CS231A Midterm Review

May 19, 2017

Midterm Logistics

- In-class midterm at Skilling Auditorium at 3:00-4:20 PM on May 22, 2017
- SCPD students not taking exam at Stanford should coordinate with SCPD.
- Open book and open notes. Not open computer.
- Lectures 1 - 12 (through Image Classification & 2D Object Detection)
- 10 T/F, 10 MC, 3 longer questions (similar to last year’s)

Topics you should know for the Exam

- General knowledge of linear algebra (matrix multiplication, SVD, etc)
- Camera Models and Transformations
- Non-perspective Cameras
- Camera Calibration
- Single View Metrology
- Epipolar Geometry
- Structure from Motion
- Active Stereo and Volumetric Stereo
- Fitting and Matching
- RANSAC
- Hough Transform
- Detectors and Descriptors
- Image Classification
- 2D Object Detection

Homogeneous Coordinates

- Homogenous coordinates allow us to apply a larger variety of transformations with matrix multiplication
  - For example, we use homogeneous coordinates to handle the 3D -> 2D projection
- Any point (x, y) becomes represented as (x, y, 1)
- More generally $(a_1, a_2, \ldots, a_n, w)$ represents the point $(a_1/w, a_2/w, \ldots, a_n/w)$

Types of transformations

- Isometric transformations preserve distances
  - Rotation, translation, reflection

- Similarity transformations preserve shape
  - Rotation, translation, scaling

- Affine transformations preserve parallelism
  - Rotation, translation, scaling, shearing, etc.
  - $T(v) = Av + t$, A is invertible

