

Assignment 4. Stereo and Motion (Deadline: Nov 17 before midnight)
(Those marked with * are optional for extra credits)

Note: Turn in a PDF document (**in writing; please type**) containing a list of your .m files (not the code itself), images showing the results of your experiments, and an analysis of the results. All the writings must be soft copies in print and be sent to Prof. Zhu <cv.zhu.ccny@gmail.com> via email. For the programming part, **send ONLY your source code by email**; please don't send in your images and executable (even if you use C++). You are responsible for the loss of your submissions if you don't write "**CSc 471 Computer Vision Assignment 4**" in the subject of your email. Do write your names and IDs (last four digits) in both both of your report and the code. **Please don't zip your report with your code and other files; send me the report in a separate PDF file. The rest can be in a zipped file.**

1. (Stereo- **30 points**) Estimate the accuracy of the simple stereo system ($Z = f B/d$, Figure 3 in the lecture notes of stereo vision) assuming that the only source of noise is the localization of corresponding points in the two images, which means the error in estimating d . Please derive (10 points) and discuss (20 points) the dependence of the error in depth estimation of a 3D point as a function of **(1) the baseline width B , (2) the focal length f , (3) stereo matching error, and (4) the depth of the 3D point, Z .**

Hint: Take the partial derivatives of Z with respect to d , assuming that both B and f and constant parameters.

2. (Motion- **40 points**) Could you obtain 3D information of a scene by viewing the scene by a camera rotating around its optical center (10 points)? Discuss why or why not (10 points). What about translating the camera along the direction of its optical axis (10 points)? Explain. (10 points)

3. (Motion- 10 points) Explain that the aperture problem can be solved if a corner is visible through the aperture.

4. (Stereo Programming - **20 points + 20 bonus points**) Use the image pair ([Image 1](#), [Image 2](#)) for the following exercises.

(1). Fundamental Matrix (**20 points**). - Design and implement a program that, given a stereo pair, determines at least eight point matches, then recovers the fundamental matrix (**10 points**) and the location of the epipoles (**5 points**). Check the accuracy of the result by measuring the distance between the estimated epipolar lines and image points not used by the matrix estimation (**5 points**). Also, overlay the epipolar lines of control points and test points on one of the images (say Image 1- I already did this in the starting code below). Control points are the

correspondences (matches) used in computing the fundamental matrix, and test points are those used to check the accuracy of the computation.

Hint: As a first step, you can pick up the matches of both the control points and the test points manually. You may use my matlab code ([FmatGUI.m](#)) as a starting point - where I provided an interface to pick up point matches by mouse clicks. The epipolar lines should be (almost) parallel in this stereo pair. If not, something is wrong either with your code or the point matches.

***(2). Feature-based matching (10 bonus points).** - Design a stereo vision system to do "feature-based matching" and explain your algorithm in writing - what the feature is, how effect it is, and what are the problems. The system should have a user interface that allows a user to select a point on the first image, say by a mouse click. The system should then find and highlight the corresponding point on the second image, say using a cross hair points). Try to use the epipolar geometry derived from (1) in searching correspondences along epipolar lines.

Hint : You may use a similar interface as I did for question (1). You may use the point match searching algorithm in (1) (if you have done so), but this time you need to constrain your search windows along the epipolar lines.

***(3) Discussions (10 bonus points).** Show your results on points with different properties like those in corners, edges, smooth regions, textured regions, and occluded regions that are visible only in one of the images. Discuss for each case, why your vision system succeeds or fails in finding the correct matches. Compare the performance of your system against a human user (e.g. yourself) who marks the corresponding matches on the second image by a mouse click.