Elements of Geometry for Computer Vision and Computer Graphics

Translation of Euclid's Elements by Adelardus Bathensis (1080–1152)

2021 Lecture 3

Tomas Pajdla

pajdla@cvut.cz

Sunday 14th February, 2021

6 Camera calibration

Let us now look at a useful interpretation of image projection matrix in space and image equipped with a cartesian coordinate systems.

6.1 Camera pose

The projection formula 5.10 reveals that the perspective projection depends on matrix $\mathtt{A}$ and vector $\vec{C}_\delta$. The vector $\vec{C}_\delta$ represents the position of the camera projection center w.r.t. the world coordinate system. Columns of matrix $\mathtt{A}$ are coordinates of the basic vectors of δ in the basis β

$$\mathtt{A} = \begin{bmatrix} \vec{d}_{1_\beta} & \vec{d}_{2_\beta} & \vec{d}_{3_\beta} \end{bmatrix} \qquad \delta \to \beta \tag{6.1}$$

To recover the orientation of the camera, we will introduce the focal length f as the distance of the camera projection center C from its projection plane π (in the world units) and replace the product $f\,\mathtt{A}$ by the product of two 3×3 matrices $\mathtt{K}$ and $\mathtt{R}$

internal param.

$$f\,\mathtt{A} = \mathtt{K}\,\mathtt{R} \leftarrow \text{rotation} \tag{6.2}$$

We will see that this seemingly artificial construction is indeed justified.

Rotation matrix $\mathtt{R}$ determines the orientation of the camera in space and altogether with $\vec{C}_\delta$ defines the *camera pose*. The *camera calibration matrix* $\mathtt{K}$ does not change when moving its camera in the space.

To obtain $\mathtt{K}$ and $\mathtt{R}$, we define, Figure 6.1, the *camera cartesian coordinate system* (C, γ) with center (again) in the camera projection center C and with

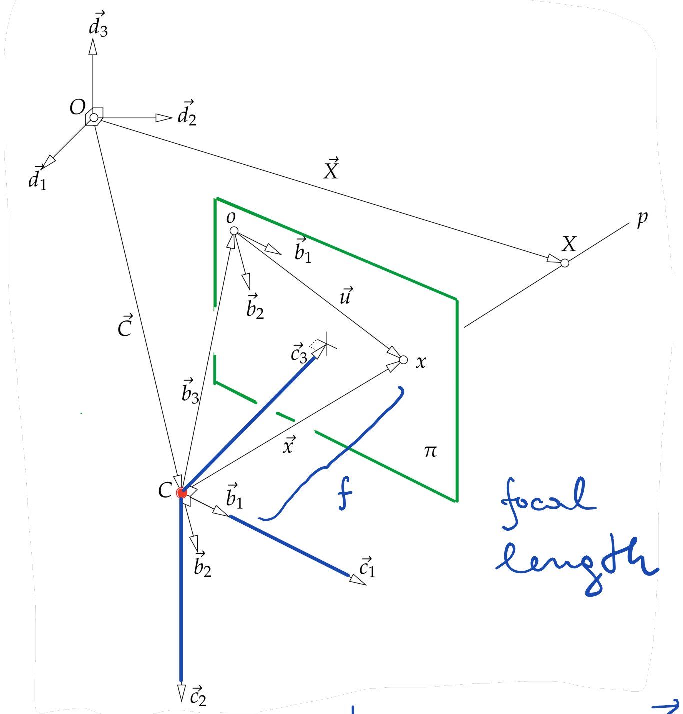


Image projection

$$\eta \begin{bmatrix} \vec{u}_\alpha \\ 1 \end{bmatrix} = \mathtt{A} \begin{bmatrix} \mathtt{I} \mid -\vec{C}_\delta \end{bmatrix} \begin{bmatrix} \vec{X}_\delta \\ 1 \end{bmatrix} \qquad \mathtt{P}_\beta = \begin{bmatrix} \underbrace{\mathtt{A}}_{3\times 3} & \mid & \underbrace{-\overbrace{\mathtt{A}\vec{C}_\delta}^{\vec{C}_\beta}}_{3\times 1} \end{bmatrix}$$

Camera cartesian coordinate system

$$(C, \gamma)$$

$$\gamma = [\vec{c}_1, \vec{c}_2, \vec{c}_3]$$

basis $\gamma = [\vec{c}_1, \vec{c}_2, \vec{c}_3]$ such that

$$\begin{aligned} \vec{c}_1 &= k_{11}\,\vec{b}_1 \\ \vec{c}_2 &= k_{12}\,\vec{b}_1 + k_{22}\,\vec{b}_2 \\ \vec{c}_3 &= k_{13}\,\vec{b}_1 + k_{23}\,\vec{b}_2 + 1\,\vec{b}_3 \end{aligned} \tag{6.3}$$

Parameters k_{ij} are determined to make the basis γ orthogonal. Notice that vector $\vec{c}_3$ is orthogonal to π since it is orthogonal to $\vec{c}_1, \vec{c}_2$, which span π, by construction. Also notice that γ is (in general) not an orthonormal basis since the length of its vectors equals the distance of C from π, i.e. the *focal length* f in the world units.

Equations 6.3 define matrix $\mathtt{K}$ as

$$\mathtt{K} = \begin{bmatrix} \vec{c}_{1_\beta} & \vec{c}_{2_\beta} & \vec{c}_{3_\beta} \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & k_{22} & k_{23} \\ 0 & 0 & 1 \end{bmatrix} \tag{6.4}$$

