PSET 1 Part 1 + Project Outline

Yan Wang

CS231A

01/17/2023
Overview

- Lecture Review
- PSET 1 Part 1
- Project Logistics
- Types of Projects
- Class Coverage and Ideas
- Where to Get Projects
- Helpful Resources
Pinhole camera

\[
P = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow P' = \begin{bmatrix} x' \\ y' \end{bmatrix}
\]

\[
\begin{align*}
x' &= f \frac{x}{z} \\
y' &= f \frac{y}{z}
\end{align*}
\]  
[Eq. 1]

Derived using similar triangles
Homogeneous coordinates

\[ E \rightarrow H \]

\[
(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}
\]

homogeneous image coordinates

\[
(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}
\]

homogeneous scene coordinates

• Converting back \textit{from} homogeneous coordinates

\[ H \rightarrow E \]

\[
\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)
\]

\[
\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)
\]
Projective transformation in the homogenous coordinate system

\[
P_h' = \begin{bmatrix}
\alpha x + c_x z \\
\beta y + c_y z \\
z
\end{bmatrix}
= \begin{bmatrix}
\alpha & 0 & c_x & 0 \\
0 & \beta & c_y & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
1
\end{bmatrix}
\]

\[P_h \]

[Eq.8]

Homogenous \quad Euclidian

\[P_h' \rightarrow P' = (\alpha \frac{X}{Z}, \beta \frac{Y}{Z})\]

\[M = \begin{bmatrix}
\alpha & 0 & c_x & 0 \\
0 & \beta & c_y & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}\]
The Camera Matrix

[Eq. 9]

\[ P' = MP \]

\[ = K \begin{bmatrix} I & 0 \end{bmatrix} P \]

\[ P' = \begin{bmatrix} \alpha & 0 & c_x & 0 \\ 0 & \beta & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \]
3D Translation of Points

\[
T = \begin{bmatrix}
T_x \\
T_y \\
T_z \\
\end{bmatrix}
\]

\[
P' \rightarrow \begin{bmatrix}
1 & T \\
0 & 1
\end{bmatrix}_{4 \times 4}
\begin{bmatrix}
x \\
y \\
z \\
1
\end{bmatrix}
\]

A translation vector in 3D has 3 degrees of freedom
3D Rotation of Points

Rotation around the coordinate axes, counter-clockwise:

\[ R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \]

\[ R_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \]

\[ R_z(\gamma) = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \]

\[ R = R_x(\alpha) R_y(\beta) R_z(\gamma) \]

\[ P' \rightarrow \begin{bmatrix} R & 0 \\ 0 & 1 \end{bmatrix}_{4\times4} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \]

A rotation matrix in 3D has 3 degrees of freedom
3D Translation and Rotation

\[ R = R_x(\alpha) \quad R_y(\beta) \quad R_z(\gamma) \quad \]

\[ T = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \]

\[ P' \rightarrow \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}_{4\times4} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \]
World reference system

In 4D homogeneous coordinates:

\[ P = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}_{4 \times 4} P_w \]

\[ P' = K \begin{bmatrix} I & 0 \\ 0 & 1 \end{bmatrix} P = K \begin{bmatrix} I & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}_{4 \times 4} P_w = \begin{bmatrix} R & T \end{bmatrix}_{4 \times 4} P_w \]

[Eq. 9] Internal parameters

[Eq. 11] External parameters
Q1 - Projective Geometry

1 Projective Geometry Problems [20 points]

In this question, we will examine properties of projective transformations. We define a camera coordinate system, which is only rotated and translated from a world coordinate system (with a rigid transform).

