C4B | 2006

**3.** Figure 1 shows a scene point $\boldsymbol{X}$ imaged at points $\boldsymbol{x}$ and $\boldsymbol{x}'$ in cameras C and C$'$. The cameras have projection matrices $\mathbf{P} = \mathrm{K}[\mathrm{I}|0]$ and $\mathbf{P}' = \mathrm{K}'[\mathrm{R}|\boldsymbol{t}]$, respectively.

(a) i) Draw a labelled diagram (based on Figure 1) and write short notes to explain the *epipolar geometry* arising from the observation $\boldsymbol{x}$ 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 $\boldsymbol{x}$ appear in the second camera C$'$.

[6 marks]

(b) A point at infinity along the ray through $\boldsymbol{x}$ is $\boldsymbol{Q} = \begin{bmatrix} \mathrm{K}^{-1}\boldsymbol{x} \\ 0 \end{bmatrix}$. Derive, in the image of the second camera C$'$,

i) the projection $\boldsymbol{q}'$ of the point at infinity $\boldsymbol{Q}$

ii) the epipole $\boldsymbol{e}'$.

Hence show that a homogeneous expression for the epipolar line in C$'$ is

$$\boldsymbol{l}' = (\mathrm{K}'\boldsymbol{t}) \times \mathrm{K}'\mathrm{R}\mathrm{K}^{-1}\boldsymbol{x}\,.$$

[5 marks]

(c) i) Use the result of part (b), and the general identity $\mathrm{M}\boldsymbol{a} \times \mathrm{M}\boldsymbol{b} = \mathrm{M}^{-\mathrm{T}}(\boldsymbol{a} \times \boldsymbol{b})$, to derive a compact expression for the Fundamental Matrix $\mathbf{F}$.

ii) Show that $\boldsymbol{x}'^{\mathrm{T}}\mathbf{F}\boldsymbol{x} = 0$.

[4 marks]

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

i) Use $\boldsymbol{x}'^{\mathrm{T}}\mathbf{F}\boldsymbol{x} = 0$ 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 marks]

Note. The skew-symmetric ("vector product") matrix is $[\boldsymbol{a}_\times] = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix}$.

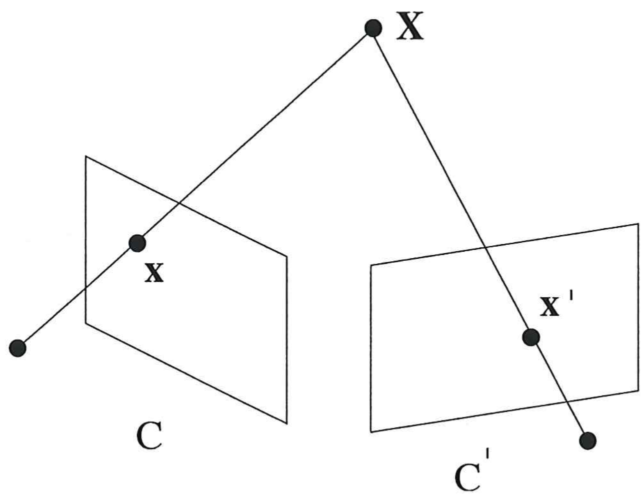


**Figure 1**

C4B COMPUTER VISION
2007

Comth.

**5.** (a) The projection matrix $\mathbf{P}_1$ of a camera $C_1$ is to be found up to scale using the measured image positions $\boldsymbol{x}_i$ of a number of known 3D scene points $\boldsymbol{X}_i$ in a Euclidean world frame, where the projection is $\lambda_i \boldsymbol{x}_i = \mathbf{P}_1 \boldsymbol{X}_i$.