By this construction, we have

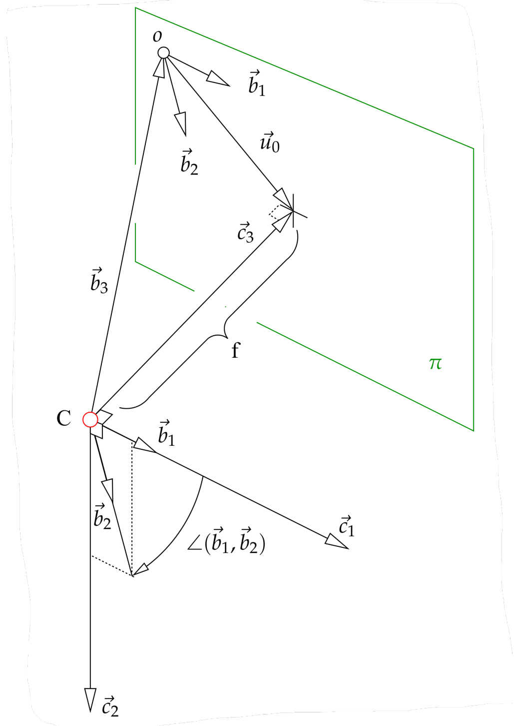
$$\vec{x}_\beta = \mathtt{A}\,\vec{x}_\delta = \mathtt{K}\,\vec{x}_\gamma \tag{6.5}$$

$$\vec{x}_\gamma = = \frac{1}{f}\mathtt{R}\,\vec{x}_\delta \tag{6.6}$$

The world cartesian coordinate system has basic vectors of unit length. The camera cartesian coordinate system (C, γ) has basic vectors of length equal to f. Therefore,

$$\begin{bmatrix} \vec{d}_{1_\gamma} & \vec{d}_{2_\gamma} & \vec{d}_{3_\gamma} \end{bmatrix} = \frac{1}{f}\,\mathtt{R} = \begin{bmatrix} \mathbf{r}_1^\top/f \\ \mathbf{r}_2^\top/f \\ \mathbf{r}_3^\top/f \end{bmatrix} \tag{6.7}$$

for some 3×3 orthonormal matrix $\mathtt{R}$ with rows $\mathbf{r}_1^\top, \mathbf{r}_2^\top, \mathbf{r}_3^\top$.



$$\begin{tikzcd} \delta \arrow[r, "\frac{1}{f}R"] \arrow[rr, bend right, "A"'] & \gamma \arrow[r, "K"] & \beta \end{tikzcd}$$

Consider that

$$\mathtt{A} = \begin{bmatrix} \vec{d}_{1_\beta} & \vec{d}_{2_\beta} & \vec{d}_{3_\beta} \end{bmatrix} = \mathtt{K} \begin{bmatrix} \vec{d}_{1_\gamma} & \vec{d}_{2_\gamma} & \vec{d}_{3_\gamma} \end{bmatrix} = \frac{1}{f}\mathtt{K}\,\mathtt{R} \tag{6.8}$$

We can view the matrices $\frac{1}{f}\mathtt{R}$ and $\mathtt{K}$ as coordinate transformation matrices, which transform a general vector $\vec{y}$ from the coordinates w.r.t. δ to γ and then to β, i.e.

$$\boxed{\vec{y}_\beta = \mathtt{K}\,\vec{y}_\gamma = \frac{1}{f}\mathtt{K}\,\mathtt{R}\,\vec{y}_\delta} \tag{6.9}$$

$$\begin{tikzcd} \delta \arrow[r, "\frac{1}{f}R"] \arrow[rr, bend right, "A"'] & \gamma \arrow[r, "K"] & \beta \end{tikzcd}$$

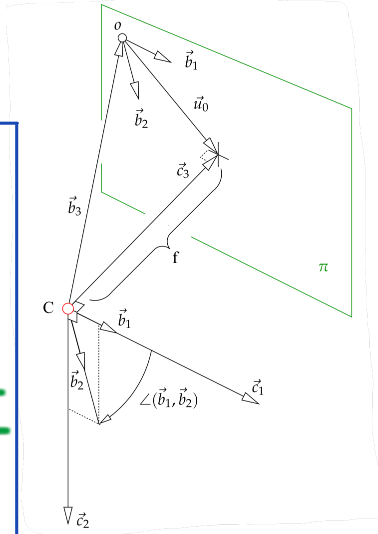
The basis γ is orthogonal and all basic vectors have the same length, which is equal to f. It follows from the orthogonality of the basis γ that $\vec{c}_1 \cdot \vec{c}_1 = f^2$, $\vec{c}_1 \cdot \vec{c}_2 = 0$ and $\vec{c}_2 \cdot \vec{c}_2 = f^2$ and hence using Equation 6.3 leads, for a positive f, to

$$\begin{aligned} k_{11}\,\|\vec{b}_1\| - f &= 0 \\ k_{11}^2\,k_{22}\,(\vec{b}_1 \cdot \vec{b}_2) + k_{12}\,f^2 &= 0 \\ k_{11}^2\,k_{22}^2\,\|\vec{b}_2\|^2 - (k_{12}^2 + k_{11}^2)\,f^2 &= 0 \end{aligned} \tag{6.10}$$

Let us solve Equations 6.10 for k_{11}, k_{12} and k_{22}. The first equation in (6.10) provides k_{11}. Substituting the square of f from the first equation into the second one and dividing it by k_{11}^2 gives the second equation of (6.11), which allows to compute k_{12} from k_{22}. To get k_{22}, we construct the third equation of (6.11) as follows. We express k_{11} from the first equation of (6.10) and k_{12} from the second equation of (6.11) and substitute them into the third equation of (6.10), which we then multiply by $||\vec{b}_1||^4/f^2$. Altogether, we get

$$\begin{aligned} k_{11}\,\|\vec{b}_1\| - f &= 0 \\ k_{12}\,\|\vec{b}_1\|^2 + k_{22}\,(\vec{b}_1 \cdot \vec{b}_2) &= 0 \\ k_{22}^2\,(\|\vec{b}_1\|^2\,\|\vec{b}_2\|^2 - (\vec{b}_1 \cdot \vec{b}_2)^2) - f^2\,\|\vec{b}_1\|^2 &= 0 \end{aligned} \tag{6.11}$$

Geometrical interpretation of K



$$\mathtt{K} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & k_{22} & k_{23} \\ 0 & 0 & 1 \end{bmatrix}$$