- Projective transformations map lines to lines
  - Pretty much everything else

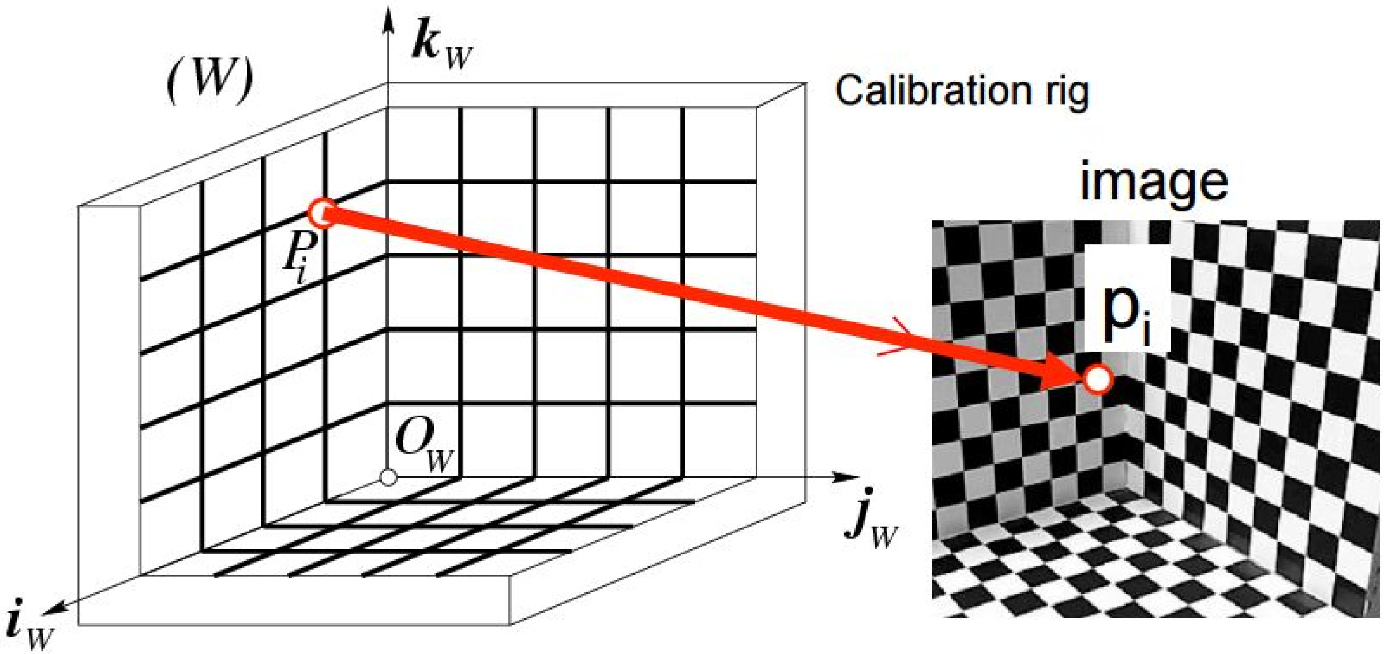
Camera Parameters

- Extrinsic parameters
  - Rotation and translation from the world frame
- Intrinsic parameters
  - Focal length in x and y direction, camera center offset, skew, distortion
  - Most people assume only 5 parameters (for the sake of this class)

$$K = \begin{bmatrix} \alpha & -\alpha \cot\theta & c_x \\ 0 & \frac{\beta}{\sin\theta} & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$P' = K \begin{bmatrix} R & T \end{bmatrix} P_w = M P_w$$

Camera Calibration



$$p_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} \dfrac{\mathbf{m}_1 \, \mathrm{P}_i}{\mathbf{m}_3 \, \mathrm{P}_i} \\ \dfrac{\mathbf{m}_2 \, \mathrm{P}_i}{\mathbf{m}_3 \, \mathrm{P}_i} \end{bmatrix} = M \; P_i \qquad \mathrm{M} = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{bmatrix}$$

in pixels

Camera Calibration

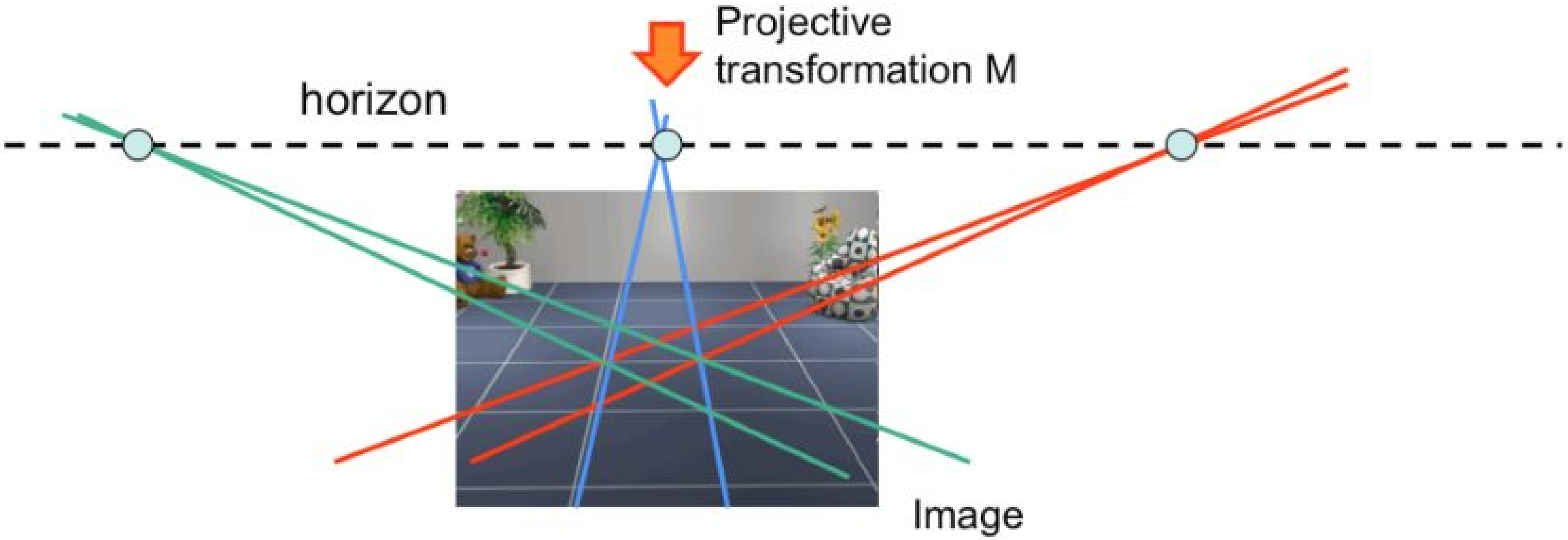
$$u_1(m_3P_1) - m_1P_1 = 0$$
$$v_1(m_3P_1) - m_2P_1 = 0$$
$$\vdots$$
$$u_n(m_3P_n) - m_1P_n = 0$$
$$v_n(m_3P_n) - m_2P_n = 0$$