Writing the elements of $\mathbf{P}_1$ as $p_{11}$, $p_{12}$, and so on, show that

$$\begin{bmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -x_iX_i & -x_iY_i & -x_iZ_i & -x_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -y_iX_i & -y_iY_i & -y_iZ_i & -y_i \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ \vdots \\ p_{34} \end{bmatrix} = \mathbf{0}\,.$$

Briefly outline an algorithm to recover $\mathbf{P}_1$ and to decompose it into the intrinsic calibration $\mathbf{K}_1$ and extrinsic calibration $\mathbf{R}_1, \boldsymbol{t}_1$.

[6 marks]

(b) A second camera $C_2$, in a different position, is calibrated using the same scene points, giving calibrations $\mathbf{K}_2$ and $\mathbf{R}_2, \boldsymbol{t}_2$.

For further analysis, a researcher applies a Euclidean transformation $X' = \mathbf{E}X$ to the scene points so that, for scene points $X'$, $C_1$'s projection matrix is $\mathbf{K}_1[\mathbf{I}|\mathbf{0}]$ and $C_2$'s projection matrix is $\mathbf{K}_2[\mathbf{R}|\boldsymbol{t}]$. Show that

$$\text{i)}\ \mathbf{E} = \begin{bmatrix} \mathbf{R}_1 & \boldsymbol{t}_1 \\ \mathbf{0}^\mathrm{T} & 1 \end{bmatrix} \quad \text{and ii)}\ [\mathbf{R}\,|\,\boldsymbol{t}\,] = [\mathbf{R}_2\mathbf{R}_1^\mathrm{T} \mid (-\mathbf{R}_2\mathbf{R}_1^\mathrm{T}\boldsymbol{t}_1 + \boldsymbol{t}_2)]\,.$$

[5 marks]

(c) With the aid of a labelled diagram, describe the basis of the epipolar geometry between two views and state why it is useful in feature matching.

[4 marks]

(d) In a particular case, the camera calibrations are

$$\mathbf{K}_1 = \mathbf{K}_2 = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix};\ \mathbf{R}_1 = \mathbf{R}_2;\ \boldsymbol{t}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix};\ \boldsymbol{t}_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}.$$

(i) Derive the Fundamental matrix $\mathbf{F} = \mathbf{K}_2^{-\mathrm{T}}\,[\boldsymbol{t}_\times]\,\mathbf{R}\,\mathbf{K}_1^{-1}$ where

$$[\boldsymbol{t}_\times] = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

(ii) Given a general point $\boldsymbol{x}_1 = [x_1, y_1, 1]^\mathrm{T}$ in the image of camera $C_1$, show that $\boldsymbol{x}_2^\mathrm{T}\mathbf{F}\boldsymbol{x}_1 = 0$ generates an expression for $(y_2/x_2)$ which is consistent with the camera motion.

[5 marks]

C4B 2008 DWM Computer Vision.

5. (a) A camera captures an image $I$, then rotates, translates, and changes its lens zoom setting before capturing image $I'$. The projection matrices in the two positions are $\mathbf{P} = \mathbf{K}[\mathbf{I}|\mathbf{0}]$ and $\mathbf{P}' = \mathbf{K}'[\mathbf{R}|\boldsymbol{t}]$, and a scene point is imaged at the corresponding points $\boldsymbol{x}$ and $\boldsymbol{x}'$.

(i) Determine the point at infinity $\boldsymbol{Q}$ along the back-projected ray from image point $\boldsymbol{x}$. Also find $\boldsymbol{q}'$, the projection of $\boldsymbol{Q}$ into the image $I'$.

(ii) The optic centre of the first camera projects to the epipole $\boldsymbol{e}'$ in image $I'$. Determine an expression for $\boldsymbol{e}'$.

