Camera calibration and pose

Image projection matrix $\mathtt{P}_\beta = [\mathtt{A}\,|\,\vec{c}\,]$ transforms coordinates of points w.r.t. world coordinate system O, δ to camera coordinate system C, β. Two facts are included in such transformation: a camera pose (moving of the camera and its orientation in the space) and camera calibration (the internal geometrical features of the camera, focal length, pixels width etc.). It is possible to separate these two facts into two matrices: $\mathtt{R}$ (matrix of rotation) and $\mathtt{K}$ (camera calibration matrix) such a way, that $\mathtt{A} = \mathtt{K}\,\mathtt{R}$ (matrix product). The moving from origin O to origin C is done by fourth column of $\mathtt{P}_\beta$ which is equal to $-\mathtt{A}\,C_\delta$. The features of matrices $\mathtt{R}$ and $\mathtt{K}$ will be studied in this text.

1 Rotation and its matrix

Matrix $\mathtt{R} \in \mathbb{R}^{n\times n}$ is called *orthogonal*, if its columns are orthonormal vectors in $\mathbb{R}^n$. Note1: “orthogonality” and “orthonormality” of vectors in $\mathbb{R}^n$ are measured using standard dot product: $\vec{x}\cdot\vec{y} = \vec{x}^\top\vec{y}$. Note2: the terminology is somewhat confusing because *orthogonal* matrix has *orthonormal* columns, but this is traditional notation.

It follows from matrix multiplication that if $\mathtt{R}$ is orthogonal then $\mathtt{R}^\top\mathtt{R} = \mathtt{I}$, so $\mathtt{R}$ is regular and its inverse is $\mathtt{R}^\top$. Matrix calculus says that there is only one matrix inversion $\mathtt{A}^{-1}\mathtt{A} = \mathtt{I} = \mathtt{A}\mathtt{A}^{-1}$, so for the orthogonal matrix $\mathtt{R}$, it holds $\mathtt{R}\,\mathtt{R}^\top = \mathtt{I}$ too. This means that if $\mathtt{R}$ is orthogonal then it has orthonormal rows too. Laplace Theorem says that $1 = \det\mathtt{I} = \det(\mathtt{R}^\top\mathtt{R}) = (\det\mathtt{R})(\det\mathtt{R})$, so $\det\mathtt{R}$ must be ± 1.

Let $\delta = (\vec{d}_1, \ldots, \vec{d}_n)$ be an orthonormal basis in a vector space V with a dot product, $\dim V = n$. Let $\mathtt{R} \in \mathbb{R}^{n\times n}$ be an orthogonal matrix. Denote by $T_\mathtt{R}$ the transformation $V \to V$ with its matrix $\mathtt{R}$ w.r.t. the basis δ. Remember that

$$\big(T_\mathtt{R}(\vec{d}_1), \ldots, T_\mathtt{R}(\vec{d}_n)\big) = (\vec{d}_1, \ldots, \vec{d}_n)\,\mathtt{R}.$$

Then $\forall\,\vec{u}, \vec{v} \in V$: $\vec{u}\cdot\vec{v} = T_\mathtt{R}(\vec{u})\cdot T_\mathtt{R}(\vec{v}) = \vec{u}_\delta^\top\vec{v}_\delta$. This Theorem says that transformations derived from orthogonal matrix keeps dot product unchanged. It means that it keeps all features induced by dot product (sizes and angles) unchanged. Especially if $\vec{u}, \vec{v}$ are orthonormal then $T_\mathtt{R}(\vec{u}), T_\mathtt{R}(\vec{v})$ are orthonormal too. The proof of this Theorem is simple:

$$T_\mathtt{R}(\vec{u})\cdot T_\mathtt{R}(\vec{v}) = (\mathtt{R}\,\vec{u}_\delta)^\top(\mathtt{R}\,\vec{v}_\delta) = \vec{u}_\delta^\top\,\mathtt{R}^\top\mathtt{R}\,\vec{v}_\delta = \vec{u}_\delta^\top\,\mathtt{I}\,\vec{v}_\delta = \vec{u}_\delta^\top\,\vec{v}_\delta = \vec{u}\cdot\vec{v}.$$

How does the transformations $T_\mathtt{R}$ look like from a geometrical point of view? Imagine an orthonormal basis δ and its transformation $T_\mathtt{R}(\delta)$. The $T_\mathtt{R}(\delta)$ must be orthonormal too. This is possible only if $T_\mathtt{R}$ is a rotation or rotation plus reflection.