(a) Prove that parallel lines in the world reference system are still parallel in the camera reference system (note: this is an example of a rigid transformation). [4 points]

(b) Now let us consider affine transformations, which are any transformations that preserve parallelism. Affine transformations include not only rotations and translations, but also scaling and shearing. Given some vector \( p \), an affine transformation is defined as

\[
A(p) = Mp + b
\]

where \( M \) is an invertible matrix. Prove that under any affine transformation, the ratio of parallel line segments is invariant, but the ratio of non-parallel line segments is not invariant. [6 points]

(c) Consider a unit square \( pqrst \) in the world reference system where \( p, q, r, \) and \( s \) are points. Will the same square in the camera reference system always have unit area? Prove or provide a counterexample. Similarly, will the unit square be preserved or not under an affine transformation? [4 points]

(d) You have explored whether these three properties (rotational and translational invariance, and ratio of parallel line segments) hold for affine transformations. Do these properties hold under any projective transformation? Justify briefly in one or two sentences (no proof needed). [6 points]
(a) Prove that parallel lines in the world reference system are still parallel in the camera reference system (note: this is an example of a rigid transformation). [4 points]

- Lines \( k \) and \( l \) are parallel
  - \( k_1 \) and \( k_2 \) are any two points on \( k \)
  - \( l_1 \) and \( l_2 \) are any two points on \( l \)
  - by definition of parallel lines:
    \[
    (k_1 - k_2) \times (l_1 - l_2) = 0
    \]
    
    \[
    (k_1 + p - k_2 - p) \times (l_1 + p - l_2 - p) = 0
    \]
    
    \[
    (Rk_1 - Rk_2) \times (Rl_1 - Rl_2) = 0
    \]
(b) Now let us consider affine transformations, which are any transformations that preserve parallelism. Affine transformations include not only rotations and translations, but also scaling and shearing. Given some vector \( p \), an affine transformation is defined as

\[
A(p) = Mp + b
\]

where \( M \) is an invertible matrix. **Prove that under any affine transformation, the ratio of parallel line segments is invariant, but the ratio of non-parallel line segments is not invariant.** [6 points]

Consider any two parallel lines \( k \) and \( l \). Take the segment between any two points \( k_1, k_2 \) on \( k \) and the segment between any two points \( l_1, l_2 \) on \( l \). By definition of parallel segments,

\[
k_1 - k_2 = k(l_1 - l_2)
\]

for some real number \( k \). Thus,

\[
\|k_1 - k_2\| = k\|l_1 - l_2\|
\]
(c) Consider a unit square $pqrs$ in the world reference system where $p$, $q$, $r$, and $s$ are points. Will the same square in the camera reference system always have unit area? Prove or provide a counterexample. Similarly, will the unit square be preserved or not under an affine transformation? [4 points]

- Given a square $pqrs$,
  \[
  \text{Area} = \| (q-p) \times (s-p) \| = \| q-p \| \times \| s-p \| \sin \theta = 1
  \]

- Hint: projecting to camera frame is isometric (they preserve lengths, as well as angles between lines)
- How about affine transformation?
(d) You have explored whether these three properties (rotational and translational invariance, and ratio of parallel line segments) hold for affine transformations. Do these properties hold under any projective transformation? **Justify briefly in one or two sentences** (no proof needed). [6 points]

- **Hint:** Think about the difference between Projective Transformation and Affine Transformation. Briefly explain in 1 or 2 sentences.
Project Logistics

• Overview [here](#)
• Teams of **1-3**: Number of people is taken into account when grading project
  • More members - More work
• Suggestions for project direction
  • Replicate an interesting paper
  • Compare different methods to a benchmark
  • Use a new approach to an existing problem
  • Implement an interesting system
  • Original research
Sharing a Project with Another Class

• Sharing projects is generally allowed
• Specify in reports
• Must be approved by both our staff and the other course staff
• Project must be profound enough that you can clarify which parts of the project were done for which class
  • Each part must be substantial enough to hold as a single project
  • Technical parts and experiments should be sufficient and different
  • For example, if you want to use CNN for flower classification, you can include some other components related to this course (e.g. geometry, …)
• Will need a separate write-up for each class
Project Grading - Important Dates

• Course project: 46%
  • Project proposal 4% (due Feb 3)
  • Midterm milestone 10% (due Feb 27)
  • Presentation 10% (due March 20)
  • Final report 22% (due March 22)
Project Proposal