$\gamma = [\vec{c}_1, \vec{c}_2, \vec{c}_3]$ is orthogonal

$\Downarrow$

$$\begin{aligned} \vec{c}_1 \cdot \vec{c}_1 &= f^2 \\ \vec{c}_1 \cdot \vec{c}_2 &= 0 \\ \vec{c}_2 \cdot \vec{c}_2 &= f^2 \end{aligned}$$

$$\begin{aligned} \vec{c}_1 &= k_{11}\,\vec{b}_1 \\ \vec{c}_2 &= k_{12}\,\vec{b}_1 + k_{22}\,\vec{b}_2 \\ \vec{c}_3 &= k_{13}\,\vec{b}_1 + k_{23}\,\vec{b}_2 + 1\,\vec{b}_3 \end{aligned}$$

Looking at the third equation of (6.11) we see that

$$k_{22}^2 = \frac{f^2\|\vec{b}_1\|^2}{\|\vec{b}_1\|^2\|\vec{b}_2\|^2 - (\vec{b}_1\cdot\vec{b}_2)^2} = \frac{f^2}{\|\vec{b}_2\|^2 - \|\vec{b}_2\|^2\cos^2\angle(\vec{b}_1,\vec{b}_2)} \tag{6.12}$$

and since γ was constructed to make k_{22} positive, we obtain

$$\boxed{k_{22} = \frac{f}{\|\vec{b}_2\|\sin\angle(\vec{b}_1,\vec{b}_2)}} \tag{6.13}$$

$$K = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & k_{22} & k_{23} \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} k_{11} & 0 & \cdot \\ 0 & k_{11} & \cdot \\ 0 & 0 & 1 \end{bmatrix} \qquad \square$$

The second equation of (6.10) now gives

$$k_{12} = -k_{22}\frac{\vec{b}_1\cdot\vec{b}_2}{\|\vec{b}_1\|^2} = -k_{22}\frac{\|\vec{b}_2\|\cos\angle(\vec{b}_1,\vec{b}_2)}{\|\vec{b}_1\|} \tag{6.14}$$

$$= -\frac{f\,\cos\angle(\vec{b}_1,\vec{b}_2)}{\|\vec{b}_1\|\sin\angle(\vec{b}_1,\vec{b}_2)} \tag{6.15}$$

$$\vec{c}_3 = \vec{u}_0 + \vec{b}_3 \Rightarrow \vec{u}_0 = \vec{c}_3 - \vec{b}_3$$

$$\vec{u}_{0\beta} = (\vec{c}_3 - \vec{b}_3)_\beta = K(\vec{c}_{3\gamma} - \vec{b}_{3\gamma})$$

$$= \left(\begin{bmatrix} k_{13} \\ k_{23} \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} \vec{u}_{0\alpha} \\ 1 \end{bmatrix}$$

$$\Rightarrow \vec{u}_{0\alpha} = \begin{bmatrix} k_{13} \\ k_{23} \end{bmatrix}$$

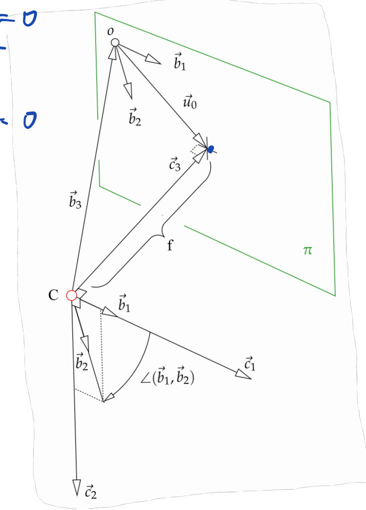
$$\begin{aligned} \vec{c}_1 &= k_{11}\,\vec{b}_1 \\ \vec{c}_2 &= k_{12}\,\vec{b}_1 + k_{22}\,\vec{b}_2 \\ \vec{c}_3 &= k_{13}\,\vec{b}_1 + k_{23}\,\vec{b}_2 + 1\,\vec{b}_3 \end{aligned}$$

Rect. $\vec{b}_2$ $\square$ $\vec{b}_1$ $k_{12} = 0$

Finally k_{11} follows from (6.11)

$$\boxed{k_{11} = \frac{f}{\|\vec{b}_1\|}} \tag{6.16}$$

Square $\square$ $k_{12} = 0$



principal point

Considering Figure 6.1 and Equation 6.3, we see that the coordinates of the vector $\vec{u}_0$, corresponding to the *principal point*, which is the perpendicular projection of C onto π, are in β

$$\vec{u}_{0\beta} = \begin{bmatrix} k_{13} \\ k_{23} \\ 0 \end{bmatrix}, \text{ i.e. } \vec{u}_{0\alpha} = \begin{bmatrix} k_{13} \\ k_{23} \end{bmatrix} \tag{6.17}$$

The horizontal pixel size corresponds to $\|\vec{b}_1\|$. Quantity k_{11} can thus be understood as f expressed in the horizontal image units. The angle between the image axes $\vec{b}_1, \vec{b}_2$ is obtained from $k_{12}/k_{11} = -\operatorname{cotan}\angle(\vec{b}_1, \vec{b}_2)$. The ratio of the lengths of the image axes is determined by $\|\vec{b}_2\|/\|\vec{b}_1\| = \sqrt{k_{11}\,(k_{11}^2 + k_{12}^2)}/k_{22}$.

Let us now return to Equation 5.11 and substitute there the above results to arrive at the final projection equation

$$\eta\,\vec{x}_\beta = \mathtt{P}_\beta \begin{bmatrix}\vec{X}_\delta\\1\end{bmatrix} \tag{6.18}$$

$$\eta \begin{bmatrix}\vec{u}_\alpha\\1\end{bmatrix} = \mathtt{A}\,(\vec{X}_\delta - \vec{C}_\delta) \tag{6.19}$$

$$f\,\eta \begin{bmatrix}\vec{u}_\alpha\\1\end{bmatrix} = f\;\mathtt{A}\,(\vec{X}_\delta - \vec{C}_\delta) \tag{6.20}$$

$$f\,\eta \begin{bmatrix}\vec{u}_\alpha\\1\end{bmatrix} = \mathtt{K}\,\mathtt{R}\,(\vec{X}_\delta - \vec{C}_\delta) \tag{6.21}$$

$f\,\mathtt{A} = \mathtt{K}\,\mathtt{R}$

$$\zeta \begin{bmatrix}\vec{u}_\alpha\\1\end{bmatrix} = \mathtt{K}\,\mathtt{R}\,(\vec{X}_\delta - \vec{C}_\delta) \tag{6.22}$$

