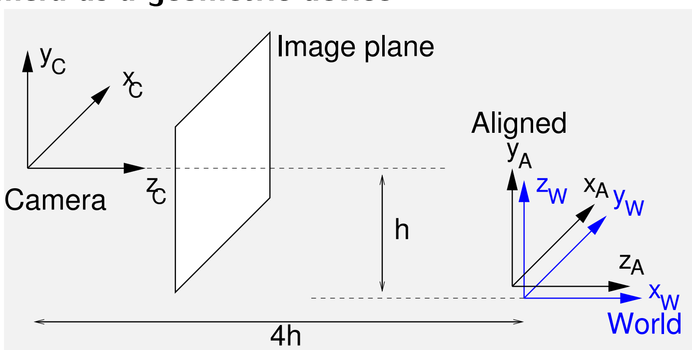
C18 Machine Vision 1: Solutions

Bugs/queries to david.murray@eng.ox.ac.uk
For answers, hints, corrections see the course page
at www.robots.ox.ac.uk/∼dwm/Courses/4CV

MT 2015
David Murray

1. Lecture 1: The camera as a geometric device



(a) **By deriving a succession of Euclidean transformations, find the camera's extrinsic camera calibration matrix** $[\mathrm{R}|\mathbf{t}]$**.**

First find the rotation from World to aligned frame A (either in steps — or in one go here as it is obvious). Second handle the translation.

Using inhomogeneous 3×1 scene vectors

$$\mathbf{X}_A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \mathbf{X}_W \qquad \mathbf{X}_C = \mathbf{X}_A + \begin{bmatrix} 0 \\ -h \\ 4h \end{bmatrix}$$

Now using 4×1 scene vectors

$$\Rightarrow \mathbf{X}_C = \begin{bmatrix} \mathrm{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \mathbf{X}_W = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -h \\ 1 & 0 & 0 & 4h \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{X}_W$$

(b) **The "ideal" image plane is placed at** $z_C = 1$**. Derive the image coordinates of the vanishing point of the family of lines parallel to** $(X_W, Y_W, Z_W) = (2+4t, 3+2t, 4+3t)$.

Point on the lines are $\mathbf{X}_W \stackrel{P}{=} \begin{bmatrix} 2+4t \\ 3+2t \\ 4+3t \\ 1 \end{bmatrix}$. Let $t \to \infty$, and the point at infinity is

$$\mathbf{X}_W^\infty \stackrel{P}{=} \begin{bmatrix} 4 \\ 2 \\ 3 \\ 0 \end{bmatrix}$$

. Projection as vanishing point is

$$\mathbf{x} \stackrel{P}{=} [\mathtt{I}|\mathbf{0}] \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -h \\ 1 & 0 & 0 & 4h \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \quad \Rightarrow \quad \mathbf{x} \stackrel{P}{=} \begin{bmatrix} 1/2 \\ 3/4 \\ 1 \end{bmatrix}$$

(c) **The intrinsic calibration matrix maps ideal image positions onto actual positions in pixels. Explain ...** (Lecture notes.)

(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).**

Using an explicit per-point scale λ to convert $\stackrel{P}{=}$ to $=$ ...

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 800 & 0 & 350 \\ 0 & 720 & 250 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1/2 \\ 3/4 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \Rightarrow \qquad \lambda x &= 800(1/2) + 350(1) \\ \lambda y &= 720(3/4) + 250(1) \\ \lambda &= 1 \end{aligned}$$

$$\Rightarrow \quad [x, y] = [750, 790]$$

---

2. Lecture 1: Camera calibration

(a) **Step 1: recover projection matrix up to scale.** Write

$$\lambda_i \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \mathtt{P}\mathbf{X}_i = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \mathbf{X}_i$$

Hence

$$\begin{aligned} (p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34})x_i &= (p_{11}X_i + p_{12}Y_i + p_{13}Z_i + p_{14}) \\ (p_{31}X_i + p_{32}Y_i + p_{33}Z_i + p_{34})y_i &= (p_{21}X_i + p_{22}Y_i + p_{23}Z_i + p_{24}) \end{aligned}$$

Re-arrange to give for point i

$$\begin{bmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -X_i x_i & -Y_i x_i & -Z_i x_i & -x_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -X_i y_i & -Y_i y_i & -Z_i y_i & -y_i \end{bmatrix} \mathbf{p} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

where vector $\mathbf{p} = (p_{11}, p_{12} \ldots, p_{33}, p_{34})^\top$ contains the unknown elements. Solve for 6 known points $i = 1 \ldots 6$ build up $\mathtt{A}\mathbf{p} = \mathbf{0}$ and find $\mathbf{p}$ as the null-space of $\mathtt{A}$

$$\mathbf{p} = \ker(\mathtt{A})$$

and refill projection matrix $\mathtt{P}$.

Step 2. Construct $\mathtt{P}_{\mathsf{LEFT}}$ from the leftmost3×3 block of $\mathtt{P}$. Invert it to obtain $\mathtt{P}^{-1}_{\mathsf{LEFT}}$. (It equals $\mathtt{R}^{-1}\mathtt{K}^{-1}$).