$$\begin{bmatrix} P_1^T & 0^T & -u_1P_1^T \\ 0^T & P_1^T & -v_1P_1^T \\ & \vdots & \\ P_1^T & 0^T & -u_1P_1^T \\ 0^T & P_1^T & -v_1P_1^T \end{bmatrix} \begin{bmatrix} m_1^T \\ m_2^T \\ m_3^T \end{bmatrix} = \mathbf{P}m = 0$$

Solve by SVD!

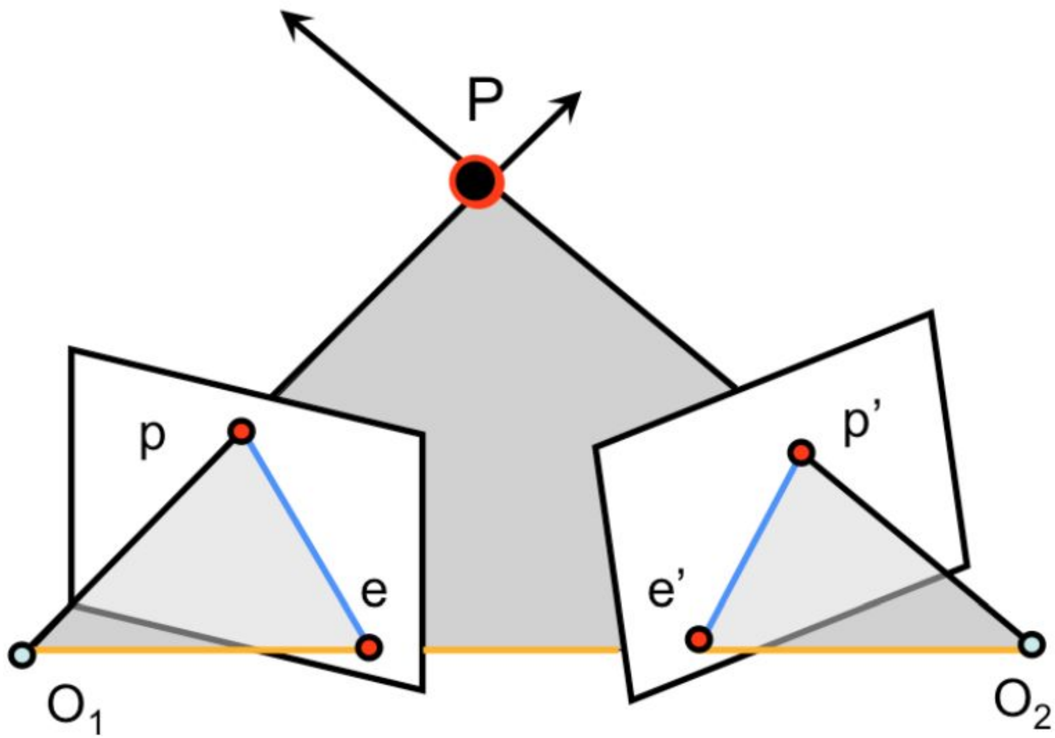
Single View Metrology

- Vanishing points and vanishing lines (horizon)