Camera projection matrix

$$\zeta \begin{bmatrix}\vec{u}_\alpha\\1\end{bmatrix} = \mathtt{K}\,\mathtt{R}\,\left[\mathtt{I}\,|-\vec{C}_\delta\right] \begin{bmatrix}\vec{X}_\delta\\1\end{bmatrix} \tag{6.23}$$

Image projection matrix

We have introduced a new parameter $\zeta = f\,\eta$, which is the depth of X in the world units. We conclude that

$$\mathtt{P}_\beta = \left[\tfrac{1}{f}\,\mathtt{K}\,\mathtt{R}\,|-\tfrac{1}{f}\,\mathtt{K}\,\mathtt{R}\,\vec{C}_\delta\right] \tag{6.24}$$

Notice that the last row $\mathtt{a}_3^\top$ of $\mathtt{A}$ provides f since

$$\mathtt{A} = \begin{bmatrix}\mathtt{a}_1^\top\\ \mathtt{a}_2^\top\\ \mathtt{a}_3^\top\end{bmatrix} = \frac{1}{f}\begin{bmatrix}k_{11} & k_{12} & k_{13}\\ 0 & k_{22} & k_{23}\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\mathbf{r}_1^\top\\ \mathbf{r}_2^\top\\ \mathbf{r}_3^\top\end{bmatrix} = \frac{1}{f}\begin{bmatrix}k_{11}\mathbf{r}_1^\top + k_{12}\mathbf{r}_2^\top + k_{13}\mathbf{r}_3^\top\\ k_{22}\mathbf{r}_2^\top + k_{23}\mathbf{r}_3^\top\\ \mathbf{r}_3^\top\end{bmatrix} \tag{6.25}$$

$$\boxed{\|\mathtt{P}_\beta(3, 1:3)\| = \tfrac{1}{f}.}$$

and hence $\|\mathtt{a}_3^\top\| = \frac{1}{f}$. Therefore $\|\mathtt{P}_\beta(3,1:3)\| = \frac{1}{f}$.

Equation 6.23 is very important in many practical situations when we do not have access to physical dimensions of the camera but only to images. Then, it is possible to recover matrix $\mathtt{K}\,\mathtt{R}\left[\mathtt{I}\quad| - \vec{C}_\delta\right]$ but not image projection matrix $\mathtt{P}_\beta$. This is so important the we introduce the *camera projection matrix*

$$\boxed{\mathtt{P} = \left[\mathtt{K}\,\mathtt{R}\,|\, - \mathtt{K}\,\mathtt{R}\,\vec{C}_\delta\right]} \tag{6.26}$$

$\|P(3,1:3)\| = 1$

which is related to the image projection matrix as

$$\boxed{\mathtt{P} = f\,\mathtt{P}_\beta} \tag{6.27}$$

$Q = \left[\quad\right]$

$\|\quad\| = 1$

P

In this text, it would be more consistent to associate subscript ν with the camera projection matrix but we will not do that since we want to use the nomenclature of [13] here whenever possible.

Let us write $\mathtt{K}$ explicitly,

$$\mathtt{K} = \begin{bmatrix} \frac{f}{\|\vec{b}_1\|} & -\frac{f\,\cos\angle(\vec{b}_1,\vec{b}_2)}{\|\vec{b}_1\|\,\sin\angle(\vec{b}_1,\vec{b}_2)} & u_0 \\ 0 & \frac{f}{\|\vec{b}_2\|\,\sin\angle(\vec{b}_1,\vec{b}_2)} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6.28}$$

where $\vec{u}_{0\alpha} = \begin{bmatrix} u_0 & v_0 \end{bmatrix}^\top$. We see that we can neither recover f nor $\|\vec{b}_1\|$ from $\mathtt{P}$.

Let us introduce *image calibration matrix*

$$\mathtt{K}_\beta = \frac{1}{f}\,\mathtt{K} \tag{6.29}$$

to have

$$\mathtt{P}_\beta = \left[\mathtt{K}_\beta\,\mathtt{R}\,|\, - \mathtt{K}_\beta\,\mathtt{R}\,\vec{C}_\delta\right] \tag{6.30}$$

Writing image calibration matrix $\mathtt{K}_\beta$ explicitly,

$$\boxed{\mathtt{K}_\beta = \frac{1}{f}\,\mathtt{K} = \begin{bmatrix} \frac{1}{\|\vec{b}_1\|} & -\frac{\cos\angle(\vec{b}_1,\vec{b}_2)}{\|\vec{b}_1\|\sin\angle(\vec{b}_1,\vec{b}_2)} & \frac{u_0}{f} \\ 0 & \frac{1}{\|\vec{b}_2\|\sin\angle(\vec{b}_1,\vec{b}_2)} & \frac{v_0}{f} \\ 0 & 0 & \frac{1}{f} \end{bmatrix}} \tag{6.31}$$

shows that it is possible to recover both

$$\|\vec{b}_1\| = \frac{1}{\mathtt{K}_{\beta 11}} \text{ and } f = \frac{1}{\mathtt{K}_{\beta 33}} \tag{6.32}$$

from image calibration matrix.

There is an important difference between $\mathtt{K}_\beta$ and $\mathtt{K}$ regarding the representation of internal camera calibration information. Image calibration matrix $\mathtt{K}_\beta$, and also image projection matrix $\mathtt{P}_\beta$, captures all calibration information about a perspective image whereas camera calibration matrix $\mathtt{K}$, and also camera projection matrix $\mathtt{P}$, captures only the calibration information that can be recovered by auto-calibration from images as we will see later. When the focal length is known in world units or when pixel sizes are known in world units, it is more appropriate to use image calibration $\mathtt{K}_\beta$, or image projection matrix $\mathtt{P}_\beta$, to represent full internal calibration information.

