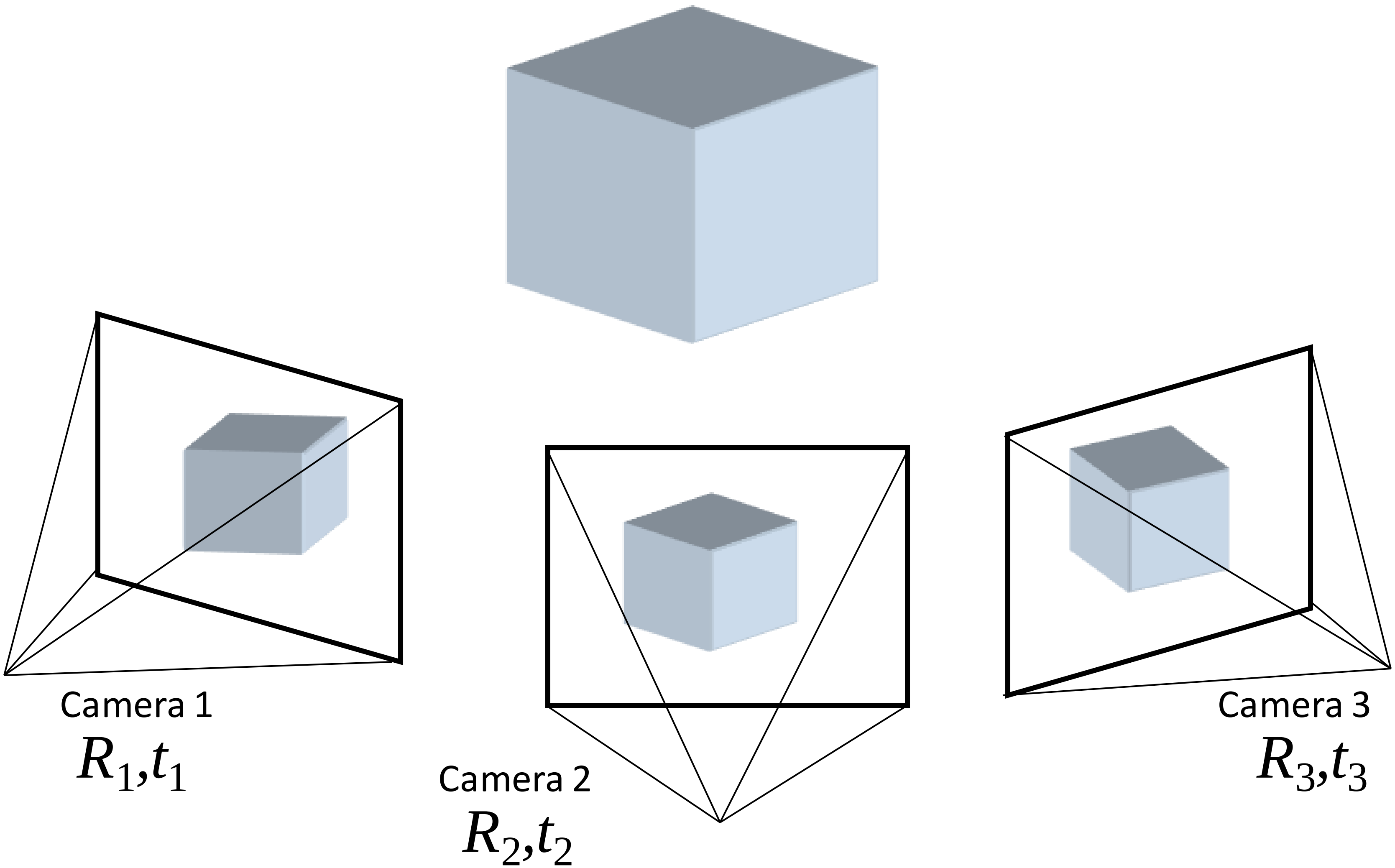
$$\frac{\sum_{x,y}(W_1(x,y) - \overline{W_1})(W_2(x,y) - \overline{W_2})}{\sigma_{W_1}\sigma_{W_2}}$$