Step 3. Apply the $\mathcal{QR}$-decomposition to $\mathtt{P}^{-1}_{\mathsf{LEFT}}$

$$\mathcal{QR} \leftarrow \mathtt{P}^{-1}_{\mathsf{LEFT}}$$

Step 4. Recover the rotation

$$\mathtt{R} \leftarrow \mathcal{Q}^{-1}$$

Step 5. Recover the intrinsic matrix, but not $\mathtt{K} \leftarrow \mathcal{R}^{-1}$ directly. Instead

$$\begin{aligned} \mathtt{Temp} &\leftarrow \mathcal{R}^{-1} \\ \forall i,j : K_{ij} &\leftarrow \mathtt{Temp}_{ij}/\mathtt{Temp}_{33} \\ \forall i,j : P_{ij\ new} &\leftarrow P_{ij\ old}/\mathtt{Temp}_{33} \end{aligned}$$

Step 6. Recover the translation. The scale of $\mathtt{P}$ is now consistent with the scale $\mathtt{K}$. Recall $\mathtt{P} = \mathtt{K}[\mathtt{R}|\mathbf{t}]$, hence

$$\mathbf{t} \leftarrow \mathtt{K}^{-1}[P_{14} \quad P_{24} \quad P_{34}]^\top \qquad \text{or you could use the pre-scaled quantities } \mathbf{t} \leftarrow \mathcal{R}[p_4 \quad p_8 \quad p_{12}]^\top$$

(b) Code in calibration.m also fixes some sign ambiguities which were mentioned but not explained in the lecture (they would not be asked for). They arise because you can change the sign of a row or column of a rotation matrix and it remains a rotation matrix. To understand them, write the product $\mathtt{KR}$ in terms of symbols

$$\mathtt{KR} = \begin{bmatrix} (fR_{11} + sR_{21} + u_0R_{31}) & (fR_{12} + sR_{22} + u_0R_{32}) & (fR_{13} + sR_{23} + u_0R_{33}) \\ (\gamma f R_{21} + v_0R_{31}) & (\gamma f R_{22} + v_0R_{32}) & (\gamma f R_{23} + v_0R_{33}) \\ K_{33}R_{31} & K_{33}R_{32} & K_{33}R_{33} \end{bmatrix}$$

and argue that f, γf and K_{33} must be positive. So

- if $(f < 0)$, change the signs of $f, R_{11}, R_{12}, R_{13}$.
- if $(\gamma f < 0)$, change the signs of $\gamma f, s, R_{21}, R_{22}, R_{23}$.
- if $(K_{33} < 0)$, change the signs of $K_{33}, u_0, v_0, R_{31}, R_{32}, R_{33}$.

(c) Projection is

$$\begin{aligned} (\lambda_1 \cdots \lambda_6)\mathbf{x} &= \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix} [\mathtt{I}|\mathbf{0}] \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 \end{bmatrix} \left[\ \mathbf{X}_1\ \right] \cdots \left[\ \mathbf{X}_6\ \right] \\ &= \begin{bmatrix} 2 & 1 & 0 & 8 \\ 1 & 0 & 2 & 2 \\ 1 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 8 \\ 2 \\ 4 \end{bmatrix} \begin{bmatrix} 9 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 9 \\ 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 8 \\ 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 10 \\ 5 \\ 5 \end{bmatrix}, \begin{bmatrix} 11 \\ 5 \\ 5 \end{bmatrix} \end{aligned}$$

So measured inhomogeneous points are

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1/2 \end{bmatrix}, \begin{bmatrix} 9/4 \\ 1/2 \end{bmatrix}, \begin{bmatrix} 9/4 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 11/5 \\ 1 \end{bmatrix}$$

Yes, two image points are the same. No matter.

(d) Results from Matlab using the above data ...

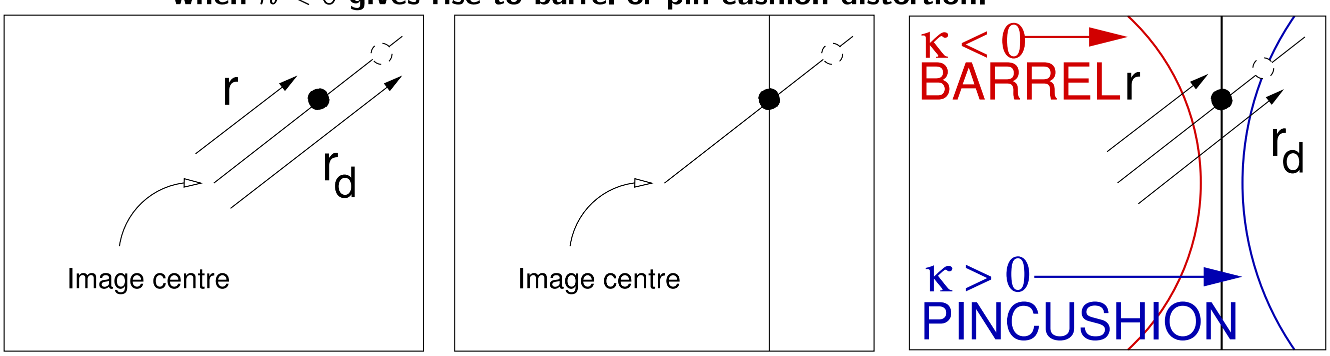
```
intrinsic =                           Rot =
1.0000    0.0000    2.0000           0.0000    1.0000    0.0000
     0    2.0000    1.0000          -0.0000   -0.0000    1.0000
     0         0    1.0000           1.0000   -0.0000    0.0000

translation =
 0.0000
-1.0000
 4.0000
```

