Assignment 2: CS 763, Computer Vision

Due: 4th February before 11:00 pm

Remember the honor code while submitting this (and every other) assignment. All members of the group should work on and understand all parts of the assignment. We will adopt a zero-tolerance policy against any violation.

Submission instructions: You should ideally type out all the answers in Word (with the equation editor) or using Latex. In either case, prepare a pdf file. For assignment submission, follow the instructions for arrangement of folders and subfolders as given in `https://www.dropbox.com/s/nbviv5h9g3a34ll/HW2_Alignment.tar.gz?dl=0`. Create a single zip or rar file obeying the aforementioned structure and name it as follows: A2-IdNumberOfFirstStudent-IdNumberOfSecondStudent-IdNumberOfThirdStudent.zip. (If you are doing the assignment alone, the name of the zip file is A2-IdNumber.zip). Upload the file on moodle BEFORE 11:00 pm on 4th February. Late assignments will be assessed a penalty of 50% per day late. Note that only one student per group should upload their work on moodle. Please preserve a copy of all your work until the end of the semester. If you have difficulties, please do not hesitate to seek help from me.

1. **Question 4:** Here you will develop an application of the concept of vanishing points to camera calibration. Consider two images (I_P and I_Q) of a non-planar scene taken with two pinhole cameras having unknown focal lengths f_p and f_q respectively. Both cameras produce images on a Cartesian grid with aspect ratio of 1 and unknown resolution s_p and s_q respectively. The orientations and positions of the two cameras are related by an unknown rotation (given by a 3×3 rotation matrix $\mathbf{R}$) and an unknown translation (given by a 1×3 vector $\mathbf{t}$). Note that 'position' here refers to the location of the camera pinhole, and 'orientation' refers to the XYZ axes of the camera coordinate system. In both I_P and I_Q, suppose you accurately mark out the corresponding vanishing points of three mutually perpendicular directions ℓ_1, ℓ_2, ℓ_3 in the scene. (Obviously, all three directions were visible from both cameras). Let the vanishing points have coordinates (p_{1x}, p_{1y}), (p_{2x}, p_{2y}) and (p_{3x}, p_{3y}) in I_P, and (q_{1x}, q_{1y}), (q_{2x}, q_{2y}) and (q_{3x}, q_{3y}) in I_Q. Note that these coordinates are in terms of pixel units and the correspondences are known. Given all this information, can you infer $\mathbf{R}$? Can you infer $\mathbf{t}$? Can you infer f_p and f_q? Can you infer s_p and s_q? Explain how (or why not). (Hint: You can start off by assuming that you knew all the intrinsic parameters and work yourself upwards from there). **[20 points]**

   Ans: Let (o_{px}, o_{py}) and (o_{qx}, o_{qy}) be the optical centers of the two cameras in their respective pixel coordinate systems. Corresponding to the three vanishing points, the direction vectors of the 3 lines in the coordinate systems of the first camera are $p_1 = (s_p(p_{1x} - o_{px}), s_p(p_{1y}, o_{py}), f_p)$, $p_2 = (s_p(p_{2x} - o_{px}), s_p(p_{2y}, o_{py}), f_p)$ and $p_3 = (s_p(p_{3x} - o_{px}), s_p(p_{3y}, o_{py}), f_p)$ respectively. Corresponding to the three vanishing points, the direction vectors of the 3 lines in the coordinate system of the second camera are $q_1 = (s_q(q_{1x} - o_{qx}), s_q(q_{1y}, o_{qy}), f_q)$, $q_2 = (s_q(q_{2x} - o_{qx}), s_q(q_{2y}, o_{qy}), f_q)$ and $q_3 = (s_q(q_{3x} - o_{qx}), s_q(q_{3y}, o_{qy}), f_q)$ respectively. We will first assume that all the intrinsic parameters are known, i.e $f_p, f_q, s_p, s_q, o_{px}, o_{py}, o_{qx}, o_{qy}$ are all known. We will unit normalize all these vectors by dividing the vectors by their respective magnitude. Let us denote the resultant unit vectors are $\tilde{p_1}, \tilde{p_2}, \tilde{p_3}, \tilde{q_1}, \tilde{q_2}, \tilde{q_3}$ respectively. The corresponding direction vectors will now be related as follows:

   $$(\tilde{p_1^t}|\tilde{p_2^t}|\tilde{p_3^t}) = \mathbf{R}(\tilde{q_1^t}|\tilde{q_2^t}|\tilde{q_3^t})$$

   Using this, you can determine $\mathbf{R}$ using matrix inversion (or transposition), assuming that the three sets of lines in 3D space were not coplanar (else the matrix on the RHS of the previous equation will not be invertible).

However, the translation parameters **cannot** be determined, because translation of a line without changing its direction does not change the coordinates of the vanishing point. The vanishing points therefore carry no information pertaining to translation

Now what if the intrinsic parameters are unknown? In pixel coordinates, we can determine $o_{px}, o_{py} and o_{qx}, o_{qy}$ from the property that the optical center is the orthocenter of the vanishing points corresponding to three mutually perpendicular lines. Also, the directions $\tilde{p}_i and \tilde{P}_j$ are perpendicular for $i \neq j$, we have $s_p^2(p_{1x} - o_{px})(p_{2x} - o_{px} + s_p^2(p_{1y} - o_{px})(p_{2y} - o_{py} + f_p^2 = 0$, from which we can determine s_p/f_p, and likewise for s_q/f_q. Thus, we can express s_p and s_q as constant multiples of f_p and f_q respectively. Now, as all three coordinates of an image point contain a term in f_p (or f_q), we can determine the directions of the three lines even if we don't know f_p (or f_q).