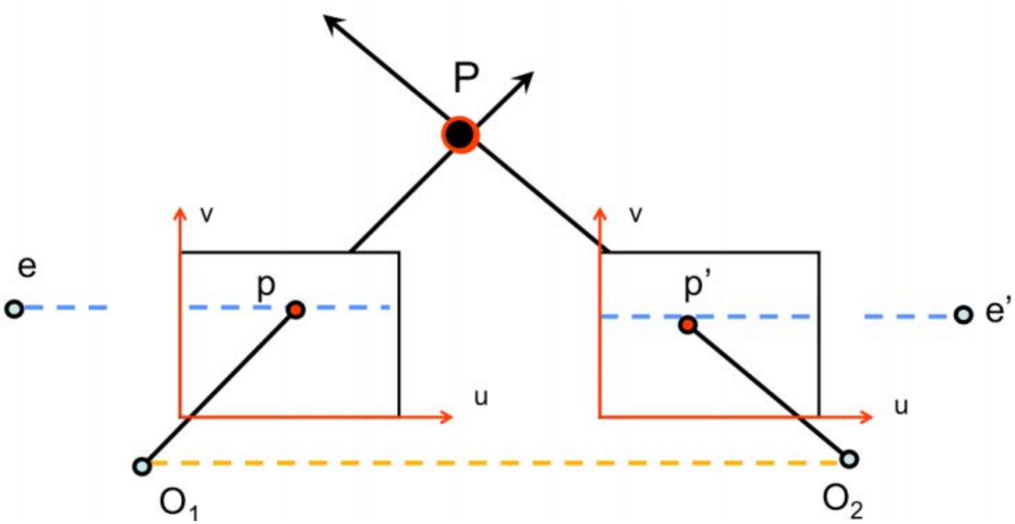
- This leads to being able to find angles between lines and planes (recall PS1)
- You can also calibrate the camera from a single image!

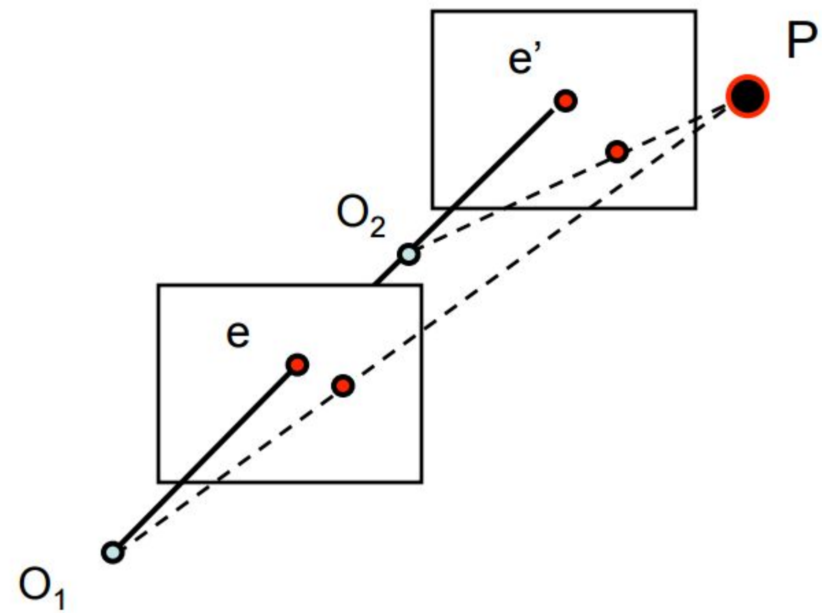
Epipolar Geometry



- Understanding the geometry of the scene and the cameras
- Should have knowledge of this entire scene and basic triangulation
  - Epipoles, epipolar lines, reprojection error, etc.

Unique Cases of Epipolar Geometry





- Parallel cameras make the epipoles at infinity
- Forward translation make the epipoles in the same location