Hurrah, it works.

---

3. Lecture 2: Edges, Radial Distortion & Corners

(a) i) **An undistorted image point at radius r is moved in the radial direction to $r_d = r/\sqrt{1-2\kappa r^2}$. By considering the distortion of the straight line in (b), determine when $\kappa < 0$ gives rise to barrel or pin cushion distortion.**



From the formula for $\kappa > 0$, the larger r gets the smaller the denominator becomes. The distorted point gets pushed ever further from the centre. Hence this is pincushion distortion. Obviously, vice versa for barrel distortion.

ii) **Find the inverse ...** No need to expand ... just rearrange

$$r_d = \frac{r}{\sqrt{1-2\kappa r^2}} \quad \Rightarrow \quad \frac{1}{r_d^2} = \frac{1}{r^2} - 2\kappa \quad \Rightarrow \quad \frac{1}{r^2} = \frac{1}{r_d^2} + 2\kappa \quad \Rightarrow \quad r = \frac{r_d}{\sqrt{1+2\kappa r_d^2}}$$

iii) **... skeleton of an algorithm to determine κ ...**

- Choose a scene with many straight lines. Capture an image.
- Detect edgels, and join into strings. Break strings where the curvature is high.
- Set $\kappa = 0$
- "Undistort" the edgel positions using the current κ $r = r_d/\sqrt{1+2\kappa r_d^2}$.

- Fit straight lines to the strings.
- Sum up the squared orthogonal deviations of edgels E_{iL} from their lines to generate a objective function $C = \sum_L \sum_{i \in L} \epsilon_{iL}^2$ to minimize.
- Select a new κ to reduce C, and repeat from (vi).

(b) Aperture problem ...

(c) From the notes.

---

4. Lecture 3: Epipolar geometry

(a) i) From the notes. Practice drawing.

ii) Different points $\mathbf{x}$ in general give rise to different epipolar planes which pierce the images at different epipolar lines. However, all planes share the base line, and so all epipolar lines pass through the epipole. Add diagram shownning "cat's whiskers".

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

$$\mathbf{Q} = \begin{bmatrix} \mathtt{K}^{-1}\mathbf{x} \\ 0 \end{bmatrix} \quad \Rightarrow \quad \mathbf{q}' = \mathtt{K}'[\mathtt{R}|\mathbf{t}] \begin{bmatrix} \mathtt{K}^{-1}\mathbf{x} \\ 0 \end{bmatrix} = \mathtt{K}'\mathtt{R}\mathtt{K}^{-1}\mathbf{x}'$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \quad \Rightarrow \quad \mathbf{e}' = \mathtt{K}'[\mathtt{R}|\mathbf{t}] \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = \mathtt{K}'\mathbf{t}$$

$$\Rightarrow \mathbf{l}' \;=\; \mathbf{e}' \times \mathbf{q}' \;=\; \mathtt{K}'\mathbf{t} \times \mathtt{K}'\mathtt{R}\mathtt{K}^{-1}\mathbf{x}' \;=\; [\mathtt{K}']^{-\top}(\mathbf{t} \times \mathtt{R}\mathtt{K}^{-1}\mathbf{x}) \;=\; [\mathtt{K}']^{-\top}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1}\mathbf{x} \;=\; \mathtt{F}\mathbf{x}$$

(c) i) **... derive a compact expression for the Fundamental Matrix** $\mathtt{F}$ ... done already

ii) **Show that $\mathbf{x}'^\top \mathtt{F}\mathbf{x} = 0$.**
Point $\mathbf{x}'$ is on the epipolar line, hence $\mathbf{x}'^\top \mathbf{l}' = 0$. Hence $\mathbf{x}'^\top \mathtt{F}\mathbf{x} = 0$.

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

i) **Use $\mathbf{x}'^\top \mathtt{F}\mathbf{x} = 0$ to derive an explicit relationship relating x', y', x, and y.** First find $\mathtt{F}$.

