Midterm Logistics

- In-class midterm at Skilling Auditorium at 1:30-2:50 PM on Monday.
- SCPD students not taking exam at Stanford should coordinate with SCPD. Let us know if you are coming to stanford so we can accommodate.
- Open book and open notes. Not open computer.
- Lectures 1 - 12 (through Image Classification & 2D Object Detection)
- 10 T/F, 10 MC, 5 short questions, 4 longer questions.
- Bring a calculator to solve numerical questions.
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

- Homogeneous 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

Examples in class are 2D, need to know how to generalize to 3D
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)

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

\[
P' = K \begin{bmatrix} R & T \end{bmatrix} P_w = MP_w
\]
Camera Calibration

\[ p_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} \frac{m_1}{m_i} P_i \\ \frac{m_2}{m_i} P_i \\ \frac{m_3}{m_i} P_i \end{bmatrix} = M \cdot P_i \]

\[ M = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} \]
Camera Calibration

\[
\begin{align*}
u_1(m_3 P_1) - m_1 P_1 &= 0 \\
v_1(m_3 P_1) - m_2 P_1 &= 0 \\
& \vdots \\
u_n(m_3 P_n) - m_1 P_n &= 0 \\
v_n(m_3 P_n) - m_2 P_n &= 0
\end{align*}
\]

\[
\begin{bmatrix}
P_1^T & 0^T & -u_1 P_1^T \\
0^T & P_1^T & -v_1 P_1^T \\
P_1^T & 0^T & -u_1 P_1^T \\
0^T & P_1^T & -v_1 P_1^T
\end{bmatrix}
\begin{bmatrix}
m_1^T \\
m_2^T \\
m_3^T
\end{bmatrix} = \mathbf{P} \mathbf{m} = 0
\]

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!

\[
\cos \theta = \frac{V_1^T \omega V_2}{\sqrt{V_1^T \omega V_1} \sqrt{V_2^T \omega V_2}}
\]

\[\omega = (K K^T)^{-1}\]

[Eq. 28]
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 makes the epipoles in the same location
The Fundamental Matrix

\[ p^T F p' = 0 \]

\[ F = K^{-T} \cdot [T_x] \cdot R \cdot K'^{-1} \]

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

\[
\begin{bmatrix}
  u_1 u'_1 & v_1 u'_1 & u'_1 & u_1 v'_1 & v'_1 & u_1 & v_1 & 1 \\
  u_2 u'_2 & v_2 u'_2 & u'_2 & u_2 v'_2 & v'_2 & u_2 & v_2 & 1 \\
  u_3 u'_3 & v_3 u'_3 & u'_3 & u_3 v'_3 & v'_3 & u_3 & v_3 & 1 \\
  u_4 u'_4 & v_4 u'_4 & u'_4 & u_4 v'_4 & v'_4 & u_4 & v_4 & 1 \\
  u_5 u'_5 & v_5 u'_5 & u'_5 & u_5 v'_5 & v'_5 & u_5 & v_5 & 1 \\
  u_6 u'_6 & v_6 u'_6 & u'_6 & u_6 v'_6 & v'_6 & u_6 & v_6 & 1 \\
  u_7 u'_7 & v_7 u'_7 & u'_7 & u_7 v'_7 & v'_7 & u_7 & v_7 & 1 \\
  u_8 u'_8 & v_8 u'_8 & u'_8 & u_8 v'_8 & v'_8 & u_8 & v_8 & 1 \\
\end{bmatrix}
\begin{bmatrix}
  F_{11} \\
  F_{12} \\
  F_{13} \\
  F_{21} \\
  F_{22} \\
  F_{23} \\
  F_{31} \\
  F_{32} \\
  F_{33} \\
\end{bmatrix} = 0
\]

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
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
  - Steps to calculate SIFT descriptor

- **HOG**
  - Implemented in PS3 - you should know about it!
  - Steps to calculate HOG feature
Image Classification and 2D Object Detection

- 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!