CS 5970/6970
Fall 2025
Assignment 3 – 3D Vision – Stereo and SfM in the Wild
Due: 11/02/2025 11:59 pm

In this assignment, you will capture your own images using a smartphone and implement key components of classical 3D vision: stereo matching and structure-from-motion (SfM). This hands-on task will help you appreciate how camera placement, motion, and calibration influence 3D reconstruction in practice.

Task 1: Stereo Vision [25 Points]

Capture two stereo images of a tabletop scene and compute a disparity map.

Instructions:

1. Choose a static scene with 2–3 textured objects on a table.
2. Place your phone at two different positions (left and right), keeping it roughly aligned horizontally.
3. Record the approximate distance between the two camera positions (the baseline).
4. Capture two images: one from each viewpoint.
5. Implement a stereo matching pipeline:
    a. Detect and match features.
    b. Estimate the fundamental matrix and draw epipolar lines to verify correspondences.
    c. Use OpenCV's StereoBM or StereoSGBM to compute the disparity map.
    d. Analyze how baseline distance and camera orientation affect disparity quality.
6. Write-up (1–2 paragraphs):
    a. How did you ensure camera alignment?
    b. Did the disparity map make sense?
    c. What happens if the baseline is very small or too large?

Task 2: Structure from Motion **[25 Points]**

Capture a sequence of 3–5 images of a small object or your workspace from different viewpoints and reconstruct the 3D structure via SfM.

Instructions:

1. Move your phone around the object, keeping it pointed at the scene. Try to keep a smooth, circular motion.
2. Extract 3–5 key frames from your video (or capture still images manually).
3. Write a pipeline that:
    a. Extracts and matches SIFT features across consecutive pairs.
    b. Estimates relative pose (rotation and translation) between views.
    c. Triangulates matched points to reconstruct a sparse 3D point cloud.
    d. Tracks and visualizes camera trajectory.
4. Write-up (1–2 paragraphs):
    a. Did your reconstruction look realistic?
    b. What kinds of motion caused errors or degenerate configurations?
    c. What were your main failure modes?

Task 3: Calibration Challenge **[10 points]**

Improve your stereo and SfM results by estimating your camera intrinsics.

Instructions:

1. Estimate the intrinsic camera matrix K for your phone using any calibration approach you prefer. You do not need to write a calibration pipeline from scratch. You can use OpenCV to implement any functions that you need.
2. Rerun all you experiments from Task 1 and task 2 with this estimated K and provide a detailed comparison of reconstruction quality before and after calibration.
    a. What changed when using better intrinsics?
    b. How sensitive is your pipeline to poor calibration?

Deliverables:

1. An IPYNB notebook that can run on Google Colab.

    a. You can use the same notebook for each part. Make sure to label them correctly with a description using a text cell.
2. Your code that you used to perform all three tasks.
    a. Provide proper documentation on how to run your code, any dependencies, etc. as a text cell *before* your code cell.
3. All your data that is needed to run and reproduce your code and analysis. You can embed the link as Google Drive folder or upload the images with your code in a single ZIP file.
4. A description of your approach, what worked and what didn't and GenAI usage declaration, and a statement of contribution for each member in your team. If it is an individual project, then just say "This was an individual project and I did everything in it."