(iii) Hence show that the homogeneous representation of the epipolar line in image $I'$ is

$$\boldsymbol{l}' \sim \mathbf{K}'\boldsymbol{t} \times \mathbf{K}'\mathbf{R}\mathbf{K}^{-1}\boldsymbol{x} .$$

[5 marks]

(b) Assuming the general identity $\mathbf{M}\boldsymbol{a} \times \mathbf{M}\boldsymbol{b} = \mathbf{M}^{-\mathrm{T}}(\boldsymbol{a} \times \boldsymbol{b})$, use the result of part (a) to derive a compact expression for the Fundamental Matrix $\mathbf{F}$, and show that corresponding points are related by $\boldsymbol{x}'^{\mathrm{T}}\mathbf{F}\boldsymbol{x} = 0$.

[5 marks]

(c) For a scene point $[\boldsymbol{X}^{\mathrm{T}}, 1]^{\mathrm{T}}$ the projections can be written with explicit scales as

$$\lambda\boldsymbol{x} = \mathbf{K}\boldsymbol{X} \quad \text{and} \quad \lambda'\boldsymbol{x}' = \mathbf{K}'(\mathbf{R}\boldsymbol{X} + \boldsymbol{t}) .$$

When the changes between images $I$ and $I'$ consist of pure rotation and zoom only, show that the corresponding points are related by an homography $\boldsymbol{x}' \sim \mathbf{H}\boldsymbol{x}$, and give the expression for $\mathbf{H}$ (up to scale).

[4 marks]

(d) The camera now rotates, zooms, *and* translates between views, but the scene is a planar surface $\hat{\boldsymbol{n}}^{\mathrm{T}}\boldsymbol{X} = d$, where $\hat{\boldsymbol{n}}$ is the plane's unit normal.

(i) Using the first projection equation of part (c), and then the equation of the plane, show that

$$\frac{\boldsymbol{X}}{\lambda} = \mathbf{K}^{-1}\boldsymbol{x} \quad \text{and} \quad \frac{\boldsymbol{t}}{\lambda} = \frac{\boldsymbol{t}\,\hat{\boldsymbol{n}}^{\mathrm{T}}\mathbf{K}^{-1}\boldsymbol{x}}{d} .$$

(ii) Using the second projection equation of part (c), develop an expression for $\lambda'\boldsymbol{x}'/\lambda$. Hence show that the corresponding points are now related by an homography $\boldsymbol{x}' \sim \mathbf{J}\boldsymbol{x}$, and give the expression for $\mathbf{J}$ (up to scale).

[6 marks]

**5.** (a) The image projection $\boldsymbol{x}_i$ of a 3D Euclidean scene point $\boldsymbol{X}_i$ is given using homogeneous coordinates as $\lambda_i \boldsymbol{x}_i = \mathbf{P}\boldsymbol{X}_i$.

(i) Explain briefly how the projection matrix $\mathbf{P}$ is built up from the camera's extrinsic parameters $\mathbf{R}$ and $\boldsymbol{t}$, and intrinsic parameters $\mathbf{K}$.

(ii) Suppose that $\mathbf{P}$ has already been recovered (up to scale) using known scene and image points. Outline how the extrinsic and instrinsic parameters can be recovered from $\mathbf{P}$. (There is no need to consider sign ambiguities.)

[7 marks]

(b) Using a single calibration object to define a frame $B$, the extrinsic parameters of two cameras are found to be $\{\mathbf{R}_1, \boldsymbol{t}_1\}$, and $\{\mathbf{R}_2, \boldsymbol{t}_2\}$, such that points in the two camera frames are

$$\boldsymbol{X}^1 = \begin{bmatrix} \mathbf{R}_1 & \boldsymbol{t}_1 \\ \mathbf{0}^\mathrm{T} & 1 \end{bmatrix} \boldsymbol{X}^B \quad \text{and} \quad \boldsymbol{X}^2 = \begin{bmatrix} \mathbf{R}_2 & \boldsymbol{t}_2 \\ \mathbf{0}^\mathrm{T} & 1 \end{bmatrix} \boldsymbol{X}^B$$