Because $\det\mathtt{R} = \pm 1$ and it measures the oriented volume of a n-cube given by the basis $T_\mathtt{R}(\delta)$, then we can see that $T_\mathtt{R}$ is rotation if $\det\mathtt{R} = 1$ and $T_\mathtt{R}$ is rotation plus reflection if $\det\mathtt{R} = -1$.

We can do a reverse walk: from a transformation $T : V \to V$ which is rotation to its matrix $\mathtt{R}$. We will show that such matrix must be orthogonal and $\det\mathtt{R} = 1$. Because T is rotation then it keeps sizes and angles, so it keeps dot product. This means that the following equation must be true:

$$T_\mathtt{R}(\vec{u})\cdot T_\mathtt{R}(\vec{v}) = (\mathtt{R}\vec{u}_\delta)^\top(\mathtt{R}\,\vec{v}_\delta) = \vec{u}_\delta^\top\,\mathtt{R}^\top\mathtt{R}\,\vec{v}_\delta = \vec{u}_\delta^\top\,\vec{v}_\delta = \vec{u}\cdot\vec{v}$$

for every $\vec{u}, \vec{v}$ in V. Only a matrix with property $\mathtt{R}^\top\mathtt{R} = \mathtt{I}$ can do this, so $\mathtt{R}$ must be orthogonal. Because rotation does not change a positivity of oriented volumes then $\det\mathtt{R} = 1$.

2 KR decomposition

The image projection matrix $\mathtt{P}_\beta = [\mathtt{A}\,|\,\vec{c}\,]$ is given ($\mathtt{A}$ is regular). We need to find the focal length $f \neq 0$ and the matrices $\mathtt{K}, \mathtt{R}$, where $\mathtt{K}$ is regular upper triangular with $\mathtt{K}_{3,3} = 1$ (camera calibration matrix) and $\mathtt{R}$ is orthogonal with $\det\mathtt{R} = 1$ (matrix of rotation) and $f\mathtt{A} = \mathtt{K}\,\mathtt{R}$. The matrices $\mathtt{A}, \mathtt{K}, \mathtt{R} \in \mathbb{R}^{3\times 3}$.

The decomposition of a given regular matrix $\mathtt{A}$ to the product of two matrices, the first is orthogonal, the second is upper triangular, is well known as QR decomposition (which is equal to Gram-Schmidt orthogonalization process from matrices point of view). But our problem needs to do the decomposition with reverse order of matrices: the first matrix is upper triangular and the second matrix is orthogonal. We will show two methods: a direct computation and a conversion of given problem to QR decomposition using transposed and permutation matrices.

1. Direct computation. The equation $f\mathtt{A} = \mathtt{KR}$ can be written in more detail as:

$$\begin{bmatrix} f\,\vec{a}_1^\top \\ f\,\vec{a}_2^\top \\ f\,\vec{a}_3^\top \end{bmatrix} = \begin{bmatrix} k & l & m \\ 0 & n & p \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \vec{r}_1^\top \\ \vec{r}_2^\top \\ \vec{r}_3^\top \end{bmatrix}$$

where $\vec{a}_i^\top$ are rows of the given matrix $\mathtt{A}$ and $\vec{r}_i^\top$ are rows of the (unknown) orthogonal matrix $\mathtt{R}$. From the last row of this equation we see that

$$f\,\vec{a}_3^\top = 0 + 0 + 1\,\vec{r}_3^\top$$

Because $\|\vec{r}_3\| = 1$ ($\mathtt{R}$ is orthogonal), we have the first result: $f = 1/\|\vec{a}_3\|$ and $\vec{r}_3 = f\vec{a}_3$.