The Fundamental Matrix

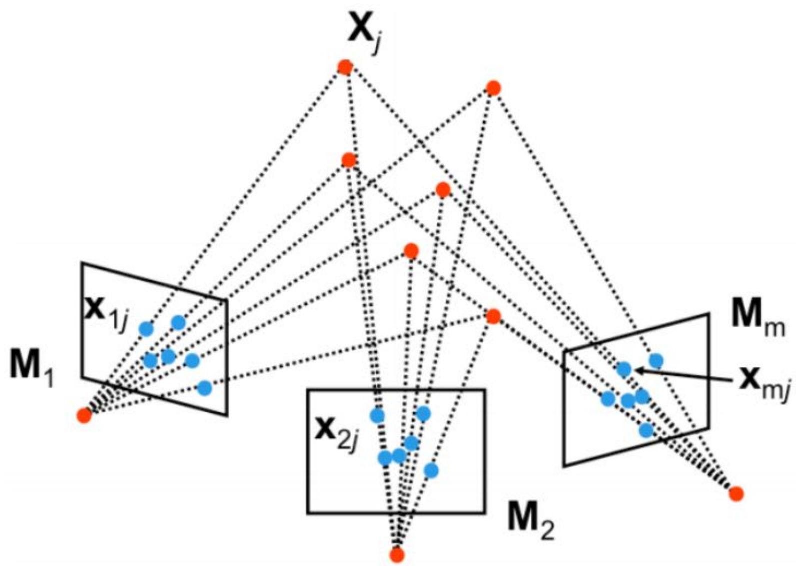
$$p^T F p' = 0 \qquad\qquad F = K^{-T} \cdot [T_\times] \cdot R K'^{-1}$$

- Relates corresponding points with a single constraint
- 7 degrees of freedom
- Can be found using Eight Point Algorithm and Normalized Eight-Point algorithm

$$\begin{bmatrix} u_1u_1' & v_1u_1' & u_1' & u_1v_1' & v_1v_1' & v_1' & u_1 & v_1 & 1 \\ u_2u_2' & v_2u_2' & u_2' & u_2v_2' & v_2v_2' & v_2' & u_2 & v_2 & 1 \\ u_3u_3' & v_3u_3' & u_3' & u_3v_3' & v_3v_3' & v_3' & u_3 & v_3 & 1 \\ u_4u_4' & v_4u_4' & u_4' & u_4v_4' & v_4v_4' & v_4' & u_4 & v_4 & 1 \\ u_5u_5' & v_5u_5' & u_5' & u_5v_5' & v_5v_5' & v_5' & u_5 & v_5 & 1 \\ u_6u_6' & v_6u_6' & u_6' & u_6v_6' & v_6v_6' & v_6' & u_6 & v_6 & 1 \\ u_7u_7' & v_7u_7' & u_7' & u_7v_7' & v_7v_7' & v_7' & u_7 & v_7 & 1 \\ u_8u_8' & v_8u_8' & u_8' & u_8v_8' & v_8v_8' & v_8' & u_8 & v_8 & 1 \end{bmatrix} \begin{bmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{bmatrix} = 0$$

Solve by SVD!