$K = \begin{bmatrix} \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot \\ 0 & 0 & 1 \end{bmatrix}$

$K_\beta = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 & 1/f \end{bmatrix}$

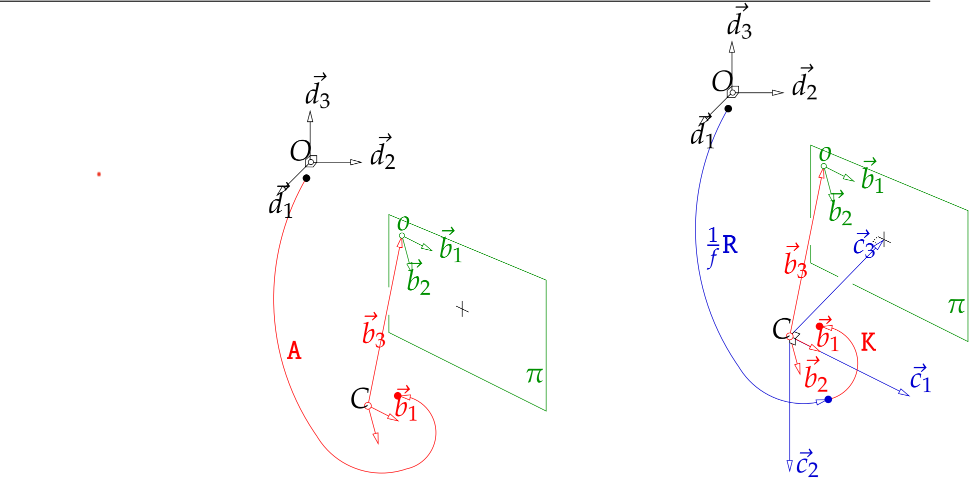
imag...

K_β, P_β

K, P

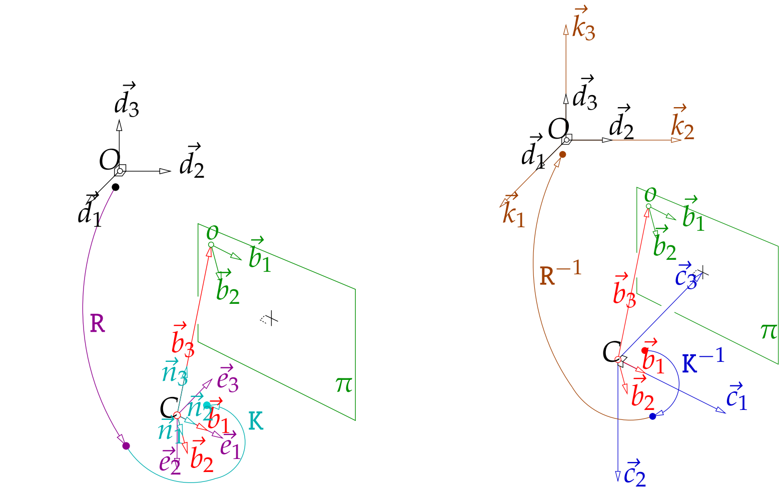
camera ...

Quiz



(a) $\beta = [\vec{b}_1, \vec{b}_2, \vec{b}_3]$, $\delta = [\vec{d}_1, \vec{d}_2, \vec{d}_3]$: $\vec{y}_\beta = \mathtt{A}\,\vec{y}_\delta$

(b) $\gamma = [\vec{c}_1, \vec{c}_2, \vec{c}_3]$: $\vec{y}_\gamma = \frac{1}{f}\mathtt{R}\,\vec{y}_\delta$
$\vec{y}_\beta = \mathtt{K}\,\vec{y}_\gamma$



(c) $\epsilon = [\vec{e}_1, \vec{e}_2, \vec{e}_3]$: $\vec{y}_\epsilon = \mathtt{R}\,\vec{y}_\delta$,
$\nu = [\vec{n}_1, \vec{n}_2, \vec{n}_3]$: $\vec{y}_\nu = \mathtt{K}\,\vec{y}_\epsilon$

(d) $\kappa = [\vec{k}_1, \vec{k}_2, \vec{k}_3]$: $\vec{y}_\gamma = \mathtt{K}^{-1}\,\vec{y}_\beta$,
$\vec{y}_\kappa = \mathtt{R}^{-1}\,\vec{y}_\gamma$

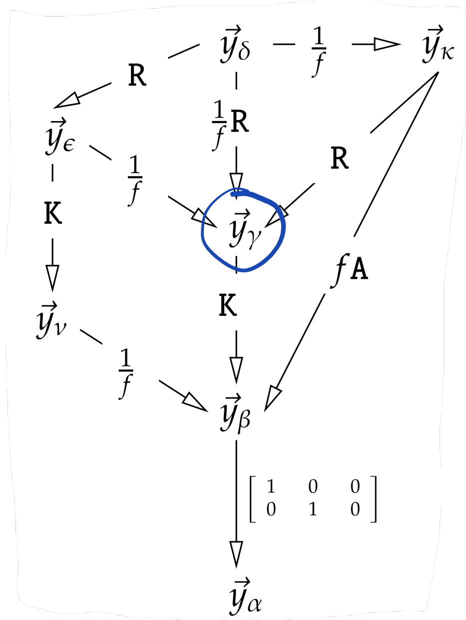
§1 Coordinate systems generated by applying $\mathtt{K}\,\mathtt{R}$ to $\vec{y}_\delta$ and $\mathtt{R}^{-1}\mathtt{K}^{-1}$ to $\vec{y}_\beta$ We have seen that the decomposition of $\mathtt{A}$ to $\mathtt{K}$ and $\mathtt{R}$ introduced the camera cartesian coordinate system (C, γ), Figure 6.2(b)

$$\boxed{\begin{aligned} \vec{y}_\gamma &= \frac{1}{f}\mathtt{R}\,\vec{y}_\delta \\ \vec{y}_\beta &= \mathtt{K}\,\vec{y}_\gamma \end{aligned}}$$

$\frac{1}{f}\mathtt{K}\mathtt{R}$

$\delta \rightarrow \gamma \rightarrow \beta$

Camera coordinate systems



There are three more coordinate systems to consider when looking at how matrices $\mathtt{R}$, $\mathtt{K}$, and their inverses $\mathtt{R}^{-1}$, $\mathtt{K}^{-1}$, apply to vectors $\vec{y}_\delta$ and $\vec{y}_\beta$, Figure 6.2.