Denote $\mathtt{B} = f\mathtt{A}$ (now, this is known matrix). We are finding matrices $\mathtt{K}, \mathtt{R}$ such that $\mathtt{B} = \mathtt{K}\,\mathtt{R}$. Moreover the last row of the matrix $\mathtt{B}$ is equal to the last row of the matrix $\mathtt{R}$ and the problem can be written as:

$$\begin{bmatrix} \vec{b}_1^\top \\ \vec{b}_2^\top \\ \vec{b}_3^\top \end{bmatrix} = \begin{bmatrix} k & l & m \\ 0 & n & p \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \vec{r}_1^\top \\ \vec{r}_2^\top \\ \vec{r}_3^\top \end{bmatrix} \tag{1}$$

where $\vec{b}_i^\top = f\vec{a}_i^\top$ are lines of the matrix $\mathtt{B}$. Now, we apply multiplication of the first and the second line by $\vec{r}_3$ (it is a known vector at this state of computation):

$$\begin{aligned} \vec{b}_1^\top \vec{r}_3 &= k\,\vec{r}_1^\top \vec{r}_3 + l\,\vec{r}_2^\top \vec{r}_3^\top + m\,\vec{r}_3^\top \vec{r}_3 = m \\ \vec{b}_2^\top \vec{r}_3 &= 0 + n\,\vec{r}_2^\top\,\vec{r}_3^\top + p\,\vec{r}_3^\top \vec{r}_3 = p \end{aligned}$$

The orthogonality of matrix $\mathtt{R}$ was used, i. e. $\vec{r}_i$ are orthonormal vectors. The coefficients m and p are computed from the equations above. Next step is Gauss-Jordan elimination applied on both sides of the equation (1): $-p$ times the last row is added to the second and $-m$ times the last row is added to the first row.

$$\begin{bmatrix} \vec{b}_1^\top - m\,\vec{b}_3^\top \\ \vec{b}_2^\top - p\,\vec{b}_3^\top \\ \vec{b}_3^\top \end{bmatrix} = \begin{bmatrix} k & l & 0 \\ 0 & n & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \vec{r}_1^\top \\ \vec{r}_2^\top \\ \vec{r}_3^\top \end{bmatrix} \tag{2}$$

The second row says that the known vector $\vec{b}_2^\top - p\,\vec{b}_3^\top$ is equal to $n\,\vec{r}_2^\top$. This gives the result $n = \|\vec{b}_2 - p\,\vec{b}_3\|$ (because $\|\vec{r}_2\| = 1$) and $\vec{r}_2 = (1/n)(\vec{b}_2 - p\,\vec{b}_3)$. Now, we know values of $f, p, m, n, \vec{r}_3, \vec{r}_2$ and all vectors at the left side of the equation (2). We will repeat the similar steps again. Multiply the first line of the equation (2) by previously calculated vector $\vec{r}_2$ with result: $(\vec{b}_1^\top - m\,\vec{b}_3^\top)\,\vec{r}_2 = k\,\vec{r}_1^\top \vec{r}_2 + l\,\vec{r}_2^\top \vec{r}_2 = l$. This equation calculates l value. Do the Gauss-Jordan elimination at both sides of the equation (2):

$$\begin{bmatrix} n\,(\vec{b}_1^\top - m\,\vec{b}_3^\top) - l\,(\vec{b}_2^\top - p\vec{b}_3^\top) \\ \vec{b}_2^\top - p\,\vec{b}_3^\top \\ \vec{b}_3^\top \end{bmatrix} = \begin{bmatrix} k & 0 & 0 \\ 0 & n & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \vec{r}_1^\top \\ \vec{r}_2^\top \\ \vec{r}_3^\top \end{bmatrix}$$

Denote the vector $\vec{u} = n\,(\vec{b}_1 - m\,\vec{b}_3) - l\,(\vec{b}_2 - p\vec{b}_3)$. This vector is known (and nonzero, because $\mathtt{B}$ is a regular matrix). The first row says $\vec{u}^\top = k\,\vec{r}_1^\top$ and we get the final result $k = \|\vec{u}\|$ and $\vec{r}_1 = (1/k)\vec{u}$. All elements of matrices $\mathtt{K}, \mathtt{R}$ are known, the decomposition is completed. Because n, k are positive, then $\det \mathtt{K} = kn > 0$. If $\det \mathtt{A} > 0$ then must be $\det \mathtt{R} = +1$.