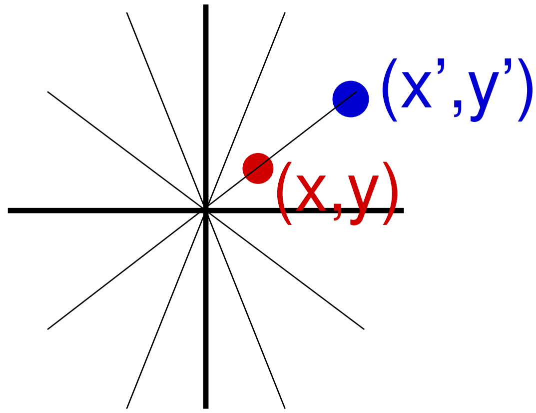
$$[\mathtt{K}']^{-\top}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1} = [\mathbf{t}]_\times = \begin{bmatrix} 0 & -t_z & 0 \\ t_z & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence

$$\mathbf{x}'^{\top}\mathtt{F}\mathbf{x} = t_z(-x'y + y'x) = 0 \quad \textit{hence} \;\; x'/y' = x/y \; .$$

ii) **Briefly relate your result to the expected epipolar geometry for this camera motion.**

The camera has moved along the optic axis, so we expect the origin to be the Focus of Expansion of a flow field. This is exactly what $x'/y' = x/y$ expresses.

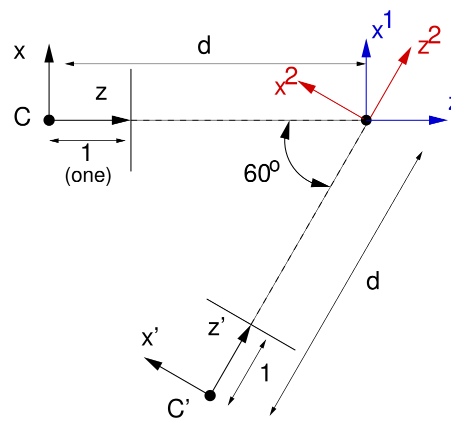


---

5. Lecture 3: Epipolar geometry

(a) **The figure shows a pair of 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) C has $\mathtt{K} = \mathtt{I}$ and $\mathtt{P} = [\mathtt{I}|\mathbf{0}]$. **Find** $\mathtt{P}'$**, find** $\mathtt{F}$**.**



Camera C' has $\mathtt{P}' = \mathtt{K}'[\mathtt{R}|\mathbf{t}]$ with $\mathtt{K}' = \mathtt{I}$, so this is all about finding the extrinsic matrix for the second camera, such that

$$\mathbf{X}' = \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^{\top} & 1 \end{bmatrix} \mathbf{X}$$

Build in stages.

$$\mathbf{X}^1 = \begin{bmatrix} \mathtt{I} & \begin{bmatrix} 0 \\ 0 \\ -d \end{bmatrix} \\ \mathbf{0}^{\top} & 1 \end{bmatrix} \mathbf{X} \;\; ; \;\; \mathbf{X}^2 = \begin{bmatrix} \cos 60 & 0 & -\sin 60 & 0 \\ 0 & 1 & 0 & 0 \\ \sin 60 & 0 & \cos 60 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{X}^1 \;\; ; \;\; \mathbf{X}' = \begin{bmatrix} \mathtt{I} & \begin{bmatrix} 0 \\ 0 \\ +d \end{bmatrix} \\ \mathbf{0}^{\top} & 1 \end{bmatrix} \mathbf{X}^2$$

Combine

$$\mathbf{X}' = \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^{\top} & 1 \end{bmatrix} \mathbf{X} = \begin{bmatrix} 1/2 & 0 & -\sqrt{3}/2 & +(\sqrt{3}/2)d \\ 0 & 1 & 0 & 0 \\ \sqrt{3}/2 & 0 & 1/2 & +(1/2)d \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{X}$$

$$
\begin{aligned}
\mathtt{F} &\stackrel{P}{=} [\mathtt{K}']^{-\top}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1} \\
&= \mathtt{I}^\top \begin{bmatrix} 0 & -(1/2)d & 0 \\ +(1/2)d & 0 & -(\sqrt{3}/2)d \\ 0 & +(\sqrt{3}/2)d & 0 \end{bmatrix} \begin{bmatrix} (1/2) & 0 & (-\sqrt{3}/2) \\ 0 & 1 & 0 \\ (\sqrt{3}/2) & 0 & (1/2) \end{bmatrix} \mathtt{I}^{-1} \\
&= \begin{bmatrix} 0 & -(1/2)d & 0 \\ -(1/2)d & 0 & -(\sqrt{3}/2)d \\ 0 & +(\sqrt{3}/2)d & 0 \end{bmatrix}
\end{aligned}
$$

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

$$
\mathbf{l}' \stackrel{P}{=} \mathtt{F}\mathbf{x} = \begin{bmatrix} 0 & -(1/2)d & 0 \\ -(1/2)d & 0 & -(\sqrt{3}/2)d \\ 0 & +(\sqrt{3}/2)d & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \stackrel{P}{=} \begin{bmatrix} +1 \\ +(1+\sqrt{3}) \\ -\sqrt{3} \end{bmatrix}
$$