Let us first consider coordinates of a vector $\vec{y}$ w.r.t. basis δ and apply successively $\mathtt{R}$ and $\mathtt{K}$. Coordinate vector $\mathtt{R}\,\vec{y}_\delta$ can be interpreted as coordinates of $\vec{y}$ w.r.t. a new basis $\epsilon = [\vec{e}_1, \vec{e}_2, \vec{e}_3]$, Figure 6.2(c). Applying further $\mathtt{K}$ to $\vec{y}_\epsilon$ gives the coordinate vector $\mathtt{K}\,\vec{y}_\epsilon$, which can be interpreted as $\vec{y}$ w.r.t. yet another new basis $\nu = [\vec{n}_1, \vec{n}_2, \vec{n}_3]$. We get from ν to β by using $\frac{1}{f}\,\mathtt{I}$

$$\vec{y}_\epsilon = \mathtt{R}\,\vec{y}_\delta \tag{6.35}$$
$$\vec{y}_\nu = \mathtt{K}\,\vec{y}_\epsilon \tag{6.36}$$
$$\vec{y}_\beta = \frac{1}{f}\mathtt{I}\,\vec{y}_\nu \tag{6.37}$$

We have introduced two new coordinate systems (C, ν), $\nu = [\vec{n}_1, \vec{n}_2, \vec{n}_3]$ and (C, ϵ), $\epsilon = [\vec{e}_1, \vec{e}_2, \vec{e}_3]$.

Next we consider coordinates of a vector $\vec{y}$ w.r.t. basis β and apply successively $\mathtt{K}^{-1}$ and $\mathtt{R}^{-1}$. Coordinate vector $\mathtt{K}^{-1}\,\vec{y}_\beta$ gives $\vec{y}_\gamma$. Coordinate vector $\mathtt{R}^{-1}\,\vec{y}_\gamma$ can be interpreted as coordinates of $\vec{y}$ w.r.t. a new basis $\kappa = [\vec{k}_1, \vec{k}_2, \vec{k}_3]$, Figure 6.2(d). To get from $\vec{y}_\kappa$ to $\vec{y}_\delta$ we need to employ $f\,\mathtt{I}$

$$\vec{y}_\gamma = \mathtt{K}^{-1}\,\vec{y}_\beta \tag{6.38}$$
$$\vec{y}_\kappa = \mathtt{R}^{-1}\,\vec{y}_\gamma \tag{6.39}$$
$$\vec{y}_\delta = f\,\mathtt{I}\,\vec{y}_\kappa \tag{6.40}$$

We have thus introduced a new coordinate system (O, κ), $\kappa = [\vec{k}_1, \vec{k}_2, \vec{k}_3]$.

Figure 6.3 summarizes the relationship between coordinates of a vector and between bases associated with a perspective camera.

We can now see why we have chosen to denote the image projection matrix as $\mathtt{P}_\beta$ and the camera projection matrix as $\mathtt{P}$. The image projection matrix provides the ray direction vector $\vec{x}$ in basis β while the camera projection matrix provides the ray direction vector $\vec{x}$ in basis ν.

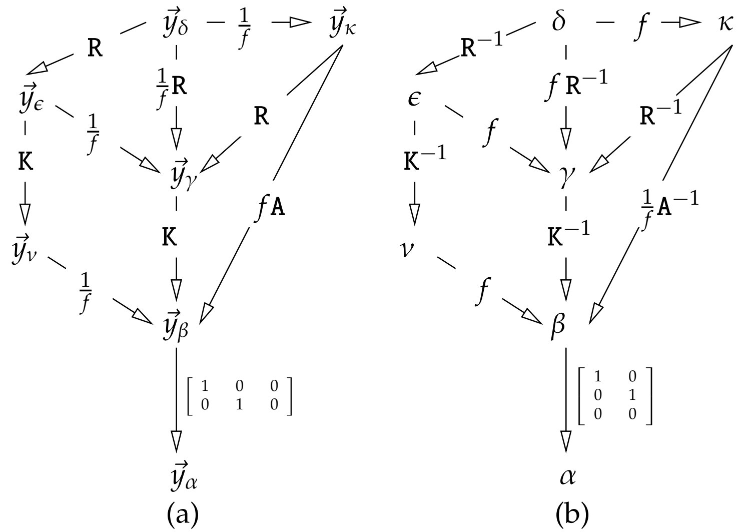


Figure 6.3: Relationships between (a) coordinates in different bases. e.g. $\vec{y}_\beta = \texttt{K}\,\vec{y}_\gamma$ and (b) bases themselves, e.g. $\beta = \gamma\,\texttt{K}^{-1}$, associated with a perspective camera.

§ 2 Recovering camera pose from its projection matrix Let us next consider that we have already computed the camera projection matrix

$$\texttt{Q} = \xi\,\texttt{P} = \xi\,\texttt{K}\,\texttt{R}\,[\texttt{I}\,|\,-\vec{C}_\delta] \tag{6.41}$$

consisting of a 3×3 matrix $\texttt{M}$ and 3×1 vector $\texttt{m}$

$$\texttt{Q} = [\texttt{M}\,|\,\texttt{m}] \tag{6.42}$$

To recover camera pose from $\texttt{Q}$, we need to get $\vec{C}_\delta$ from $\texttt{m}$ and to decompose $\texttt{Q}$ into the product of $\texttt{K}$ in the form of (6.4) and $\texttt{R}$ such that $\texttt{R}^\top\texttt{R} = \texttt{I}$ and

$|\mathtt{R}| = 1$. Consider $\mathtt{M}$ in the form

$$\mathtt{M} = \begin{bmatrix} \mathtt{m}_1^\top \\ \mathtt{m}_2^\top \\ \mathtt{m}_3^\top \end{bmatrix} \tag{6.43}$$