• Maximum of 2 pages
• Submit the report as a PDF document through Gradescope
• Include the following:
  • Title and authors
  • Sec. Introduction: Problem you want to solve and why
  • Sec. Technical Approach: How do you propose to solve it?
  • Sec. Milestones (dates and sub-goals)
  • References
• You will be assigned a project mentor
Project Milestone Report

• Maximum of 4 pages
• Submit the report as a PDF document through Gradescope
• Include the following:
  • Title and authors
  • Sec. Introduction: Problem you want to solve and why
  • Sec. Technical Approach: How do you propose to solve it?
  • Sec. Milestones achieved so far
  • Sec. Remaining Milestones (dates and sub-goals)
  • References
Project Presentations

• Short presentation with time for a brief Q&A

• Include the following:
  • Problem Motivation/Description
  • Technical Approach
  • Results
  • Maybe demo (+)!
Project Final Report

• Length of 6-8 pages
• Submit the report as a PDF document through Gradescope
• Submit your code
• Include the following:
  • Title and authors
  • Abstract
  • Sec. Introduction
  • Sec. Previous work
  • Sec. Technical Approach
  • Sec. Experiments
  • Sec. Conclusions
  • References
Class Coverage

- Camera models and calibration
  - Single camera and how we model it
- Single view metrology
  - Estimating geometry from a single view
- Epipolar Geometry (Stereo Vision)
  - Estimating geometry from two viewpoints
- Structure from Motion
  - Using motion/several viewpoints to estimate structure
- Volumetric Stereo
  - Using multiple views to map 3D points
Class Coverage

● Representations and Representation Learning
  ○ Extracting features from 2D images for downstream applications
● Monocular Depth Estimation & Feature Tracking
  ○ Estimating depth in images, tracking of pixels in videos
● Optical and Scene Flow
● Optimal Estimation
● Neural Radiance Fields
View Morphing

Image morphing techniques can generate compelling 2D transitions between images.

View Morphing
Automatic Photo Pop-Up

A fully automatic method for creating a 3D model from a single photograph
Novel Hardware
Mobile Devices

Can you take an existing vision algorithm and adapt it to a mobile device to make it more useful?
Monocular Depth estimation
Camera Calibration for Robots
Recognizing Panoramas

Image Segmentation

Partition an image into multiple segments (sets of pixels) in order to make it easier to analyze.
Tracking
Face Detection – Face Identification
Other Topics

• Pose Estimation: Estimate the skeleton angles for a person from an image/video
• Action and Gesture Recognition: Is a person standing, walking, or sitting in an image/video? Is he/she waving?
• Scene Understanding: Can you classify a scene? Can you recognize and/or segment each component of the scene?
• Trajectory Forecasting
  • ...
Negative project examples

- Projects without components related to the course
- Applying Alexnet for image classification
- Finding and running an existing Github code
- Only running OpenCV libraries for a task
- ...
Where to get Project Ideas

• Course Staff: Office hours, ideas posted on website
• Computer vision papers and conferences
  • CVPR
  • ICCV
  • ECCV
• Computer vision research groups at Stanford
  • Silvio Savarese
  • Fei-Fei Li
  • Juan Carlos Niebles
• Past projects: See course website
• Papers with Code: Computer Vision
• Come up with your own!
Datasets

- Many are available on the web
- See the following aggregators:
  - CV Datasets on the Web
  - Yet Another Computer Vision Index To Datasets (YACVID)
- References found in papers
- Course CA’s
Project Advice

• Choose your team well
• Make sure the scope of your project fits a quarter
  • Set a minimum goal, desired goal, and a moonshot
• Constrain your problem smartly
• See what datasets are available if you are doing a recognition project
  • Specially for deep learning projects
• You may need to plan ahead/learn outside materials
• Use software when available
  • OpenCV, MATLAB, Deep learning frameworks
• Come ask questions – We’re happy to talk!
Thanks!

Questions