16-822: Geometry-based Methods in Vision (F17) Released: Oct-25., Due: Nov 15th

Homework 3

Lecturer: Martial Hebert *TA: Aayush Bansal*

1 Affine cameras (35 Points)

In class, we have mostly focused on perspective cameras. For this part of the homework, we consider cameras with center lying on the plane at infinity, especially *affine cameras*. An affine camera is one that has a camera matrix of the form $M = [A \,|\, b]$ in which A is singular and the last row of M is of the form $(0, 0, 0, 1)$. The following questions explore affine cameras a little more in depth.

- What happens to points on a plane perpendicular to the image plane that also passes through the world origin?
- Given a set of points X_i in $3D$, and their projection x_i in the image, what is the projection of the centroid of X_i's? (no equations, just 1 short line.)
- Show that an affine camera maps parallel world lines to parallel image lines; i.e., when using an affine camera, parallel lines in $3D$ project to parallel lines in image plane.[1]
- Show that for parallel lines mapped by an affine camera, the ratio of lengths on line segments is an invariant.

For the following questions, consider a system of 2 affine cameras, with camera matrix M and M': (Last row of each camera matrix is of the form $(0, 0, 0, 1)$.)

- Show that epipolar planes (and lines) are parallel.
- Show that the affine fundamental matrix F defined by two affine cameras is invariant to an affine transformation of the world coordinates.

2 Two camera geometry (10 + 20 + 10 Points)

2.1 2 singular values of E are equal (10 Points)

Let E be the *essential* matrix of a pair of cameras. We know that E is singular so that one of its singular value is zero. Show that the other 2 singular values are equal. (Remember that the singular values of a matrix A are the square roots of the *eigenvalues* of the matrix AA^T. Apply this property to E by apply EE^T to the appropriate vectors to figure out what the eigenvalues are.)

[1] Actually, the converse is also true; if the camera preserves parallelism of lines, then it is an affine cameras.

2.2 Recovering R, t from E (20 Points)

We address the problem of recovering the rotation R and translation t from the essential matrix $E = [t]_\times R$.

1. Show how to recover t from E (one single equation).
2. Show (using the properties of the essential matrix stated in class), that the SVD of E is of the form: $E = U\mathtt{diag}(1, 1, 0)V^T$.
3. Show that the two matrices $R_1 = UWV^T$ and $R_2 = UW^TV^T$ are solutions in R, where $W = [0\ -1\ 0;\ 1\ 0\ 0;\ 0\ 0\ 1]$.
4. Why is there 2 solutions in R (in words or drawing)?

2.3 Extra Credits - Π $\|$ to the baseline (10 Points)

Consider a plane Π viewed in a two-camera system. Show that if the vanishing line of Π in one image (i.e., the projection of the line at infinity of Π in the image) contains the epipole in that image, then Π is parallel to the baseline between the 2 cameras (the line joining the 2 camera centers).

3 Fundamental matrices between three cameras (25 Points)

In this problem we consider the relationship between 3 images. Let us denote by F_{ij} the fundamental matrix between image i and j (from 1 to 3), and by e_{ij} the epipole generated by image j in image i. We will now show that the 3 fundamental matrices induced by 3 cameras are not independent.

- Draw the 3 cameras with e_{ij} labelled.
- Show that: $F_{ij}e_{ik} = e_{jk} \times e_{ji}$ for all triplets. (this can be shown purely geometrically, without involved algebra)
- Based on these 3 quadratic relations, give a counting argument showing that the 3-camera geometry is characterized by 18 degrees of freedom. And thus conclude that the 3 fundamental matrices are not independent.

4 Auto-calibration (20 + 10 + 10)

4.1 General Questions (20 Points)

1. No equations, use a counting argument only: Suppose that we have m images from which we have generated a projective reconstruction. Suppose that the intrinsic parameters (K_i) are fixed throughout all of the images but unknown except for the skew which is known to be 0. How many images are needed to solve for the calibration parameters (auto-calibration toward metric reconstruction)?

2. Can it be solved in a linear fashion in this case?

3. Suppose that all the calibration parameters in K are allowed to vary across the images, but that we know that the two coordinates of the principal point are equal in each of the images: $x_o = y_o$. How many images are needed to solve the auto-calibration problem?[2]

4. Can it be solved linearly?

5. Linear approaches to the auto-calibration problem solve for $L = Q_3 Q_3^T$, where Q_3 is the 4×3 matrix of the first three columns of the 4×4 rectifying matrix Q. After solving for L, how can we recover Q_3?

4.2 Extra Credits - Auto-calibration, special case (10 Points)

We showed that, given m cameras, if we know the homography of the plane at infinity, H^i_∞, between the first image and each image i, we can derive the relations: $\omega^* = H^i_\infty \omega^* (H^i_\infty)^T$, where $\omega^* = KK^T$, assuming constant intrinsic parameters. Although, in principle we can solve for K given enough equations, there are cases in which the problem degenerates and an infinite family of solutions exist.

1. Does the homography of the plane at infinity between any two images depend on the translation between the images?

2. Show that, if the cameras are in translation only without rotation, the auto-calibration problem cannot be solved using the relations based on the plane at infinity as shown above (basically no equations, use the result of the previous question)[3].

4.3 Extra Credits - Weird constraints for auto-calibration (10 Points)

1. How many views m are needed to generate a metric reconstruction assuming that 1) the skew is zero and 2) the principal point is constrained to be on a circle of radius r centered at the origin $(0, 0)$?

2. Can it be solved linearly?

[2] FYI, this particular constraint does not correspond to a realistic physical constraint on the cameras, but it is still a valid constraint in theory.

[3] FYI, in fact, not only we need some amount of rotation to be able to solve unambiguously, we need rotations about at least three different axes. In other words, the camera needs to "move enough" to be able to self-calibrate unambiguously!