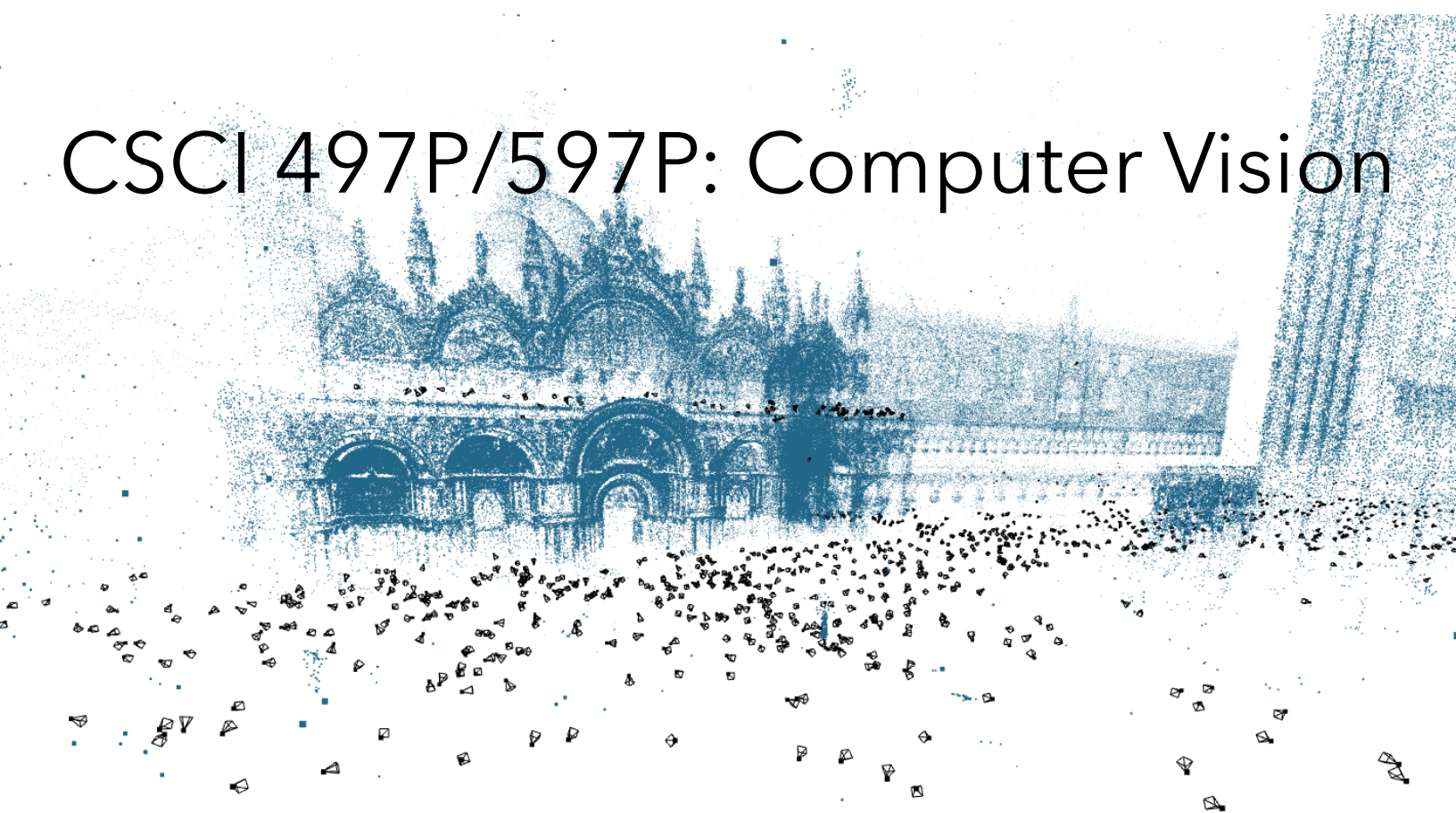
CSCI 497P/597P: Computer Vision



Lecture 25

Structure From Motion

Multiview Stereo

Introduction to Recognition

Announcements

- P3 due in 1 week
- HW4 due Friday

Goals

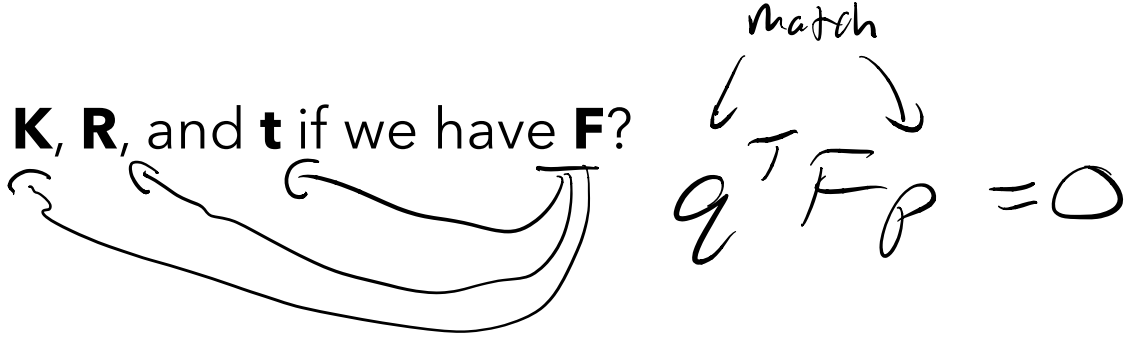
- Understand the Structure From Motion problem and the general idea behind how it is solved.
- Understand the Multiview Stereo problem and the general idea behind how it's solved.

Two questions:

- We derived **F** assuming that **K**, **R**, and **t** are known.
- Can we find **F** without them?

  yes - normalized 8-point algorithm, among others

- Can we find **K**, **R**, and **t** if we have **F**?