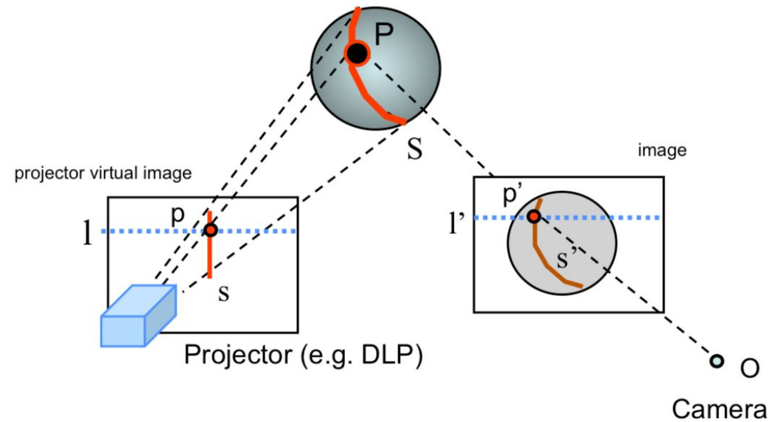
Structure from Motion



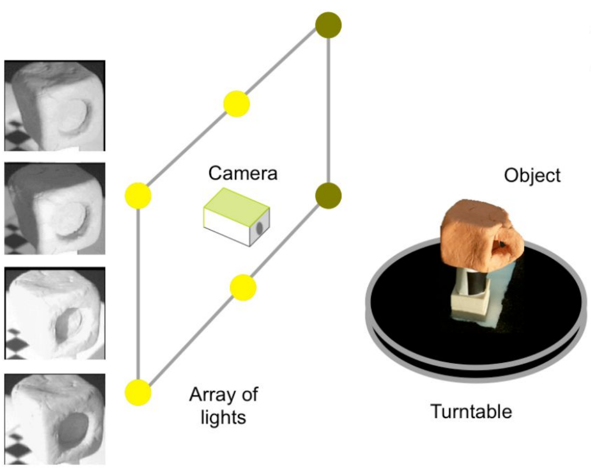
- Estimating both the camera positions and the 3D structure simultaneously from point correspondences
- You've implemented a few algorithms:
  - Factorization method
  - An iterative triangulation method

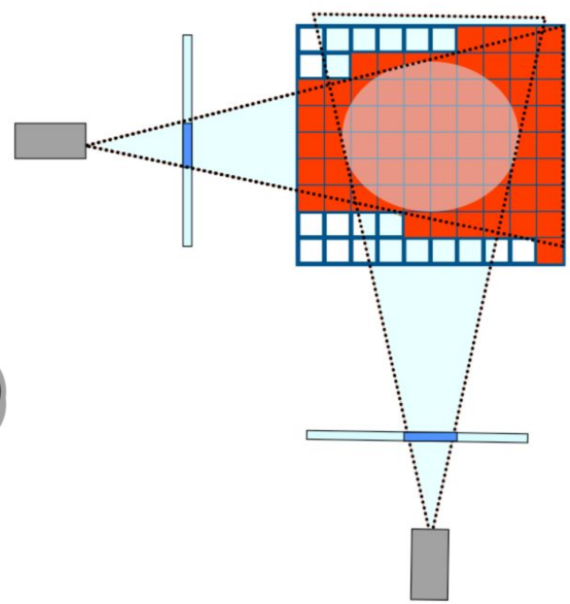
Active Stereo and Volumetric Stereo

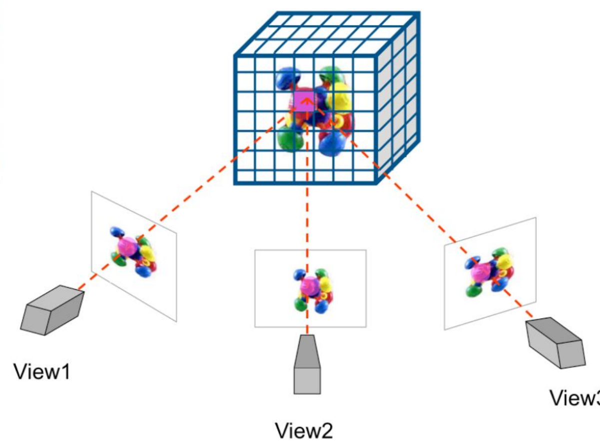
- Active Stereo
  - Replaces one camera with a projector



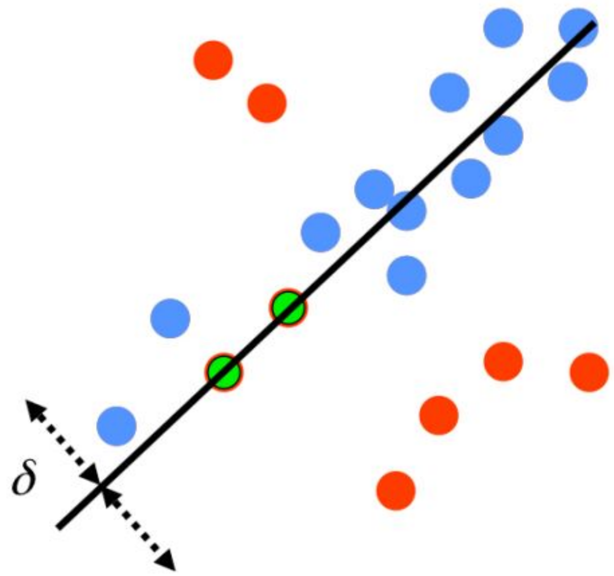
- Volumetric Stereo
  - Space carving
  - Shadow carving
  - Voxel coloring