But a little problem can arise in a special situation when $\det \mathtt{A} < 0$. Then $\det \mathtt{R} = -1$, but we need a matrix of rotation here, no reflection. We can modify orthogonal matrix $\mathtt{R}$ in order to it keeps its orthogonality: (i) first row can be multiplied by -1 or (ii) second row can be multiplied by -1 or

(iii) whole matrix $\mathtt{R}$ can be multiplied by -1. The new matrix $\mathtt{R}$ is matrix of rotation. In order to keep the equation $f\mathtt{A} = \mathtt{K}\,\mathtt{R}$ we have to do modifications which depends on the modifications of matrix $\mathtt{R}$ mentioned above: (i) multiply k by -1 or (ii) multiply n and l by -1 or (iii) multiply f by -1. Which alternative is used depends on our knowledge about the geometry of basis β.

Note that the projection matrix $\mathtt{P}_\beta = [\mathtt{A}\,|\,\vec{c}\,]$ has its fourth column $\vec{c} = -\mathtt{A}\,C_\delta$. If we need to calculate the coordinates of the projection center point C w.r.t. the world coordinate system (i. e. C_δ) and the matrix $\mathtt{P}_\beta$ is given then we need to calculate $C_\delta = -\mathtt{A}^{-1}\vec{c}$ or we need to compute the solution of the linear equation system with extended matrix $[\mathtt{A}\,|\,-\vec{c}\,]$.

2. KR decomposition using QR decomposition. Assume that we have a computer which is able to provide QR decomposition of given matrix $\mathtt{A}$, for example `[Q R] = qr(A)` in Matlab. Note that QR decomposition is quite common algorithm but RQ decomposition is rarely accessible. Important notice: we have a notation mishmash here: $\mathtt{R}$ is an orthogonal matrix in our notation but $\mathtt{R}$ is upper a diagonal matrix and $\mathtt{Q}$ is an orthogonal matrix in common notation used in QR algorithm. This is the reason why we introduce new name of the output of QR algorithm: `[Q U] = qr(A)`.

A regular matrix $\mathtt{A}$ is given and we need to find an orthogonal matrix $\mathtt{Q}$ and an upper triangular matrix $\mathtt{U}$ with the condition $\mathtt{A} = \mathtt{U}\,\mathtt{Q}$. We can use a (more common) QR algorithm which does $\mathtt{A} = \mathtt{Q}\,\mathtt{U}$ decomposition in a computer.

Let us consider a permutation matrix

$$\mathtt{P} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

The problem can be simply generalized for arbitrary dimension of matrices but matrices in $\mathbb{R}^{3\times 3}$ are sufficient for our purpose. Note that $\mathtt{PA}$ reverses the order of rows and $\mathtt{AP}$ reverses the order of columns of the matrix $\mathtt{A}$. Permutation matrices are orthogonal: $\mathtt{PP}^\top = \mathtt{P}^\top\mathtt{P} = \mathtt{I}$. Moreover $\mathtt{P} = \mathtt{P}^\top$, so $\mathtt{PP} = \mathtt{I}$. Reverse rows of the the given matrix $\mathtt{A}$ to get $\mathtt{A}_1 = \mathtt{PA}$. Then use QR decomposition on the transposition matrix $\mathtt{A}_1^\top = \mathtt{Q}_1\mathtt{U}_1$. We will show that the matrix $\mathtt{Q} = \mathtt{PQ}_1^\top$ is orthogonal and $\mathtt{U} = \mathtt{PU}_1^\top\mathtt{P}$ is upper triangular and these matrices hold the desired equation $\mathtt{A} = \mathtt{UQ}$. Why the matrix $\mathtt{PQ}_1^\top$ is orthogonal? Because

$$(\mathtt{PQ}_1^\top)(\mathtt{PQ}_1^\top)^\top = \mathtt{PQ}_1^\top\mathtt{Q}_1\mathtt{P}^\top = \mathtt{PIP}^\top = \mathtt{I}.$$