Show that if $\boldsymbol{X}^1$ is used to define the world frame then $\mathbf{P}_1 = [\mathbf{I}|\mathbf{0}]$ and $\mathbf{P}_2 = [\mathbf{R}|\boldsymbol{t}]$ where

$$\mathbf{R} = \mathbf{R}_2\mathbf{R}_1{}^\mathrm{T} \quad \text{and} \quad \boldsymbol{t} = -\mathbf{R}_2\mathbf{R}_1{}^\mathrm{T}\boldsymbol{t}_1 + \boldsymbol{t}_2 \,.$$

[5 marks]

(c) A point is imaged at $\boldsymbol{x}_1$ in the first camera. By considering the projection into the second camera of the optic centre and a point at infinity of the back-projected ray in the first camera, show that the epipolar line in the second camera is

$$\boldsymbol{l}_2 \sim \mathbf{K}_2\boldsymbol{t} \times \mathbf{K}_2\mathbf{R}\mathbf{K}_1{}^{-1}\boldsymbol{x}_1 \,.$$

Hence, assuming the general identity $\mathbf{M}\boldsymbol{a} \times \mathbf{M}\boldsymbol{b} = \mathbf{M}^{-\mathrm{T}}(\boldsymbol{a} \times \boldsymbol{b})$, derive an expression for the Fundamental matrix $\mathbf{F}$.

[4 marks]

(d) Two parallel cameras are separated by a unit distance along the $x$-axis, so that $\boldsymbol{t} = [1, 0, 0]^\mathrm{T}$. The first camera has $\mathbf{K}_1 = \mathbf{I}$, but the second camera has a focal length twice that of the first camera.

Derive

(i) the fundamental matrix $\mathbf{F}$, and

(ii) the epipolar line corresponding to $\boldsymbol{x}_1 \sim [1, 1, 1]^\mathrm{T}$. Comment on your result.

[4 marks]

**5.** A number of static single cameras survey a town centre. Referred to a coordinate frame attached to one building, the rotation and translation of the $n^{\text{th}}$ camera are $\{\mathbf{R}_n, \boldsymbol{t}_n\}$, such that a point's description in that camera's frame is

$$\boldsymbol{X}^n = \begin{bmatrix} \mathbf{R}_n & \boldsymbol{t}_n \\ \mathbf{0}^{\mathrm{T}} & 1 \end{bmatrix} \boldsymbol{X}^B$$

(a) Consider just the cameras 1 and 2. If the frame of camera 1 is used to define the world frame, such that the projection matrices become $\mathbf{K}_1[\mathbf{I}|\mathbf{0}]$ and $\mathbf{K}_2[\mathbf{R}|\boldsymbol{t}]$, respectively, show that $\mathbf{R} = \mathbf{R}_2\mathbf{R}_1{}^{\mathrm{T}}$ and find the expression for $\boldsymbol{t}$.

[3 marks]

(b) A point in the scene is imaged at $\boldsymbol{x}_1$ in the first camera. Using the projections into camera 2 of the optic centre and a point at infinity of the back-projected ray in camera 1, show that the epipolar line in the second camera is

$$\boldsymbol{l}_2 \sim \mathbf{F}_{21}\boldsymbol{x}_1$$

and determine a compact expression for the fundamental matrix $\mathbf{F}_{21}$ in terms of $\mathbf{R}$, $\boldsymbol{t}$, and the instrinsic matrices $\mathbf{K}_1$ and $\mathbf{K}_2$. You can assume the general identity $\mathbf{M}\boldsymbol{a} \times \mathbf{M}\boldsymbol{b} = \mathbf{M}^{-\mathrm{T}}(\boldsymbol{a} \times \boldsymbol{b})$.

[5 marks]

(c) The $n^{\text{th}}$ camera's projection matrix $\mathbf{P}_n$ is estimated up to scale using known scene points in frame $B$ and their measured projections.

(i) Give the steps in an algorithm to recover the intrinsic and extrinsic parameters from the estimated $\mathbf{P}_n$.

