HW2

Q1-1: Video(https://youtu.be/INTPF3wqitE)
The scene is a wood table in my lab.

Q1-2: I didn't do it.
Since my personal things happened during the 10/10 weekend, and other days are occupied by other homeworks, I choose not to do this part, sorry.

Q2: Video(https://youtu.be/dd49sMofK04)
Q2-1:
All codes are generated by Claude Sonnet 4.5.

1. Matching each 2d point in the validation image by descriptor to 3d by solving PnP(using P3P+RANSAC).
2. P3P+RANSAC is done like the following (I set it to at most 300 iteration)
    a. Randomly choose 3 points and use Kneip-p3p to find R and t.
    b. Use R and t to perform reprojection, for all points, and calculate errors, for each pair of points that the error is in the given threshold, mark it inlier.
    c. If the number of inliers is larger than the original max, update the max inlier sets.
    d. If the inlier ratio(number of inliers / total points) is high enough, reduce the number of iterations.
    e. Iterate until the number of iteration meets the maximum(adaptive) iteration.
    f. Calculate R and t using the largest set of inliers from the above algorithm.
3. Do 2 for each validation image.
4. Calculate the error between ground truth R,t and predicted R_hat, t_hat, and get the median. (I can see some outliers, but median doesn't capture them)
5. Store each R,t for those images. (for 2-2)

Q2-2:
All codes are generated by Claude Sonnet 4.5

1. Adjust the cube using transform_cube.py.
2. Using R,t of each frame makes an image (possibly) containing the cube.
3. Bringing those images together forms a video.
4. The cube video(ar_cube_video.mp4)(https://drive.google.com/file/d/145WY9TO2OhpzYl0tUliBElZpMfa_-roJ/view?usp=sharing)