What about more than 2 views?

- 2 views: fundamental matrix
- 3 views: trifocal tensor
- 4 views: *quadrifocal* tensor
- more views: it's called **Structure From Motion**

Structure from Motion

- Given many photos, reconstruct:
  - positions of the cameras
  - positions of 3D points

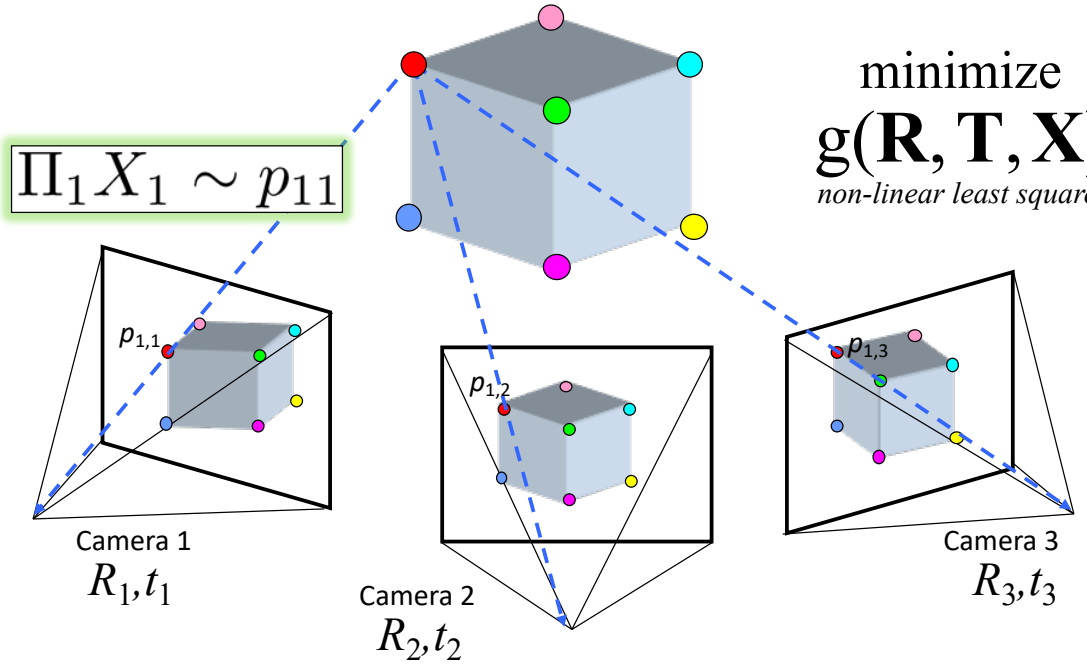
Large-scale structure from motion

- https://www.youtube.com/watch?v=sQegEro5Bfo

Chicken/Egg

- Step 1: solve for relative pose of pairs (or triples) of cameras using correspondences from feature matching.
- Step 2: alternate between solving:
  - given camera positions, solve for point locations
  - given point locations, solve for camera positions