(ii) Each camera has an image array of $768 \times 576$ pixels, and each pixel is square. The full field of view in the $x$-direction is $60°$. What values do you expect for the elements of $\mathbf{K}_n$?

[6 marks]

(d) Correspondence has been established between image points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in cameras 1 and 2. It is necessary to determine the corresponding *image* point $\boldsymbol{x}_3$ in camera 3.

By using the fundamental matrices $\mathbf{F}_{31}$ and $\mathbf{F}_{32}$, devise a method of finding $\boldsymbol{x}_3$ *that does not* involve finding the 3D point in the scene.

[6 marks]

DWM 19/1/11
C4B 2011

**1.** Two cameras $C$ and $C'$ with projection matrices $\mathbf{P} = \mathrm{K}[\mathrm{I}|0]$ and $\mathbf{P}' = \mathrm{K}'[\mathrm{R}|\boldsymbol{t}]$, image a scene point $\boldsymbol{X}$ at images points $\boldsymbol{x}$ and $\boldsymbol{x}'$, respectively. All symbols have their usual meanings.

(a) Define (up to scale) the homogenous coordinates of the point at infinity $\boldsymbol{Q}$ associated with the ray back-projected from $\boldsymbol{x}$, and derive its projection $\boldsymbol{q}'$ in the second camera. Also derive the position of the epipole $\boldsymbol{e}'$ in the second camera.

Sketch these and other relevant geometrical features on a diagram based on Figure 1.

[5 marks]

(b) Derive a homogeneous expression for the epipolar line $\boldsymbol{l}'$ in the second camera on which correspondences to $\boldsymbol{x}$ may lie, and state the endpoints of the allowed extent of correspondences.

Hence, and using the general identity $\mathrm{M}\boldsymbol{a} \times \mathrm{M}\boldsymbol{b} = \mathrm{M}^{-\mathrm{T}}(\boldsymbol{a} \times \boldsymbol{b})$, derive a compact expression for the Fundamental matrix $\mathbf{F}$.

[~~5~~ 4 marks]

(c) Stereo images are often *rectified* before correspondence matching. Indicate why this process is useful.

[~~2~~ 3 marks]

(d) If the image of camera $C$ is to be rectified to align with $C'$, the projection matrix of its rectified "camera" $C_{\text{rect}}$ should be $\mathrm{K}'[\mathrm{R}|0]$. By considering a scene point $\boldsymbol{X}$ viewed in the original and rectified cameras, derive the homography $\mathbf{H}$ which transforms points between the images as $\boldsymbol{x}_{\text{rect}} = \mathbf{H}\boldsymbol{x}$.

[4 marks]

(e) Suppose instead the image of camera $C'$ is to be rectified to align with $C$. Noting that the translation $\boldsymbol{t}$ used in the projection matrix of $C'$ is the position of the camera centre of $C$ *in the frame of* $C'$, derive (i) the desired projection matrix of its rectified "camera" $C'_{\text{rect}}$ and show that the transformation is

$$\boldsymbol{x}'_{\text{rect}} = \mathrm{K}\mathrm{R}^{\mathrm{T}}\mathrm{K}'^{-1}\boldsymbol{x}' .$$

[4 marks]

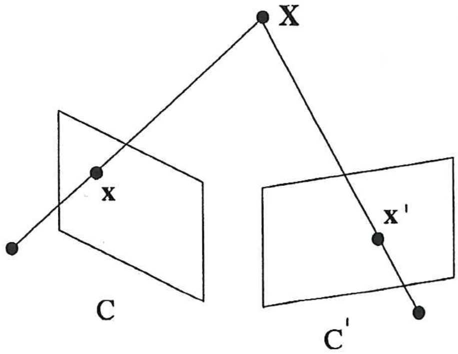


**Figure 1**

**1.** A camera captures an image, then rotates, translates and zooms before capturing a second image. The projection matrices are $\mathbf{K}[\mathbf{I}|\mathbf{0}]$ and $\mathbf{K}'[\mathbf{R}|\boldsymbol{t}]$, where symbols have their usual meanings.