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

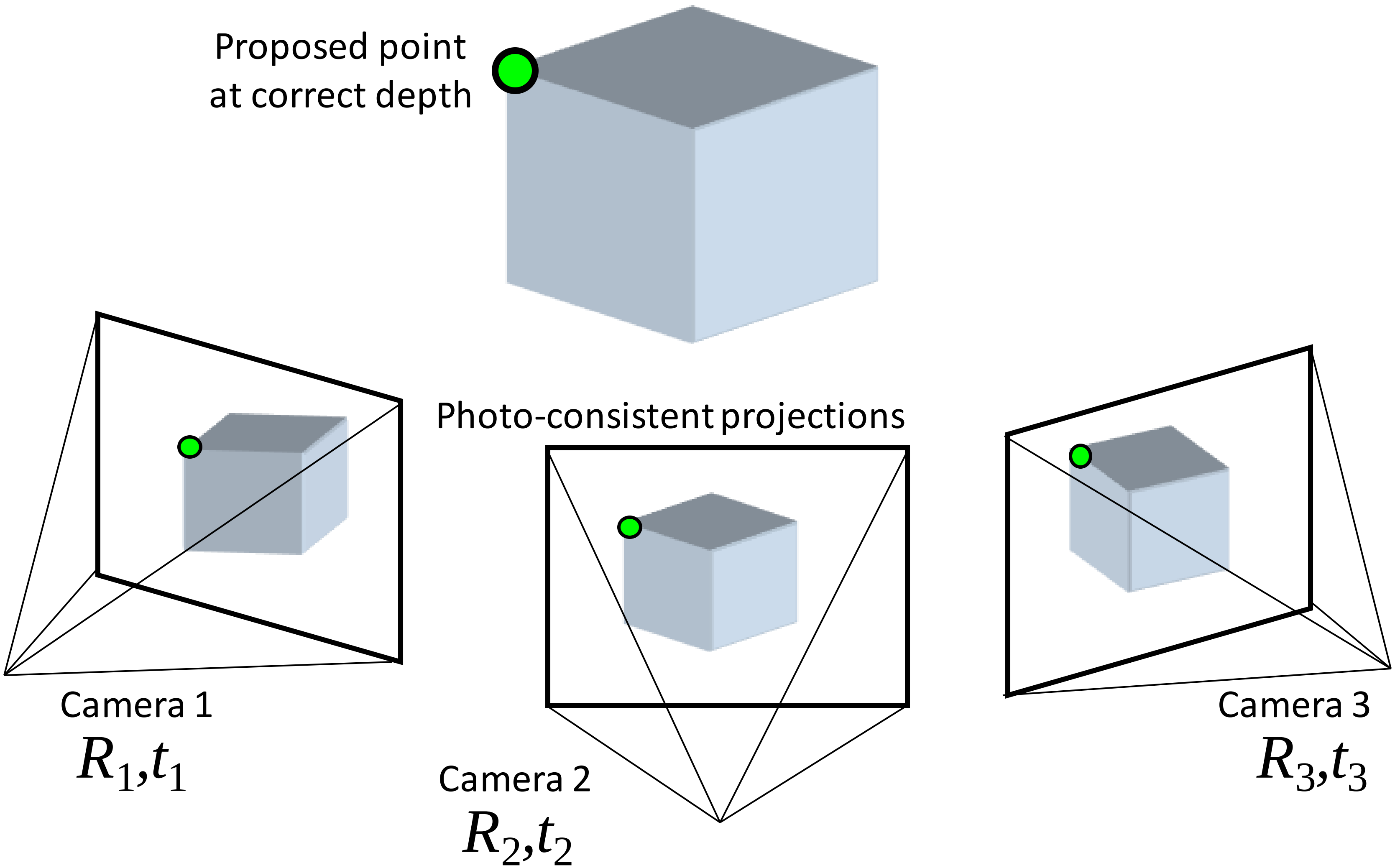
  – what advantages might NCC have?

Questions?