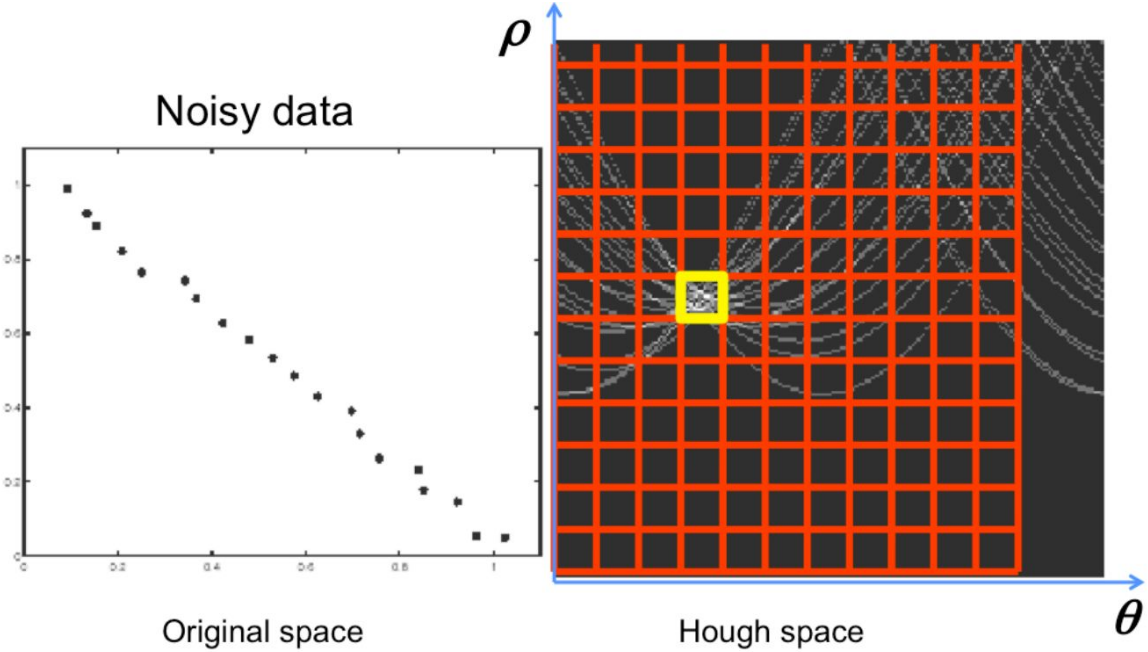


RANSAC



- Select random sample of minimum size
- Compute a model from this
- Compute the inliers within the model
- Repeat steps for a fixed amount and return the model with the most inliers

Hough Transform



- Find some parameter space that defines the line, plane, etc. that we're trying to estimate
- For each observation, plot in this parameter space
  - Could be points, lines, hyperplanes, etc.
- Grid up the parameter space and find cells with many observations
- If there is a problem, will be similar to the example found in lecture

Detectors and Descriptors

- Corner detectors
  - Harris corner detector
- Edge detectors
  - Find areas of high gradients, but should smooth before doing so to remove noise
  - Should know about Laplacian of Gaussian and Difference of Gaussian
- Blob detection
  - Similar to edge detection, but in 2D
- SIFT
  - A local descriptor around keypoints based on gradients in the image
  - Scale and in-plane rotation invariant
- HOG
  - Implemented in PS3 - you should know about it!

Image Classification and 2D Object Detection

- This questions on the midterm will be purely conceptual - nothing too difficult
- Bag of words
  - Histogram representation of “words” (features)
- Part of PS4 (don’t need to implement, but skimming the ideas in preparation for the midterm is a good idea)
  - Sliding window detectors
  - Non-maximal suppression

Exam Advice

- When studying, make a cheat sheet to quickly reference at exam time.
  - Since you have 80 minutes, you do not want to sift through pages of notes.
- You will not need to know the complex math derivations involved in the course.
  - There are some linear algebra problems, though!
- You will need to know how generally things work and explain them (camera matrices, SFM, RANSAC, Hough transforms, etc.).
- This review session is not comprehensive of all material on the exam!