(a) Drawing a carefully labelled diagram to illustrate your mathematical solution, show that the fundamental matrix describing the camera's epipolar geometry is

$$\mathbf{F} \sim \mathbf{K}'^{-\mathrm{T}}[\boldsymbol{t}_\times]\mathbf{R}\mathbf{K}^{-1}$$

Note (i) that $\sim$ denotes equality up to scale, and (ii) that you may assume the relationships

$$\mathbf{M}\boldsymbol{a} \times \mathbf{M}\boldsymbol{b} = \mathbf{M}^{-\mathrm{T}}(\boldsymbol{a} \times \boldsymbol{b}) \quad \text{and} \quad [\boldsymbol{t}_\times] = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

[7 marks]

(b) Show that corresponding image points $\boldsymbol{x} \leftrightarrow \boldsymbol{x}'$ are related by $\boldsymbol{x}'^{\mathrm{T}}\mathbf{F}\boldsymbol{x} = 0$.

[2 marks]

(c) The camera's motion between views is now a pure translation, and its intrinsic parameters remain fixed at $\mathbf{K}' = \mathbf{K} = \mathbf{I}$.

(i) For a correspondence between image points $\boldsymbol{x} = [x, y, 1]^{\mathrm{T}}$ and $\boldsymbol{x}' = [x', y', 1]^{\mathrm{T}}$, show that the result of part (b) can now be rewritten as the inner product

$$\begin{bmatrix} (y - y') & (x' - x) & (xy' - x'y) \end{bmatrix} \boldsymbol{t} = 0$$

(ii) For three such correspondences, $\boldsymbol{x}_1 \leftrightarrow \boldsymbol{x}'_1$, $\boldsymbol{x}_2 \leftrightarrow \boldsymbol{x}'_2$, and $\boldsymbol{x}_3 \leftrightarrow \boldsymbol{x}'_3$, show that the translation vector $\boldsymbol{t}$ is the null space of a $3 \times 3$ matrix $\mathbf{A}$, and write down the elements of $\mathbf{A}$.

[4 marks]

(d) Three corresponding points are found to be

| | Pair 1 | Pair 2 | Pair 3 |
|---|---|---|---|
| $[x, y]$ | $[1,\ 2]$ | $[1/2,\ -1]$ | $[0,\ 1/2]$ |
| $[x', y']$ | $[1,\ 3/2]$ | $[2/3,\ -1/3]$ | $[1/3,\ 2/3]$ |

Evaluate the matrix $\mathbf{A}$, and verify that $\boldsymbol{t} \sim [1, 1, 1]^{\mathrm{T}}$.

[3 marks]

**1.** A sensor unit contains two projective cameras $\mathcal{C}$ and $\mathcal{C}'$, with projection matrices $\mathrm{K}[\mathbf{I}|0]$ and $\mathrm{K}'[\mathrm{R}|\boldsymbol{t}]$ respectively.

(a) By considering the projection of a scene point $\boldsymbol{X}$ in the world coordinate system, briefly explain the physical origins of the matrices and vectors making up the camera projection matrices.

[3 marks]

(b) Consider homogeneous image points $\boldsymbol{x}$ in $\mathcal{C}$ and $\boldsymbol{x}'$ in $\mathcal{C}'$.

(i) With the aid of a labelled diagram, explain why the potential correspondences $\boldsymbol{x}'$ to a point $\boldsymbol{x}$ must lie on a restricted section of a line $\boldsymbol{l}'$. Ensure that your diagram contains all detail relevant to an explanation of epipolar geometry.

(ii) Hence derive the form of the Fundamental matrix in the relationship $\boldsymbol{l}' = \mathbf{F}\boldsymbol{x}$. You may use the general relationship $\mathrm{M}\boldsymbol{a} \times \mathrm{M}\boldsymbol{b} = \mathrm{M}^{-\mathrm{T}}(\boldsymbol{a} \times \boldsymbol{b})$.