Structure From Motion



$$g(\mathbf{X}, \mathbf{R}, \mathbf{T}) = \sum_{i=1}^{m} \sum_{j=1}^{n} w_{ij} \cdot \left\| \underbrace{\mathbf{P}(\mathbf{x}_i, \mathbf{R}_j, \mathbf{t}_j)}_{\text{predicted image location}} - \underbrace{\begin{bmatrix} u_{i,j} \\ v_{i,j} \end{bmatrix}}_{\text{observed image location}} \right\|^2$$

w_{ij}: *indicator variable*: is point *i* visible in image *j* ?

Applications

- Hyperlapse https://www.youtube.com/watch?v=SOpwHaQnRSY
- SLAM: https://medium.com/scape-technologies/building-the-ar-cloud-part-three-3d-maps-the-digital-scaffolding-of-the-21st-century-465fa55782dd
- Graphics, movies, games, self-driving cars, robots, ...

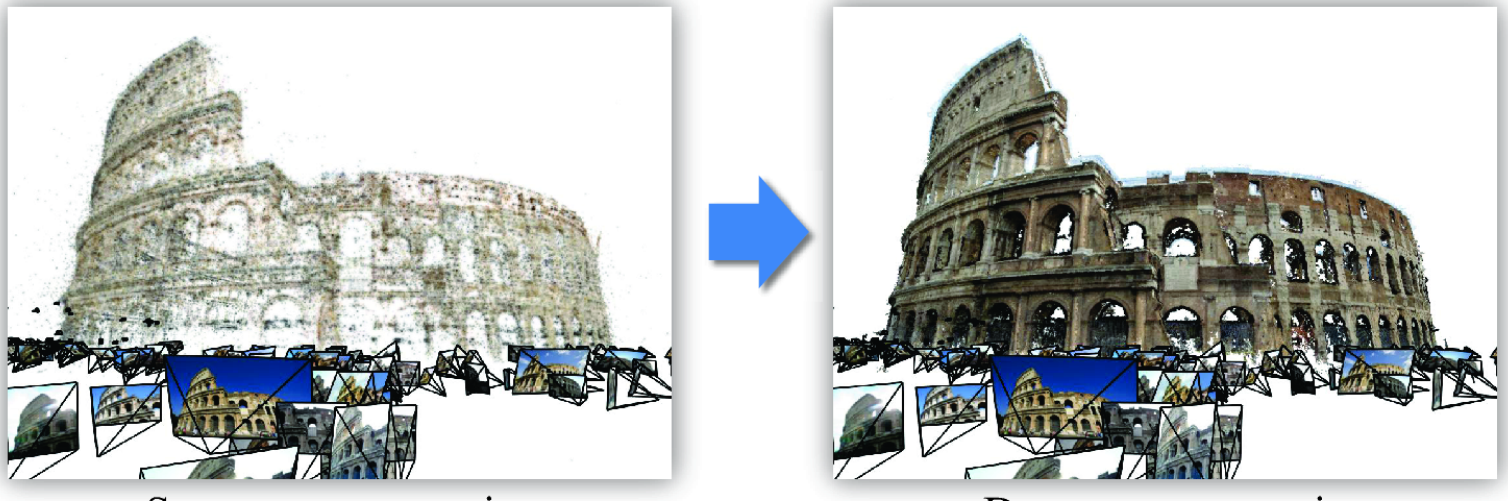
SfMFlex

Visualizing Spectral Bundle Adjustment Uncertainty
Kyle Wilson and Scott Wehrwein, 3DV 2020.