(b) i) **Derive the 4-vectors which represent the optical centres of the two cameras.**
The world frame is coincident with the frame of the first camera, so that optical centre of C is at the world origin $\begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}$. The optical centre of C' is at $\begin{bmatrix} \mathbf{C}' \\ 1 \end{bmatrix}$ such that

$$
\begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathtt{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \begin{bmatrix} \mathbf{C}' \\ 1 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} \mathbf{C}' \\ 1 \end{bmatrix} = \begin{bmatrix} -\mathtt{R}^\top \mathbf{t} \\ 1 \end{bmatrix}
$$

ii) **Use the above results to derive expressions for the epipoles of the two cameras.**
Epipole $\mathbf{e}$ is the projection of the optic centre of C' into camera C.

$$
\mathbf{e} \stackrel{P}{=} \mathtt{K}[\mathtt{I}|\mathbf{0}] \begin{bmatrix} -\mathtt{R}^\top \mathbf{t} \\ 1 \end{bmatrix} \stackrel{P}{=} \mathtt{K}\mathtt{R}^\top \mathbf{t}.
$$

Epipole $\mathbf{e}'$ is the projection of the optic centre of C into camera C'.

$$
\mathbf{e}' \stackrel{P}{=} \mathtt{K}'[\mathtt{R}|\mathbf{t}] \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} = \mathtt{K}'\mathbf{t}.
$$

iii) **Prove that the epipoles e and e′ are the right and left null-spaces respectively of the fundamental matrix $\mathtt{F}$.**
We have to show that $\mathtt{F}\mathbf{e} = \mathbf{0}$ and $\mathtt{F}^\top \mathbf{e}' = \mathbf{0}$, respectively.

$$
\mathtt{F}\mathbf{e} \stackrel{P}{=} [\mathtt{K}']^{-\top}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1}\mathbf{e} = [\mathtt{K}']^{-\top}[\mathbf{t}]_\times \mathtt{R}\mathtt{K}^{-1}\mathtt{K}\mathtt{R}^\top \mathbf{t} = [\mathtt{K}']^{-\top}[\mathbf{t}]_\times \mathbf{t} = \mathbf{0}
$$

$$
\mathtt{F}^\top \mathbf{e}' \stackrel{P}{=} [\mathtt{K}]^{-\top}\mathtt{R}^\top [\mathbf{t}]_\times{}^\top [\mathtt{K}']^{-1} \ \mathtt{K}'\mathbf{t} = [\mathtt{K}]^{-\top}\mathtt{R}^\top [\mathbf{t}]_\times{}^\top \mathbf{t}
$$

But $[\mathbf{t}]_\times{}^\top = -[\mathbf{t}]_\times$ as it is skew-symmetric. So

$$
\mathtt{F}^\top \mathbf{e}' = [\mathtt{K}]^{-\top}\mathtt{R}^\top - [\mathbf{t}]_\times \mathbf{t} = \mathbf{0}
$$

(c) **A camera rotates abouts its optical centre and changes its intrinsics so that the camera matrices before and after $\mathtt{P} = \mathtt{K}[\mathtt{I}|\mathbf{0}]$ and $\mathtt{P}' = \mathtt{K}'[\mathtt{R}|\mathbf{0}]$**

i) **Derive an expression for the homography $\mathtt{H}$ which relates the images of points before and after the motion as $\mathbf{x}' = \mathtt{H}\mathbf{x}$.**

From part (a),

$$\mathbf{x} = \mathtt{K}\mathbf{X}^{W}_{3\times 1} \quad \text{and} \quad \mathbf{x}' = \mathtt{K}'\mathtt{R}\mathbf{X}^{W}_{3\times 1}$$

Eliminate $\mathbf{X}^{W}_{3\times 1}$

$$\mathbf{x}' \stackrel{P}{=} \mathtt{K}'\mathtt{R}\mathtt{K}^{-1}\mathbf{x} = \mathtt{H}\mathbf{x}$$

The matrix $\mathtt{H}$ is a 3×3 matrix with 8 d.o.f.

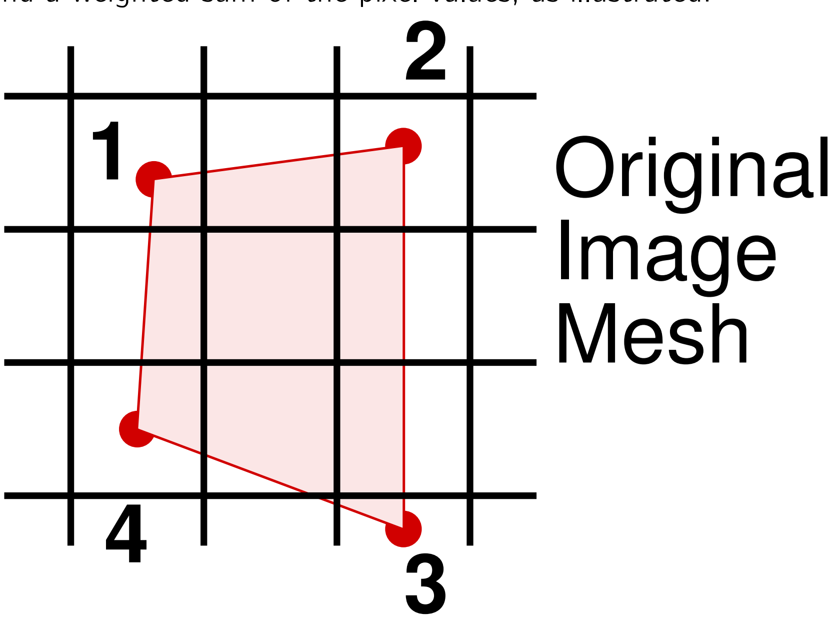
ii) **How might this formula be applied to rectify images from a stereo camera rig?**