[6 marks]

(c) Camera $\mathcal{C}$ is a "depth camera", so that the brightness of a pixel at $[x, y]^{\mathrm{T}}$ gives the depth $Z$ of the scene point $\boldsymbol{X}_3 = [X, Y, Z]^{\mathrm{T}}$ being projected into the camera.

Show that the three vector $\boldsymbol{x} = [xZ, yZ, Z]^{\mathrm{T}}$ is a valid homogeneous representation of the image point, and, starting from the projection equation of camera $\mathcal{C}$, derive the form of the relationship $\boldsymbol{X}_3 = f(\boldsymbol{x})$.

[3 marks]

(d) Camera $\mathcal{C}'$ is a colour camera, and to "colour" each depth point $\boldsymbol{X}_3$, the mapping between image points in $\mathcal{C}$ and $\mathcal{C}'$ is required.

(i) Determine the form of the relationship $\boldsymbol{x}' = \mathbf{A}\boldsymbol{x} + \boldsymbol{b}$.

(ii) Show that your result is consistent with $\boldsymbol{x}^{\mathrm{T}}\mathbf{F}^{\mathrm{T}}\boldsymbol{x}' = 0$

[4 marks]

**1.** (a) Two cameras $C$ and $C'$ have perspective projection matrices $\mathbf{K}[\mathbf{I}|\mathbf{0}]$ and $\mathbf{K}'[\mathbf{R}|\boldsymbol{t}]$. Drawing a carefully labelled diagram to illustrate your mathematical solution, show that the fundamental matrix describing the cameras' epipolar geometry is

$$\mathbf{F} \sim \mathbf{K}'^{-\mathrm{T}}[\boldsymbol{t}_\times]\mathbf{R}\mathbf{K}^{-1} ,$$

where $\sim$ denotes equality up to scale. You may assume without proof that

$$\mathbf{M}\boldsymbol{a} \times \mathbf{M}\boldsymbol{b} = \mathbf{M}^{-\mathrm{T}}(\boldsymbol{a} \times \boldsymbol{b}) .$$

[6 marks]

(b) Explain how knowledge of the fundamental matrix is used in a stereo correspondence algorithm, and briefly outline other knowledge that might be exploited to establish correspondence between image points $\boldsymbol{x}$ and $\boldsymbol{x}'$.

[3 marks]

(c) The cameras view a planar square mesh pattern (e.g., a chess board, or floor tiles). An example of the image edges recovered is given in Figure 1. Suppose that, by intersecting the edges from the two sets of parallel lines in the scene, two vanishing points $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ have already been found in the image of $C$.

(i) State why a image vanishing point in $C$ can be immediately mapped to a specific corresponding image *point* in $C'$, rather than merely to a point on an epipolar line.

(ii) Determine expressions for the positions $\boldsymbol{v}'_1$ and $\boldsymbol{v}'_2$ of the corresponding vanishing points in $C'$.

[3 marks]

(d) The line $\boldsymbol{h}$ joining $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in $C$ is the projection of the horizon belonging to the planar scene, and similarly for $\boldsymbol{h}'$ in camera $C'$.

(i) Using homogeneous notation, write down equations for $\boldsymbol{h}$ and $\boldsymbol{h}'$ in the two images.

