1. **Lecture 1: The camera as a geometric device**

The camera shown in Figure 1 has its $x$, $y$ and $z$ axes aligned with the world’s $y$, $z$ and $x$ axes respectively. The world frame’s origin is at $(0, -h, 4h)$ in the camera’s frame.

(a) By deriving a succession of Euclidean transformations, find the camera’s extrinsic camera calibration matrix $[R|t]$, such that

$$X_C = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} X_W$$

(b) The “ideal” image plane is placed at $z_C = 1$. (Ideal means that the intrinsic camera matrix $K$ is just a $3 \times 3$ identity matrix.) Derive the image coordinates of the vanishing point of the family of lines parallel to the following line, expressed parametrically as:

$$(X_W, Y_W, Z_W) = (2 + 4t, 3 + 2t, 4 + 3t).$$

(c) The intrinsic calibration matrix maps ideal image positions onto actual positions in pixels. Explain, with the aid of diagrams, the physical meaning of the focal length $f$, the aspect ratio $\gamma$, and the principal point $(u_o, v_o)$, and show why and where they appear in the calibration matrix. (Assume the skew $s$ is negligible.)

(d) The actual camera has $f = 800$ pixels, $\gamma = 0.9$, $s = 0$, and $(u_o, v_o) = (350, 250)$ pixels. Derive the coordinates in the actual image plane of the vanishing point of part (b).
2. Lecture 1: Camera calibration

(a) Starting with the projection equation \( x = P \mathbf{X} = K[R|\mathbf{t}] \mathbf{X} \) (all up to scale), describe an algorithm which uses the measured image positions \((x, y)_i\) of at least six points with known positions in the world \((X, Y, Z)_i\) to recover the extrinsic and intrinsic calibration parameters.

(b) Download the matlab code `calibration.m` from the web page at the top of the sheet, and annotate it with comments to confirm that it fits with your algorithm description in part (b).

(c) Set \( h = 1 \) in the geometry of Question (1), and determine the image positions that would be obtained in a camera with the following intrinsic matrix when viewing the following points \( \mathbf{X} \).

\[
K = \begin{bmatrix}
1 & 0 & 2 \\
0 & 2 & 1 \\
0 & 0 & 1 \\
\end{bmatrix}
\]

\[
\mathbf{X}_{1-6} = \begin{bmatrix}
0 \\
0 \\
0 \\
1 \\
1 \\
1 \\
\end{bmatrix}
\]

(d) Use the code to recover the extrinsic and intrinsic calibration from the image and world point data.

3. Lecture 2: Edges, Corners and SIFTs

(a) Modest radial distortion in a lens can be modelled by a single parameter \( \kappa \). As shown in Figure 2(a), an undistorted image point at radius \( r \) from the image centre is moved in the radial direction to \( r_d = r/\sqrt{1 - 2\kappa r^2} \).

i) By considering the distortion of the straight line \((x = k, y)\) in Figure 2(b), determine whether \( \kappa < 0 \) gives rise to barrel or pin cushion distortion.

ii) Determine the form of the inverse expression \( r = f(r_d) \) used to correct distortion.

iii) You have an edge detector and straight line fitter at your disposal, and can point the camera at any scene you wish. Devise the skeleton of an algorithm to determine \( \kappa \) automatically.

(b) Explain the statement *matching sections of straight edge constrains only one component of the 2D motion of the edge*.

(c) Write short notes on the Harris corner detector.

(d) Write one sentence to outline the FAST corner detector.

(e) Explain how feature locations are determined in the SIFT operator.

(f) Explain how the SIFT operator achieves invariance to (i) illumination; (ii) scale; (iii) rotation.

![Figure 2](image.png)
4. **Lecture 3: Epipolar geometry**

Figure 3 shows a scene point \( \mathbf{X} \) imaged at points \( \mathbf{x} \) and \( \mathbf{x}' \) in cameras \( C \) and \( C' \). The cameras have projection matrices \( P = K[I|0] \) and \( P' = K'[R|t] \), respectively.

(a)  
   i) Draw a labelled diagram and write short notes to explain the epipolar geometry arising from the observation \( \mathbf{x} \) in the first camera. Indicate entities such as the optic centre, baseline, epipole, epipolar line, and epipolar plane.
   
   ii) Assuming convergent cameras, explain how the epipolar lines corresponding to different image points \( \mathbf{x} \) appear in the second camera \( C' \).