You want to work out the pixel values for square pixels in the rectified image. The trick is to perform the inverse of the above.

- For each pixel i, j in the **rectified** image ...
- Consider its four corners at (i, j), $(i, j+1)$, $(i+1, j+1)$, $(i+1, j)$
- Derive their positions in the original image

$$\mathbf{x}_{1,2,3,4} = \mathtt{H}^{-1} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix}, \begin{bmatrix} i \\ j+1 \\ 1 \end{bmatrix}, \begin{bmatrix} i+1 \\ j+1 \\ 1 \end{bmatrix}, \begin{bmatrix} i+1 \\ j \\ 1 \end{bmatrix}$$

- Dehomogenize into $(x, y)_{1,2,3,4}$
- Find a weighted sum of the pixel values, as illustrated.



6. (a) **Find $\mathtt{F}$, books closed**. Already done.

(b) Using three vectors, and for pure rotation

$$\lambda\mathbf{x} = \mathtt{K}\mathbf{X} \qquad \text{and } \lambda'\mathbf{x}' = \mathtt{K}'(\mathtt{R}\mathbf{X}+\mathbf{t})\ . \text{ becomes } \lambda'\mathbf{x}' = \mathtt{K}'\mathtt{R}\mathbf{X}$$

and hence, inverting the first expression in order to replace $\mathbf{X}$,

$$\lambda'\mathbf{x}' = \lambda\mathtt{K}'\mathtt{R}\mathtt{K}^{-1}\mathbf{x} \quad \Rightarrow \quad \mathbf{x}' \stackrel{P}{=} \mathtt{H}\mathbf{x} \quad \text{with} \quad \mathtt{H} \stackrel{P}{=} \mathtt{K}'\mathtt{R}\mathtt{K}^{-1}$$

(c) **The camera's motion and zoom change are again as in part (a), but the scene viewed is now a planar surface $\hat{\mathbf{n}}^\top\mathbf{X} = d$, where $\hat{\mathbf{n}}$ is the plane's unit normal.**

i) **Using the projection equation for the first position show that**

$$\frac{\mathbf{X}}{\lambda} = \mathtt{K}^{-1}\mathbf{x} \quad \text{and} \quad \frac{\mathbf{t}}{\lambda} = \frac{\mathbf{t}\hat{\mathbf{n}}^\top\mathtt{K}^{-1}\mathbf{x}}{d}\ .$$

We haven't seen this before, so just obey the instructions ... Use the first projection equation

$$\lambda\mathbf{x} = \mathtt{K}\mathbf{X} \ \Rightarrow\ \lambda\mathtt{K}^{-1}\mathbf{x} = \mathbf{X} \ \Rightarrow\ \frac{\mathbf{X}}{\lambda} = \mathtt{K}^{-1}\mathbf{x}$$

Now introduce the equation of the plane ...

$$\frac{\hat{\mathbf{n}}^\top\mathbf{X}}{\lambda} = \frac{d}{\lambda} = \hat{\mathbf{n}}^\top\mathtt{K}^{-1}\mathbf{x} \quad \Rightarrow \quad \frac{1}{\lambda} = \frac{\hat{\mathbf{n}}^\top\mathtt{K}^{-1}\mathbf{x}}{d} \quad \Rightarrow \quad \frac{\mathbf{t}}{\lambda} = \frac{\mathbf{t}\,\hat{\mathbf{n}}^\top\mathtt{K}^{-1}\mathbf{x}}{d}$$

ii) **Using that for the second position, develop an expression for $\dfrac{\lambda'\mathbf{x}'}{\lambda}$**. Again just follow the instructions ...

$$\lambda'\mathbf{x}' = \mathtt{K}'(\mathtt{R}\mathbf{X}+\mathbf{t}) \quad \Rightarrow \quad \frac{\lambda'\mathbf{x}'}{\lambda} = \mathtt{K}'\left(\mathtt{R}\frac{\mathbf{X}}{\lambda} + \frac{\mathbf{t}}{\lambda}\right)$$

Suddendly you see why we being asked in the previous part ...

