Assignment 1: CS 763, Computer Vision

Due: 31st Jan before 11:55 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 `http://www.cse.iitb.ac.in/~ajitvr/CS763_Spring2017/HW1/HW1_CameraGeometry.rar`. Create a single zip or rar file obeying the aforementioned structure and name it as follows: A1-IdNumberOfFirstStudent-IdNumberOfSecondStudent-IdNumberOfThirdStudent.zip. (If you are doing the assignment alone, the name of the zip file is A1-IdNumber.zip). Upload the file on moodle BEFORE 11:55 pm on 31st January. 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. In this exercise, we will prove the orthocenter theorem pertaining to the vanishing points Q, R, S of three mutually perpendicular directions OQ, OR, OS, where O is the pinhole (origin of camera coordinate system). Let the image plane be $Z = f$ without any loss of generality. Recall that two directions v_1 and v_2 are orthogonal if $v_1^T v_2 = 0$. One can conclude that OS is orthogonal to $OR - OQ$ (why?). Also the optical axis Oo (where o is the optical center) is orthogonal to $OR - OQ$ (why?). Hence the plane formed by triangle OSo is orthogonal to $OR - OQ$ and hence line oS is perpendicular to $OR - OQ = QR$ (why?). Likewise oR and oQ are perpendicular to QS and RS. Hence we have proved that the altitudes of the triangle QRS are concurrent at the point o. QED. Now, in this proof, I considered the three perpendicular lines to be passing through O. What do you think will happen if the three lines did not pass through O? [10 points]

2. Suppose you have acquired the image of a cricket pitch at the time instant that a ball thrown by the bowler landed on the ground somewhere on the pitch at some point say L_1. Given this image, your task is to propose two different methods to determine the perpendicular distance from L_1 to the line containing the wickets on the batsman's side. Make use of the standard dimensions of a cricket pitch as seen on `https://en.wikipedia.org/wiki/Cricket_pitch#/media/File:Cricket_pitch.svg`. Assume that the ball and all the sides of the pitch were clearly visible in the image. For both methods, explain (with equations if required) why you do not need a calibrated camera for this calculation, ignoring errors due to discretization of the spatial coordinates. [10+10=20 points]

3. 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]

4. Let $\boldsymbol{P}$ be the coordinates of a point in the world coordinate system. Let $\boldsymbol{r_1}, \boldsymbol{r_2}, \boldsymbol{r_3}$ respectively be the X, Y, Z axes of a camera's coordinate system expressed in the world coordinate system. Let the center of projection of the camera in the world coordinate system be vector $\boldsymbol{t}$. Express the coordinates of an orthographic image of $\boldsymbol{P}$ on the camera plane in the world coordinate system. [10 points]

5. You are given two datasets in the folder Q5 within `http://www.cse.iitb.ac.in/~ajitvr/CS763_Spring2017/HW1/HW1_CameraGeometry.rar`. The file names are Features2D_dataset1.mat, Features3D_dataset1.mat, Features2D_dataset2.mat and Features3D_dataset2.mat. Each dataset contains (1) the XYZ coordinates of N points marked out on a calibration object, and (2) the XY coordinates of their corresponding projections onto an image plane. Your job is to write a MATLAB program which will determine the 3×4 projection matrix $\mathbf{M}$ such that $\mathbf{P_1} = \mathbf{MP}$ where $\mathbf{P}$ is a $4 \times N$ matrix containing the 3D object points (in homogeneous coordinates) and $\mathbf{P_1}$ is a $3 \times N$ matrix containing the image points (in homogeneous coordinates). Use the SVD method and print out the matrix $\mathbf{M}$ on screen (include it in your pdf file as well). Write a piece of code to verify that your computed $\mathbf{M}$ is correct. For any one dataset, repeat the computation of the matrix $\mathbf{M}$ after adding zero mean i.i.d. Gaussian noise of standard deviation $\sigma = 0.05 \times max_c$ (where max_c is the maximum absolute value of the X,Y,Z coordinate across all points) to every coordinate of $\mathbf{P}$ and $\mathbf{P_1}$ (leave the homogeneous coordinates unchanged). Comment on your results. Include these comments in your pdf file that you will submit. **Tips:** A mat file can be loaded into MATLAB memory using the 'load' command. To add Gaussian noise, use the command 'randn'. For ease of grading, use the 'uigetdir' command in MATLAB to allow a user to choose a folder containing your input data (i.e. the aforementioned mat files). [20 points]

6. Consider the image in the folder Q6 within `http://www.cse.iitb.ac.in/~ajitvr/CS763_Spring2017/HW1/HW1_CameraGeometry.rar`. It is an image of two sheets of paper placed on the surface of a table. One (top right) is an A4 sheet of paper, i.e. it has width 21 cm and height 29.7 cm (these are its actual dimensions - obviously its apparent dimensions in the image are different). The second sheet of paper (bottom left) has unknown dimension. It is safe to assume that the surface of the table and both sheets of paper are completely coplanar. Your job is to design and implement a semi-automatic method (i.e. you can ask the user to click on some salient points) to determine the dimensions of the second sheet of paper using the fact that the first one has known dimensions. For ease of grading, use the 'uigetdir' command in MATLAB to allow a user to choose a folder containing your input data (i.e. the image provided). [20 points]