Next we notice that the last row of $\mathtt{K}\,\mathtt{R}$ has unit norm since it is equal to the last row of rotation $\mathtt{R}$. Therefore, we need to divide $\mathtt{M}$ by the norm of its last row to get a matrix decomposable into the product of $\mathtt{K}\,\mathtt{R}$. Moreover, it follows from the construction of β that $k_{11} > 0$ and $k_{22} > 0$. Thus, determinant $|\mathtt{K}\,\mathtt{R}| = |\mathtt{K}|\,|\mathtt{R}| = k_{11}\,k_{22} > 0$. Therefore, we also need to multiply $\mathtt{M}$ by the sign of its determinant to get a matrix decomposable into $\mathtt{K}\,\mathtt{R}$.

$$\frac{\operatorname{sign}|\mathtt{M}|}{\|\mathtt{m}_3\|}\mathtt{M} = \frac{\operatorname{sign}|\mathtt{M}|}{\|\mathtt{m}_3\|}\begin{bmatrix} \mathtt{m}_1^\top \\ \mathtt{m}_2^\top \\ \mathtt{m}_3^\top \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & k_{22} & k_{23} \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \mathtt{r}_1^\top \\ \mathtt{r}_2^\top \\ \mathtt{r}_3^\top \end{bmatrix} \tag{6.44}$$

which provides the following set of equations

polynom. of

$$\frac{\mathtt{m}_2^\top \mathtt{m}_3}{\|\mathtt{m}_3\|^2} = k_{22}\,\underbrace{\mathtt{r}_2^\top \mathtt{r}_3}_{=0} + k_{23}\,\underbrace{\mathtt{r}_3^\top \mathtt{r}_3}_{\|\mathtt{r}_3\|^2 = 1} = k_{23} \tag{6.45}$$

$$\frac{\mathtt{m}_1^\top \mathtt{m}_3}{\|\mathtt{m}_3\|^2} = k_{13} \tag{6.46}$$

$$\frac{\mathtt{m}_2^\top \mathtt{m}_2}{\|\mathtt{m}_3\|^2} = k_{22}^2 + k_{23}^2 \tag{6.47}$$

$$\frac{\mathtt{m}_1^\top \mathtt{m}_2}{\|\mathtt{m}_3\|^2} = k_{12}\,k_{22} + k_{13}\,k_{23} \tag{6.48}$$

$$\frac{\mathtt{m}_1^\top \mathtt{m}_1}{\|\mathtt{m}_3\|^2} = k_{11}^2 + k_{12}^2 + k_{13}^2 \tag{6.49}$$

from which $k_{11}, k_{12}, k_{13}, k_{22}, k_{23}$ can be easily computed considering that the most of consumer digital cameras have $k_{11} > 0, k_{22} > 0, k_{13} > 0, k_{23} > 0$.

By solving

§ 2 Recovering camera pose from its projection matrix Let us next consider that we have already computed the camera projection matrix

$$\boxed{\mathtt{Q} = \xi\,\mathtt{P}} = \xi\,\mathtt{K}\,\mathtt{R}\,[\mathtt{I}\,|\,-\vec{C}_\delta] \tag{6.41}$$

consisting of a 3×3 matrix $\mathtt{M}$ and 3×1 vector $\mathtt{m}$

$$\boxed{\mathtt{Q} = [\mathtt{M}\,|\,\mathtt{m}]} = \begin{bmatrix} \underbrace{\boxed{3\times 3}}_{M} & \underbrace{|}_{m} \end{bmatrix} \tag{6.42}$$

To recover camera pose from $\mathtt{Q}$, we need to get $\vec{C}_\delta$ from $\mathtt{m}$ and to decompose $\mathtt{Q}$ into the product of $\mathtt{K}$ in the form of (6.4) and $\mathtt{R}$ such that $\mathtt{R}^\top\mathtt{R} = \mathtt{I}$ and

$$|KR| > 0 \Rightarrow |M| > 0$$

$|M|$:

$$> 0 \Rightarrow \left|\frac{\operatorname{sign}|M|}{\|m_3\|} M\right| = \frac{1\cdot|M|}{\|m_3\|} > 0$$

$$< 0 \Rightarrow \left|\frac{\operatorname{sign}|M|\,M}{\|m_3\|}\right| = \frac{\overbrace{-1^3}^{<0}\,\overbrace{|M|}^{<0}}{\|m_3\|} > 0$$

$$\boxed{\frac{\operatorname{sign}|M|}{\|m_3\|} M \text{ is decomposable into } K\,R}$$

Having k_{ij} computed, we recover $\mathtt{R}$ from $\mathtt{M}$ as

$$\boxed{\mathtt{R} = \mathtt{K}^{-1} \frac{\operatorname{sign} |\mathtt{M}|}{\|\mathtt{m}_3\|} \mathtt{M}} \tag{6.50}$$

$\ldots = \mathtt{KR}$

$\mathtt{K}^{-1} \ldots = \mathtt{R}$

Camera projection center can be computed in two ways. Either we get

$$\boxed{\vec{C}_\delta = -\mathtt{M}^{-1}\mathtt{m}} \tag{6.51}$$

For finite cameras, rank $\mathtt{M} = 3$

$\mathtt{P} = [\underbrace{\mathtt{KR}}_{\tilde{\mathtt{M}}} \mid \underbrace{-\mathtt{KR}\vec{C}_\delta}_{\tilde{\mathtt{m}}}]$, $\mathtt{m}$

or we obtain it by finding a basis $\mathtt{c}$ of the one-dimensional right null space of matrix $\mathtt{Q}$, i.e. solving

$$\boxed{\mathtt{Q}\,\mathtt{c} = 0} \qquad \mathtt{P}\begin{bmatrix} \vec{C}_\delta \\ 1 \end{bmatrix} = \vec{0} \tag{6.52}$$

and then computing

$$\begin{bmatrix} \vec{C}_\delta \\ 1 \end{bmatrix} = \frac{1}{\mathtt{c}_4}\mathtt{c} \tag{6.53}$$

$\mathtt{KR}\vec{C}_\delta - \mathtt{KR}\vec{C}_\delta = 0$

General (even for $\vec{C}$ at ∞) but then $\mathtt{C} \sim \mathtt{c} = \begin{bmatrix} \vdots \\ 0 \end{bmatrix}$

where $\mathtt{c}_4$ is the fourth coordinate of vector $\mathtt{c}$.