$$\frac{\lambda'\mathbf{x}'}{\lambda} = \mathtt{K}'\left(\mathtt{R}\mathtt{K}^{-1}\mathbf{x} + \frac{\mathbf{t}\,\hat{\mathbf{n}}^\top\mathtt{K}^{-1}\mathbf{x}}{d}\right)$$

and show that the corresponding points are again related by an homography

$$\mathbf{x}' \stackrel{P}{=} \mathtt{K}'\left(\mathtt{R}\mathtt{K}^{-1} + \frac{\mathbf{t}\,\hat{\mathbf{n}}^\top\mathtt{K}^{-1}}{d}\right)\mathbf{x} \quad \Rightarrow \quad \mathtt{J} \stackrel{P}{=} \mathtt{K}'\left(\mathtt{R}\mathtt{K}^{-1} + \frac{\mathbf{t}\,\hat{\mathbf{n}}^\top\mathtt{K}^{-1}}{d}\right)$$

(d) **Show that if $\mathbf{x}' = \mathtt{H}\mathbf{x}$, then $\mathtt{H}^\top\mathtt{F}$ must be an antisymmetric matrix.**

$$\mathbf{x}'^\top\mathtt{F}\mathbf{x} = 0 \quad \Rightarrow \quad \mathbf{x}^\top\mathtt{H}^\top\mathtt{F}\mathbf{x} = 0$$

Now consider the scalar quantity $\mathbf{x}^\top\mathtt{A}\mathbf{x}$, where $\mathtt{A}$ is antisymmetric. The transpose of a scalar is the same scalar, so

$$\mathbf{x}^\top\mathtt{A}\mathbf{x} = [\mathbf{x}^\top\mathtt{A}\mathbf{x}]^\top = \mathbf{x}^\top\mathtt{A}^\top\mathbf{x}$$

But $\mathtt{A}^\top = -\mathtt{A}$, so

$$\mathbf{x}^\top\mathtt{A}\mathbf{x} = -\mathbf{x}^\top\mathtt{A}\mathbf{x} \qquad \Rightarrow \ \mathbf{x}^\top\mathtt{A}\mathbf{x} = 0\ .$$

As $\mathbf{x}^\top\mathtt{H}^\top\mathtt{F}\mathbf{x} = 0$, then $\boxed{\mathtt{H}^\top\mathtt{F} \text{ is antisymmetric.}}$

7. Lecture 4: 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.**

As per notes. Invalid for (i) Discontinuous surfaces (ii) Tranparent surfaces

(b) **Prove that ZNCC is invariant under $I' = \alpha I + \beta$.**

Assume the original intensity, and let the source patch be A. The patch mean is subtracted from each pixel $\tilde{A}_{ij} = A_{ij} - \mu_A$ and the values used in

$$ZNCC = \frac{\sum_i \sum_j \tilde{A}_{ij}\tilde{B}_{ij}}{\sqrt{\sum_i \sum_j \tilde{A}_{ij}^2}\sqrt{\sum_i \sum_j \tilde{B}_{ij}^2}}$$

Now consider altering the intensities of each pixel as in $I' = \alpha I + \beta$. The patch mean is changed to

$$\mu'_A = \alpha\mu_A + \beta$$

so that subtracting the patch mean gives

$$\begin{aligned}
\tilde{A}'_{ij} &= A'_{ij} - \mu'_A \\
&= \alpha A_{ij} + \beta - (\alpha\mu_A + \beta) \\
&= \alpha(A_{ij} - \mu_A) \\
&= \alpha\tilde{A}_{ij}
\end{aligned}$$

Now

$$\begin{aligned}
ZNCC' = \frac{\sum_i \sum_j \tilde{A}'_{ij}\tilde{B}_{ij}}{\sqrt{\sum_i \sum_j \tilde{A}'^2_{ij}}\sqrt{\sum_i \sum_j \tilde{B}_{ij}^2}} &= \frac{\alpha \sum_i \sum_j \tilde{A}_{ij}\tilde{B}_{ij}}{\sqrt{\alpha^2 \sum_i \sum_j \tilde{A}_{ij}^2}\sqrt{\sum_i \sum_j \tilde{B}_{ij}^2}} \\
&= ZNCC
\end{aligned}$$

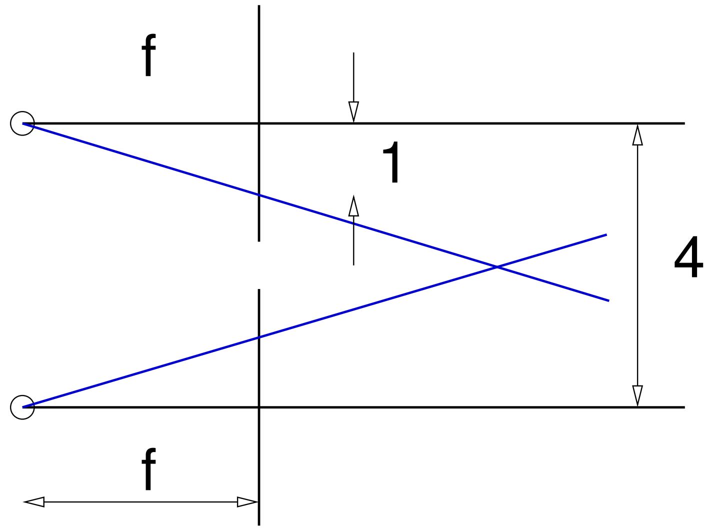
as the α's cancel top and bottom.