Why the matrix $\mathtt{PU}_1^\top\mathtt{P}$ is upper triangular? First, the transposition is applied, so $\mathtt{U}_1^\top$ is lower triangular. Then the rows are arranged in reverse order (we get the "triangularity" along collateral diagonal) and then the columns are reversed and we get upper triangular matrix again. Why $\mathtt{A} = \mathtt{UQ}$? Because

$$\mathtt{U}\,\mathtt{Q} = (\mathtt{PU}_1^\top\mathtt{P})\,(\mathtt{PQ}_1^\top) = \mathtt{PU}_1^\top\mathtt{Q}_1^\top = \mathtt{P}(\mathtt{Q}_1\mathtt{U}_1)^\top = \mathtt{P}\,\mathtt{A}_1 = \mathtt{P}(\mathtt{P}^{-1}\mathtt{A}) = \mathtt{A}.$$

The result of this process is an upper triangular matrix $\mathtt{U}$ with $\mathtt{U}_{3,3} = (1/f)$. We need to calculate $\mathtt{K} = f\,\mathtt{U}$ in order to get an upper triangular matrix $\mathtt{K}$ with the condition $\mathtt{K}_{3,3} = 1$. Moreover, $f = 1/\mathtt{U}_{3,3}$ is focal length, if $\mathtt{P}_\beta = [\mathtt{A}\,|\,\vec{c}\,]$ is an image projection matrix.

3 Elements of the calibration matrix K

If the geometry of basis $\beta = (\vec{b}_1, \vec{b}_2, \vec{b}_3)$ is known (i. e. the basis of the camera coordinate system), then the values of elements of upper triangular calibration matrix $\mathtt{K}$ can be expressed using sizes of vectors $\vec{b}_i$, angles between them and using focal length f. The corresponding equations will be derived in this section. The meanings of the elements of the matrix $\mathtt{K}$ will be clarified from these equations. The equations will also enable to recover geometry of β from a known matrix $\mathtt{K}$.

The equation $f\mathtt{A} = \mathtt{KR}$ can be expressed in another form

$$\mathtt{A} = \mathtt{K}\left(\frac{1}{f}\mathtt{R}\right) \tag{3}$$

The matrix $\mathtt{A}$ transfers coordinates w.r.t. the world basis δ to coordinates w.r.t. the camera basis β (i.e. $\mathtt{Ax}_\delta = \vec{x}_\beta$). The decomposition (3) does this transformation in two steps. Firstly, the coordinates w.r.t. δ are transformed to the coordinates to "in between" basis γ using the transformation matrix

$(1/f)\mathtt{R}$ and, secondly, these coordinates are transformed to the coordinates w.r.t β using the matrix $\mathtt{K}$. The first transformation does rotation and scaling by f. Assume that the world basis is orthonormal. The sizes of its vectors are equal to an unit by which we do measurement in our geometrical space (millimeters, inches, etc.). The basis δ is mapped by inverse rotation and scaling to basis $\gamma = (\vec{c}_1, \vec{c}_2, \vec{c}_3)$. Vectors $\vec{c}_i$ are orthogonal (like vectors from δ) but their sizes are equal to f. See Figure 6.2 at the page 36 of the main GVG text. The projection plane is perpendicular to $\vec{c}_3$ and its distance from camera center C is f.

Because $\mathtt{K}\vec{x}_\gamma = \vec{x}_\beta$ the matrix $\mathtt{K}$ includes in its columns the coordinates of vectors $\vec{c}_i$ w.r.t the basis β. Look at the Figure 6.2 in the main GVG text again and focus to the geometrical interpretation of these coordinates. The matrix $\mathtt{K}$ realizes an orthogonalization process from non-orthogonal basis β to orthogonal basis γ. We will next show it in detail.

Vectors $\vec{b}_1$, $\vec{c}_1$ are parallel and consistently oriented, so $\vec{c}_1 = k\,\vec{b}_1$, where $k = \|\vec{c}_1\|/\|\vec{b}_1\| = f/\|\vec{b}_1\|$. (Note that the sizes of vectors are measured in the world units.) The first column of matrix $\mathtt{K}$ includes coordinates $\vec{c}_1$ w.r.t β, so it must be $[k, 0, 0]^\top$.