(ii) Determine the compact expression for the matrix $\mathbf{A}$ that relates the projected horizons as

$$\boldsymbol{h}' = \mathbf{A}\boldsymbol{h} .$$

[4 marks]

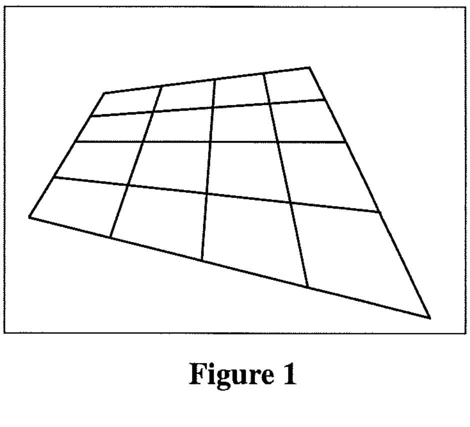

**Figure 1**

**1.** (a) The extrinsic calibration $\mathbf{R}$, $\boldsymbol{t}$ and intrinsic calibration $\mathbf{K}$ of a camera are to be recovered from 2D point features $\boldsymbol{x}_i$ imaged from known 3D scene points $\boldsymbol{X}_i$ on a calibration object.

(i) Write down the projection equation, and state the roles of $\mathbf{R}$, $\boldsymbol{t}$ and $\mathbf{K}$ in it.

(ii) Assuming the projection matrix $\mathbf{P}$ has been recovered up to scale, write down and briefly explain the principal steps in an algorithm to recover the values of the calibration parameters.

[5 marks]

(b) Two calibrated cameras $C$ and $C'$ have intrinsics $\mathbf{K}$ and $\mathbf{K}'$ and projection matrices $[\mathbf{I}_{3\times3}|\mathbf{0}]$ and $[\mathbf{R}|\boldsymbol{t}]$, respectively, where $\mathbf{I}_{3\times3}$ is a $3 \times 3$ indentity matrix.

Show that a point $\boldsymbol{x}$ in $C$ should correspond to a point on the line $\boldsymbol{l}'$ in $C'$ given by $\boldsymbol{l}' \sim \mathbf{F}\boldsymbol{x}$ and find a compact form of the fundamental matrix $\mathbf{F}$.

Note that "$\sim$" denotes equality up to scale, and you may assume the general matrix-vector identity $\mathbf{M}\boldsymbol{v}_a \times \mathbf{M}\boldsymbol{v}_b = \mathbf{M}^{-\mathrm{T}}(\boldsymbol{v}_a \times \boldsymbol{v}_b)$.

[4 marks]

(c) The two cameras view a planar scene described in $C$ by $\boldsymbol{n}^{\mathrm{T}}\boldsymbol{X}_3 = 1$, where $\boldsymbol{n}$ is the known normal to the plane (not necessarily a unit vector), and $\boldsymbol{X}_3$ is a scene point written as a 3-vector.

(i) Explain, without detailed proof, why the search for correspondence to an image point $\boldsymbol{x}$ in $C$ is now confined to a fixed point $\boldsymbol{x}'$ in camera $C'$, rather than to a point on a line.

(ii) Show that the back projection of $\boldsymbol{x}$ onto the plane gives the scene point as a homogeneous 4-vector

$$\boldsymbol{X} \sim \begin{bmatrix} \mathbf{K}^{-1}\boldsymbol{x} \\ \boldsymbol{n}^{\mathrm{T}}(\mathbf{K}^{-1}\boldsymbol{x}) \end{bmatrix} .$$

[3 marks]

(d) (i) Use the result of part (c) to show that $\boldsymbol{x}'$ and $\boldsymbol{x}$ are related by a $3 \times 3$ homography $\boldsymbol{x}' \sim \mathbf{H}\boldsymbol{x}$, where the homography can be written as

$$\mathbf{H} \sim \mathbf{K}'[\mathbf{R}|\boldsymbol{t}]\begin{bmatrix} \mathbf{I}_{3\times3} \\ \boldsymbol{n}^{\mathrm{T}} \end{bmatrix}\mathbf{K}^{-1} .$$

(ii) Use the constraint imposed on $\boldsymbol{x}'$ and $\boldsymbol{x}$ by the fundamental matrix to show that $\mathbf{H}^{\mathrm{T}}\mathbf{F}$ is antisymmetric.

[4 marks]