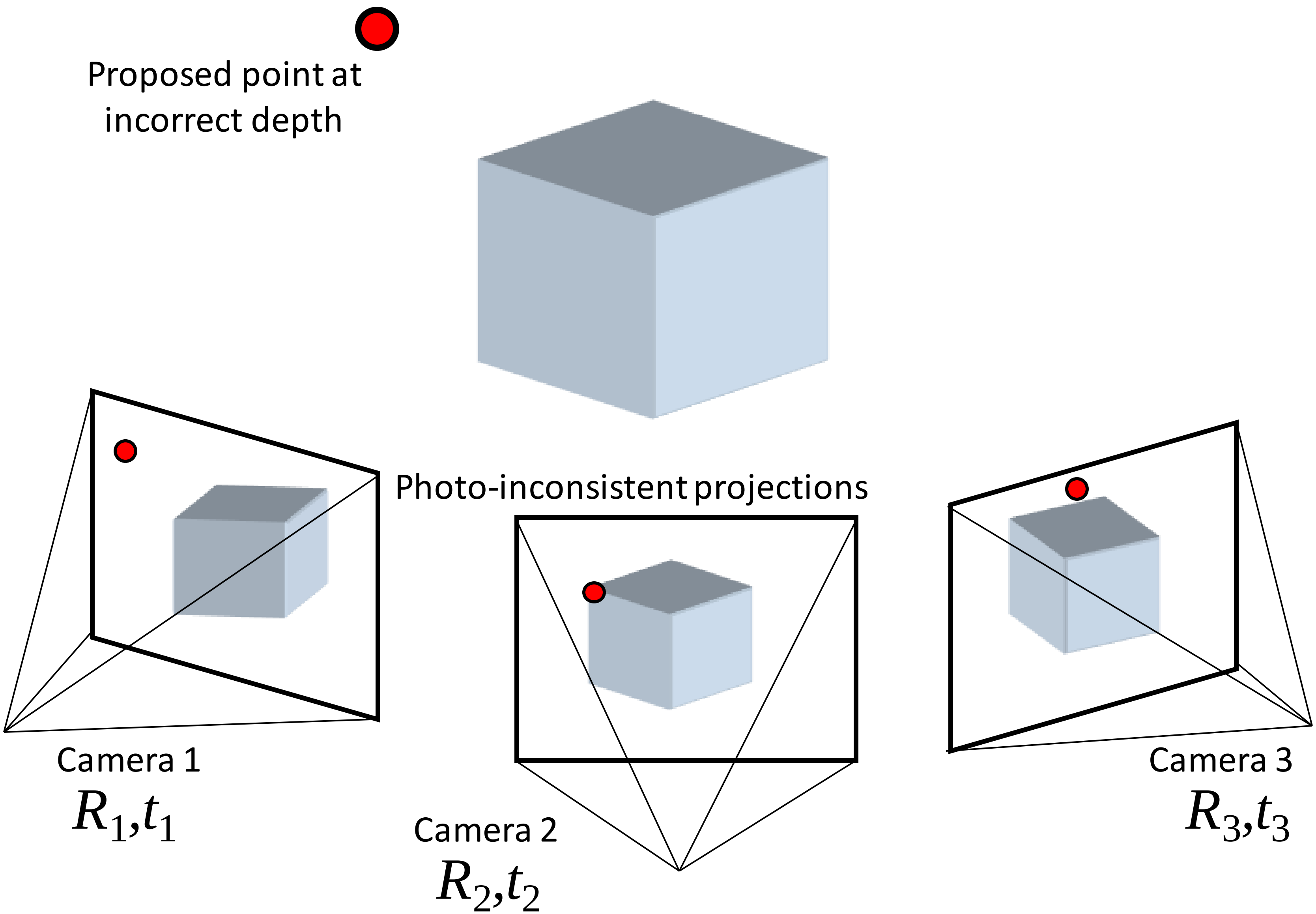
Plane-Sweep Stereo



Plane-Sweep Stereo

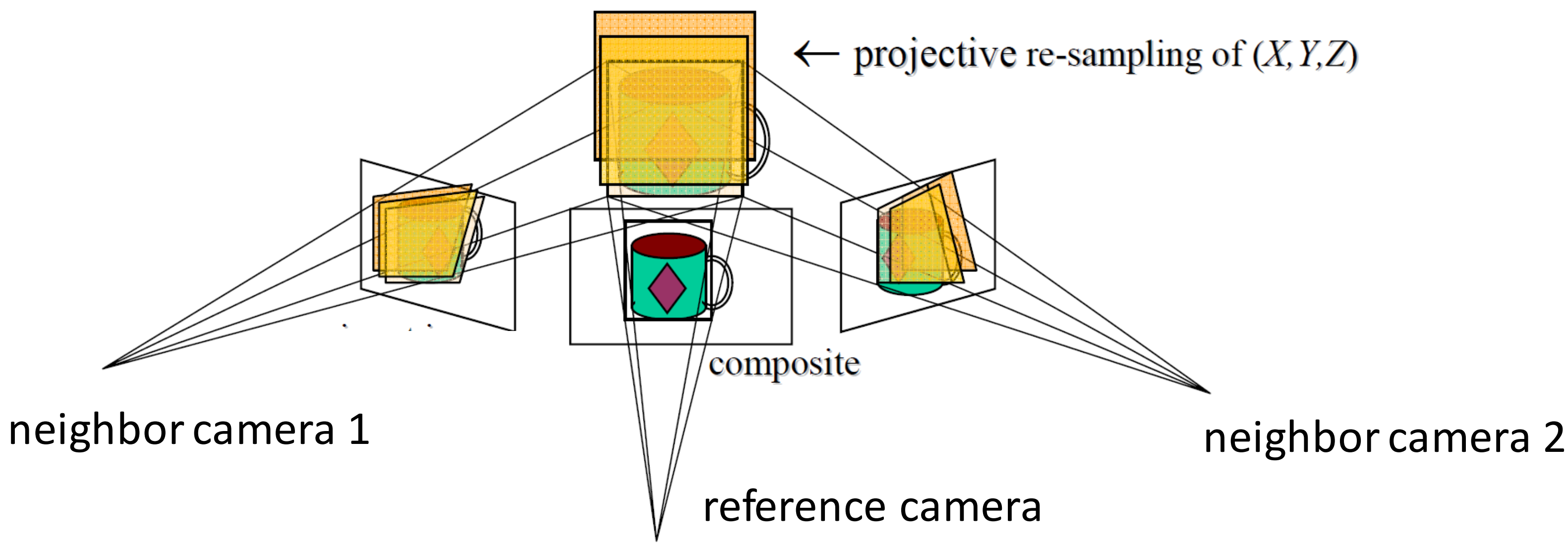


Plane-Sweep Stereo



Plane-Sweep Stereo

- Sweep family of planes parallel to the reference camera image plane
- Reproject neighbors onto each plane (via homography) and compare reprojections