6.2 Camera calibration and angle between projection rays

All calibration information

We have introduced matrices $\mathtt{P}$, $\mathtt{R}$ and $\mathtt{K}$, and vector $\vec{C}_\delta$ which determine the projection from space to images. However, since $\mathtt{K}$ is introduced with $\mathtt{K}_{33} = 1$, the triplet $(\mathtt{K}, \mathtt{R}, \vec{C}_\delta)$ does not contain all information about the camera, which can be obtained by direct measurement of its physical components in a world coordinate system equipped with a known *world unit length* $\mathbf{1}_W$. The missing element is the scale of $\mathtt{P}$, which is equivalent to knowing the value of the focal length or the size of pixels, i.e. f, $\|\vec{b}_1\|$ or $\|\vec{b}_2\|$, in $\mathbf{1}_W$.

$\mathtt{K}, \mathtt{R}, \vec{C}_\delta$ +
- f ✓ $(\mathtt{K}, f; \mathtt{R}, \vec{C}_\delta)$
- $\|\vec{b}_1\|$ ✓ $(\mathtt{K}, \|\vec{b}_1\|; \mathtt{R}, \vec{C}_\delta)$
- $\|\vec{b}_2\|$ ✓ $(\underbrace{\mathtt{K}, \|\vec{b}_2\|}_{\text{internal (intrinsic)}}; \underbrace{\mathtt{R}, \vec{C}_\delta}_{\text{external}})$

$\mathtt{K}_{33}$

Knowing $\mathtt{K}$ and f allows to recover $\|\vec{b}_1\|$ from Equations 6.3 as $\|\vec{b}_1\| = f/k_{11}$. Knowing $\mathtt{K}$ and $\|\vec{b}_1\|$, on the other hand, gives $f = \|\vec{b}_1\|\, k_{11}$.

Therefore, full calibration of the camera is encoded in matrix $\mathtt{P}_\beta$, Equation 6.24, or, e.g., in one of the following tuples: $(\mathtt{K}_\beta, \mathtt{R}, \vec{C}_\delta)$, $(\mathtt{K}, \mathtt{R}, \vec{C}_\delta, f)$, $(\mathtt{K}, \mathtt{R}, \vec{C}_\delta, \|\vec{b}_1\|)$ or $(\mathtt{K}, \mathtt{R}, \vec{C}_\delta, \|\vec{b}_2\|)$.

We defined the camera calibration matrix $\mathtt{K}$ with $\mathtt{K}_{33} = 1$ because we often do not have access to the world unit when working with images without knowing anything about the camera which was used to make them. Moreover, a number of important tasks can be done without knowing the world unit.

§ 1 Angle between projection rays Consider two image points $\vec{u}_{1\alpha}$ and $\vec{u}_{2\alpha}$. The direction vectors of the rays are in β given by

$$\vec{x}_{1\beta} = \begin{bmatrix} \vec{u}_{1\alpha} \\ 1 \end{bmatrix}, \quad \vec{x}_{2\beta} = \begin{bmatrix} \vec{u}_{2\alpha} \\ 1 \end{bmatrix} \tag{6.54}$$

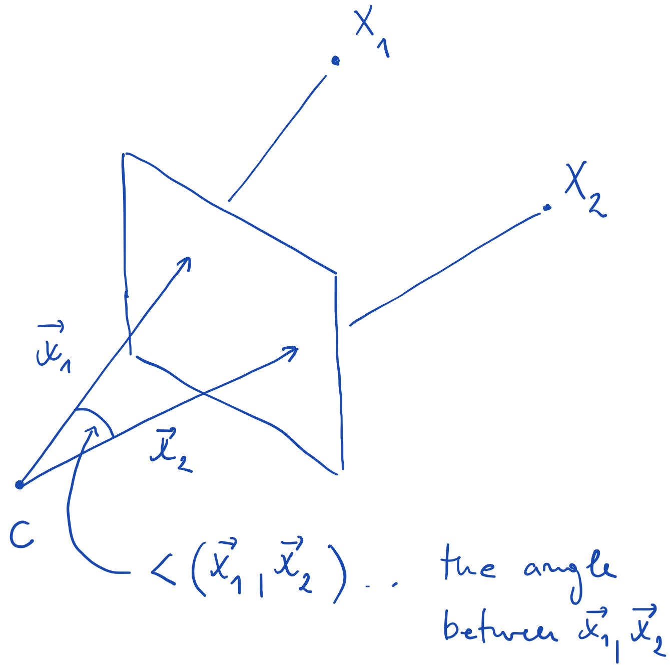
To obtain the angle between the direction vectors by evaluating the scalar product of the vectors, we need to pass to an orthogonal basis. The "closest" orthogonal basis is γ. Hence

$$\cos\angle(\vec{x}_1, \vec{x}_2) = \frac{\vec{x}_{1\gamma}^\top \vec{x}_{2\gamma}}{\|\vec{x}_{1\gamma}\|\|\vec{x}_{2\gamma}\|} = \frac{\vec{x}_{1\beta}^\top \mathtt{K}^{-\top}\mathtt{K}^{-1}\vec{x}_{2\beta}}{\|\mathtt{K}^{-1}\vec{x}_{1\beta}\|\|\mathtt{K}^{-1}\vec{x}_{2\beta}\|} \tag{6.55}$$

Notice that we could use the orthogonal basis γ to measure angles instead of, e.g., the closest orthonormal basis ϵ since the unknown scale factor f cancels in the following formula

$$\cos\angle(\vec{x}_1, \vec{x}_2) = \frac{\vec{x}_{1\epsilon}^\top \vec{x}_{2\epsilon}}{\|\vec{x}_{1\epsilon}\|\|\vec{x}_{2\epsilon}\|} = \frac{(f\,\vec{x}_{1\gamma}^\top)(f\,\vec{x}_{2\gamma})}{\|f\,\vec{x}_{1\epsilon}\|\|f\,\vec{x}_{2\gamma}\|} = \frac{\vec{x}_{1\gamma}^\top \vec{x}_{2\gamma}}{\|\vec{x}_{1\gamma}\|\|\vec{x}_{2\gamma}\|} \tag{6.56}$$

We conclude that we do not need to know f to measure angles between projection rays.



Quiz