Project webpage: https://wilsonkl.github.io/sfmflex-release/

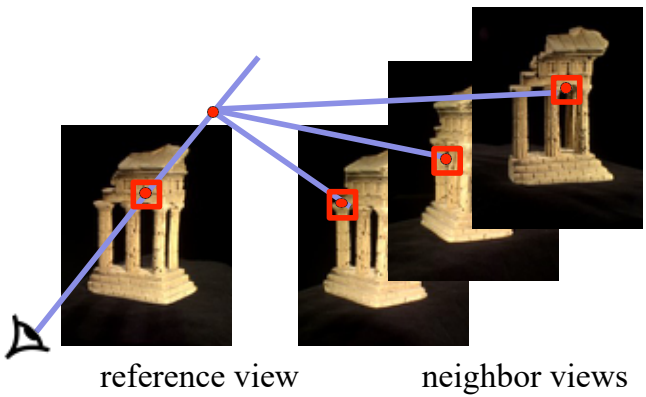
Multiview Stereo

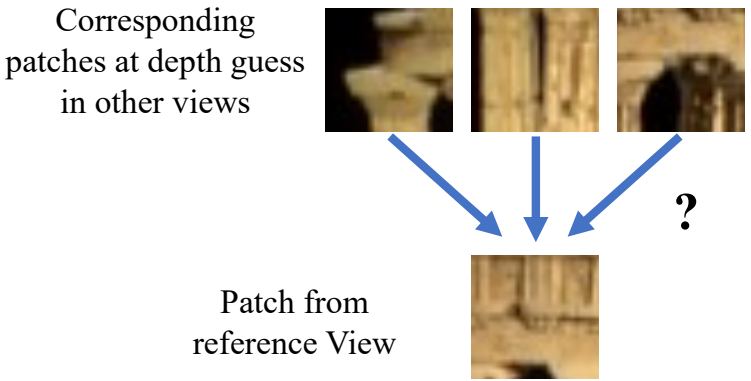
- Once you've solved for all those camera positions, how good a 3D model can you create?



Multiview Stereo: Basic Idea

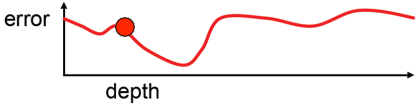
Evaluate the likelihood of geometry at a particular depth for a particular reference patch:





Source: Y. Furukawa

Photometric error across different depths

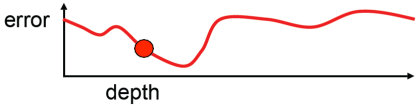


reference view

neighbor views

Source: Y. Furukawa

Photometric error across different depths

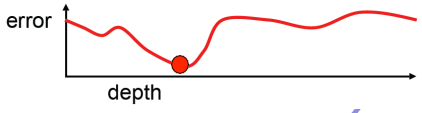


reference view

neighbor views

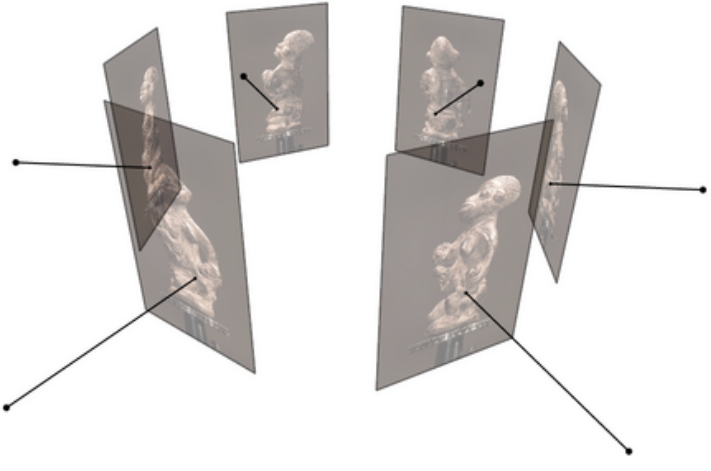
Source: Y. Furukawa

Photometric error across different depths



Depth map fusion

- Compute depth maps for multiple cameras, then fuse them into a 3D model



Figures by Carlos Hernandez

Result

- 3D Google Earth: https://www.youtube.com/watch?v=N6Douyfa7l8