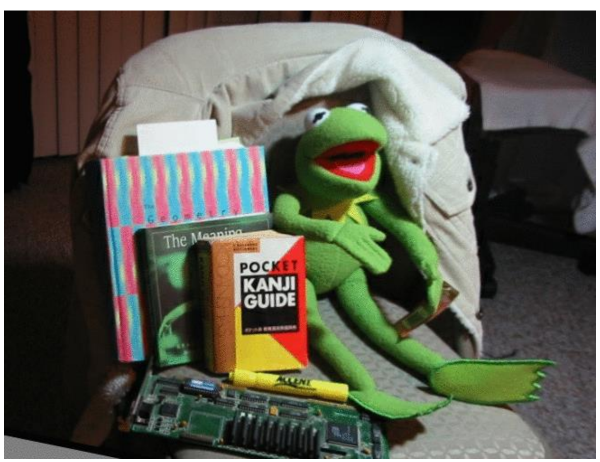


Plane-Sweep Stereo

Left neighbor

Reference image

Right neighbor



Left neighbor projected into reference image

Average images on each plane

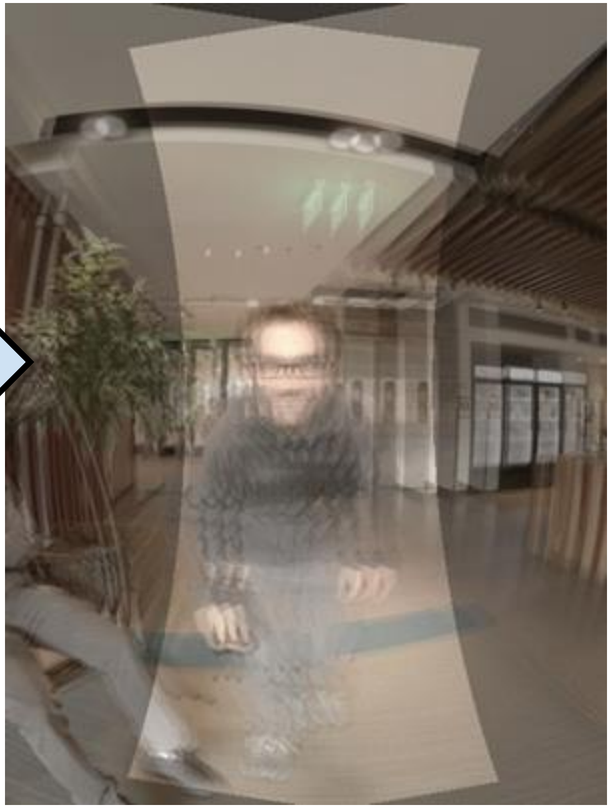
Right neighbor projected into reference image