(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.**

In a monocular sequence images are taken a short time apart (typically 30ms) by the same camera. Any change in gain and offset over that time is likely to be neglible.

Stereo involves different camera which could quite easily have different gains and offsets.

8. Lecture 4: Triangulation



(a) **When the cameras are separated by $t_x = 4$ units, find the coordinates of the 3D point for the correspondences: (i)** $[-1, 0] \leftrightarrow [1, 0]$**; (ii)** $[0, 0] \leftrightarrow [0, 0]$**.**

Easy to see that

$$\mathbf{X}_1 = \begin{bmatrix} -2 \\ 0 \\ 2f \\ 1 \end{bmatrix} \qquad \mathbf{X}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

The latter is a point at infinity.

(b) i) **Assuming the camera matrices are $\mathtt{K}[\mathtt{I}|\mathbf{0}]$ and $\mathtt{K}[\mathtt{R}|\mathbf{t}]$, derive a vector expression for the position $\mathbf{X}$ of the 3D scene point given a correspondence $\mathbf{x} \leftrightarrow \mathbf{x}'$, assuming no noise.**

We are assuming no noise in the image, so that

$$\mathbf{X}_{4\times1} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} + \alpha \begin{bmatrix} \mathtt{K}^{-1}\mathbf{x} \\ 0 \end{bmatrix} = \begin{bmatrix} -\mathtt{R}^\top\mathbf{t} \\ 1 \end{bmatrix} + \beta \begin{bmatrix} \mathtt{R}^\top\mathtt{K}'^{-1}\mathbf{x}' \\ 0 \end{bmatrix}$$

Now just relax to 3-vectors, so

$$\mathbf{X}_{3\times1} = \alpha\mathtt{K}^{-1}\mathbf{x} = -\mathtt{R}^\top\mathbf{t} + \beta\mathtt{R}^\top\mathtt{K}'^{-1}\mathbf{x}'$$

We want to find either α or β. To find α, it saves work to bring the rotation to the LHS

$$\alpha\mathtt{R}\mathtt{K}^{-1}\mathbf{x} = -\mathbf{t} + \beta\mathtt{K}'^{-1}\mathbf{x}' \quad \text{or} \quad \alpha\mathbf{p} = -\mathbf{t} + \beta\mathbf{q}$$

Then dot this equation first with $\mathbf{p}$ then $\mathbf{q}$ to get two simultaneous equations in α and β, and solve for

$$\alpha = \frac{(\mathbf{p}^\top\mathbf{q})(\mathbf{q}^\top\mathbf{t}) - (\mathbf{q}^\top\mathbf{q})(\mathbf{p}^\top\mathbf{t})}{(\mathbf{p}^\top\mathbf{p})(\mathbf{q}^\top\mathbf{q}) - (\mathbf{p}^\top\mathbf{q})^2}$$

ii) **Verify this expression using the parallel camera configuration earlier in this question**

$$\mathtt{R} = \mathtt{I} \quad \mathbf{t} = [4, 0, 0]^\top \quad \mathtt{K}^{-1} = \mathtt{K}'^{-1} = \mathrm{Diag}(1/f, 1/f, 1)$$

Using the corresponding points

$$\mathbf{x} = [-1, 0, 1]^\top \quad \mathbf{x}' = [1, 0, 1]^\top$$

you should find

$$\mathbf{p} = \mathtt{R}\mathtt{K}^{-1}\mathbf{x} = [-1/f, 0, 1]^\top \quad \mathbf{q} = [1/f, 0, 1]^\top$$

Using our formula for α we find $\alpha = 2f$.
So for the parallel cameras, $\mathbf{X}_i = \alpha \mathtt{K}^{-1}\mathbf{x} = [-2, 0, 2f]^\top$ — just as before. Which is nice.

(c) **For a parallel camera stereo configuration the "horizontal" disparity d is given by $d = x' - x = ft_x/Z$. Hence show that the error δZ in depth Z corresponding to an error δd in disparity is given by $\delta Z = -(\delta d/ft_x)Z^2$**
We are told $Z = ft_x\frac{1}{d}$ so

$$\delta Z = -ft_x \frac{1}{d^2}\,\delta d \;=\; -ft_x \frac{Z^2}{(ft_x)^2}\,\delta d \;=\; -\frac{Z^2}{ft_x}\,\delta d$$

What are the consequences of this relation when computing a stereo reconstruction?
The error in disparity δd is likely to be constant across the image — it depends principally on uncertainties in the location of image features which are independent of depth. As f and t_x are constant, the uncertainty in depth is proportional to the depth squared.

---

9. For interest only.