(b) Derive the point at infinity \( \mathbf{Q} \) along the ray through \( \mathbf{x} \), and determine its projection \( \mathbf{q}' \) in the image of the second camera \( C' \). Also determine the epipole \( \mathbf{e}' \), and hence show that a homogeneous expression for the epipolar line in \( C' \) is

   \[
   l' = (K't) \times K'RK^{-1}x .
   \]

(c)  
   i) Use the result of part (b), and the general identity \( M\mathbf{a} \times M\mathbf{b} = M^{-\top}(\mathbf{a} \times \mathbf{b}) \), to derive a compact expression for the Fundamental Matrix \( F \).
   
   ii) Show that \( \mathbf{x}'^\top F \mathbf{x} = 0 \).

(d) A single camera with \( K = I \) captures an image, and then translates along its optic axis before capturing a second image, so that \( t = [0, 0, t_z]^\top \).

   i) Use \( \mathbf{x}'^\top F \mathbf{x} = 0 \) to derive an explicit relationship relating \( x', y', x, \) and \( y \).
   
   ii) Briefly relate your result to the expected epipolar geometry for this camera motion.
5. Lecture 3: More epipolar geometry

(a) Figure 4 shows a pair of ideal cameras each with focal length unity whose principal axes meet at a point. The y-axes of both cameras are parallel and point out of the page.

i) Assuming $C$ has projection matrix $P = [I|0]$, work out the projection matrix $P'$ for $C'$ and hence evaluate the fundamental matrix (using the form derived in the previous question).

ii) Use the relationship $l' = Fx$ to compute the epipolar line in the right image corresponding to the homogeneous point $x = [1, 1, 1]^\top$ in the left image.

(b) i) Expressing the projection matrices of the two cameras as $P = K[I|0]$ and $P' = K'[R|t]$, derive the 4-vectors which represent the cameras’ optical centres, and hence derive expressions for the epipoles $e$ and $e'$.

ii) Show that $e$ and $e'$ are the right and left null-spaces respectively of the fundamental matrix $F$.

iii) Using Matlab or otherwise, verify your results using configuration of part (a).

(c) A camera rotates about its optical centre and changes its intrinsics so that the camera matrices before and after $P = K[I|0]$ and $P' = K'[R|0]$.

i) Show that the images are related by an homography $x' = Hx$.

ii) Devise the skeleton of a practical algorithm to use this expression to rectify a pixelated image. (You will need to consider where in the original unrectified image the corners of each rectified pixel appear.)
6. **Lecture 4: Dense stereo correspondence**

(a) The ordering constraint is often used in stereo correspondence algorithms to disambiguate point matches on corresponding epipolar lines. Sketch a configuration of surfaces and cameras for which the ordering constraint is valid, and a configuration for which it is not valid.

(b) Prove that if normalized cross-correlation is used to measure the similarity of image regions between two images (including adjusting the signal to have mean zero), then the measure is invariant (unchanged) under the following transformation of the intensities

\[
I(x, y) \leftarrow \alpha I(x, y) + \beta
\]

where \(\alpha\) and \(\beta\) are scalars.

(c) Explain why using normalized cross-correlation is likely to be more important for stereo matching than for matching between images in a monocular sequence.

7. **Lecture 4: RANSAC**

Using a simple point correspondence algorithm 200 point matches between two views have been established. It is estimated that up to 25% are outliers.

(a) If the 8-point algorithm is used to compute the Fundamental Matrix, determine minimum number of trials that will be required in order to be 99% confident that the correct motion has been found.

(b) Repeat the calculation for the 7-point algorithm and comment on the results.

8. **Lecture 4: Triangulation**

(a) Two parallel cameras, each with focal length \(f\) and otherwise ideal, are separated along the \(x\)-axis by \(t_x = 4\) units. By drawing a plan view of the system (ie, one with the \(x\) and \(z\) axes in the plane of the paper), and using very simple drawing construction, show that the coordinates of the 3D points reconstructed for the correspondences: (i) \([-1, 0] \leftrightarrow [1, 0]\) and (ii) \([0, 0] \leftrightarrow [0, 0]\) are (i) \([-2, 0, 2f, 1]^\top\) and (ii) \([0, 0, 1, 0]^\top\).

(b) For a parallel camera stereo configuration where each camera has focal length \(f\) the “horizontal” disparity \(d\) is given by \(d = (x' - x) = ft_x/Z\). Use differentiation to show that the error \(\delta Z\) in depth \(Z\) corresponding to an error \(\delta d\) in disparity is given by

\[
\delta Z = -\left(\frac{\delta d}{ft_x}\right)Z^2.
\]

By considering the nature of \(f\), \(t_x\) and \(\delta d\), suggest an obvious consequence of this relationship.

9. **For interest only: Back to Lecture 3, Computing the \(F\) matrix**

Download the matlab code `sevenpoint.m` from the web page at the top of the sheet. Study and run the code.

**For answers, hints, etc**

See the web page given at start of sheet.