Another example

Left neighbor

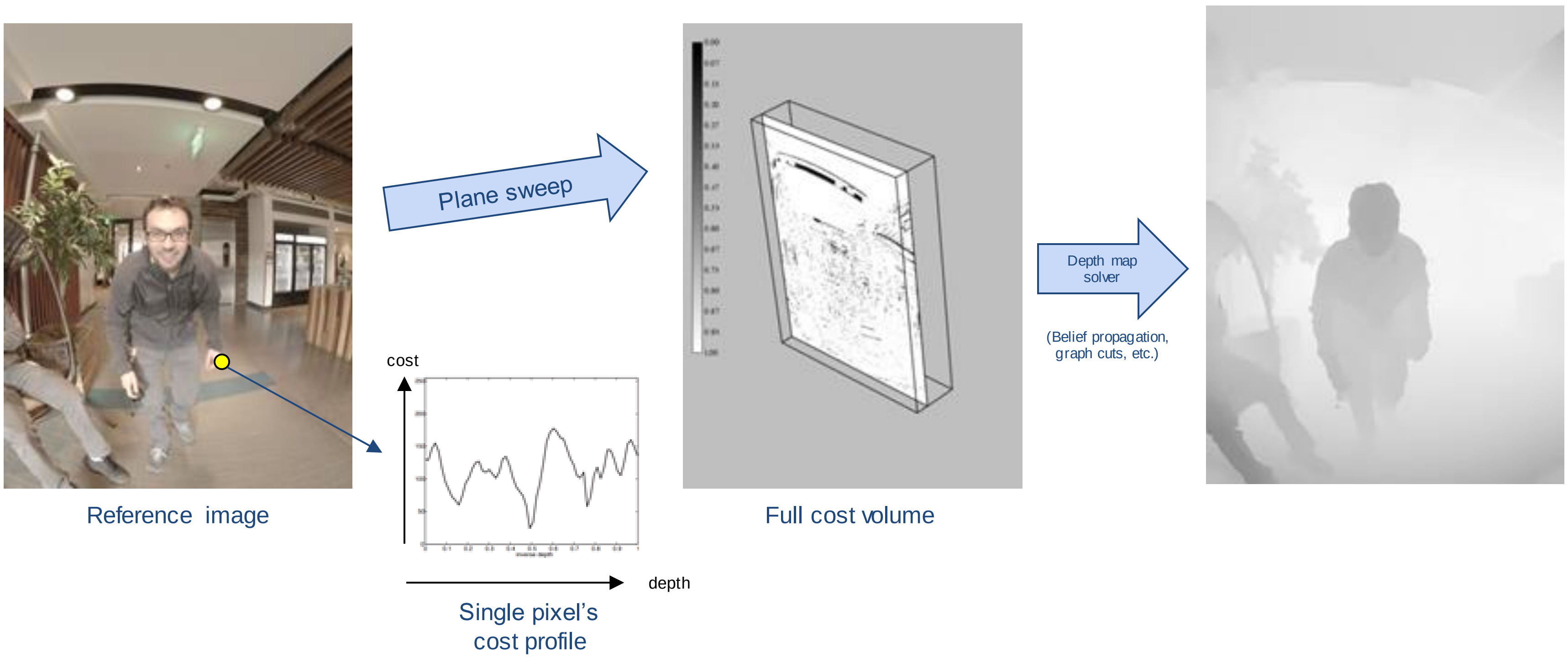
Reference image

Right neighbor



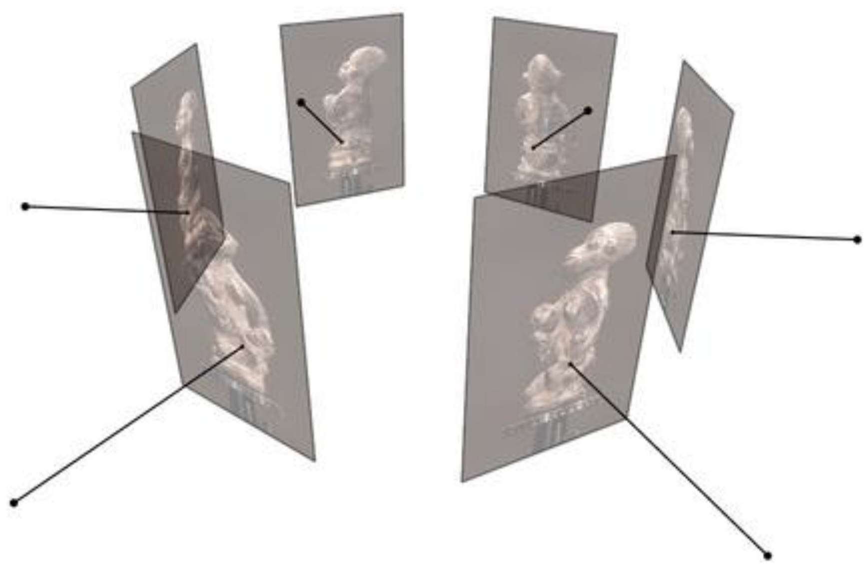
Planar image reprojections swept over depth (averaged)

Cost Volumes -> Depth Maps



Fusing multiple depth maps

- Compute depth map per image
- Fuse the depth maps into a 3D model



Figures by Carlos Hernandez

Questions?