It is possible to draw all representantives of free vectors $\vec{b}_1, \vec{b}_2, \vec{c}_1, \vec{c}_2$ into projection plane (do it!) because $\operatorname{span}(\vec{b}_1, \vec{b}_2) = \operatorname{span}(\vec{c}_1, \vec{c}_2)$. It means that there exist scalars l, n such that $\vec{c}_2 = l\,\vec{b}_1 + n\,\vec{b}_2$. Denote l_0 the orthogonal projection of $\vec{b}_2$ to $\vec{c}_1$ and n_0 the orthogonal projection of $\vec{b}_2$ to $\vec{c}_2$. It is clear (from the figure) that

$$\vec{c}_2 = \frac{f}{n_0}\left(\vec{b}_2 - l_0\frac{\vec{b}_1}{\|\vec{b}_1\|}\right). \tag{4}$$

Let φ is the angle between $\vec{b}_1$ and $\vec{b}_2$. Then $l_0 = \|\vec{b}_2\|\,\cos\varphi$ and $n_0 = \|\vec{b}_2\|\,\sin\varphi$. It follows from equation (4) that

$$\vec{c}_2 = \frac{f}{\|\vec{b}_2\|\,\sin\varphi}\left(\vec{b}_2 - \|\vec{b}_2\|\,(\cos\varphi)\frac{\vec{b}_1}{\|\vec{b}_1\|}\right) = \frac{f}{\|\vec{b}_2\|\sin\varphi}\,\vec{b}_2 - \frac{f\,\cos\varphi}{\|\vec{b}_1\|\sin\varphi}\,\vec{b}_1 = l\,\vec{b}_1 + n\,\vec{b}_2,$$

$$\text{it means that}\quad l = -\frac{f\,\cos\varphi}{\|\vec{b}_1\|\sin\varphi}, \qquad n = \frac{f}{\|\vec{b}_2\|\sin\varphi}.$$

Because the second column of the matrix $\mathtt{K}$ includes coordinates of $\vec{c}_2$ w.r.t the basis β, this column must be $[l, n, 0]^\top$.

There exists one *principal point* P on the projection plane which is the perpendicular projection of C onto the projection plane. Let the coordinates of this point w.r.t the coordinate system $(o, \vec{b}_1, \vec{b}_2)$ be m, p. It means that $P = o + m\,\vec{b}_1 + p\,\vec{b}_2$ Note that o is the origin with the condition $o = C + \vec{b}_3$. Thus $P = C + \vec{b}_3 + m\,\vec{b}_1 + p\,\vec{b}_2$ which means that $\vec{c}_3 = P - C = \vec{b}_3 + m\,\vec{b}_1 + p\,\vec{b}_2$. Because the third column of the matrix $\mathtt{K}$ includes the coordinates of vector $\vec{c}_3$ w.r.t the basis β, it must be $[m, p, 1]^\top$.

Let we summarize the results. The matrix $\mathtt{K}$ is in the form

$$\mathtt{K} = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & k_{22} & k_{23} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} k & l & m \\ 0 & n & p \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{\|\vec{b}_1\|} & -\frac{f\cos\varphi}{\|\vec{b}_1\|\sin\varphi} & m \\ 0 & \frac{f}{\|\vec{b}_2\|\sin\varphi} & p \\ 0 & 0 & 1 \end{bmatrix}$$

where

f	... is the focal length
$\|\vec{b}_1\|, \|\vec{b}_2\|$	... are sizes of basic vectors, first one is the pixel width
φ	... is the angle between $\vec{b}_1$, $\vec{b}_2$
m, p	... are coordinates of the principal point w.r.t the coordinate system $(o, \vec{b}_1, \vec{b}_2)$

It is very common that pixels are rectangles: $\varphi = 90°$, $l = 0$ and $n = f\,/\,\|\vec{b}_2\|$.

Note that the elements of matrix $\mathtt{K}$ are dimensionless, because they include ratios of lengths. The element $k_{11} = k$ denotes the focal length expressed in the unit of the pixel width. The element $k_{22} = n$ denotes the focal length in the unit of pixel height (when pixel is a rectangle). The element m and p determine the principal point position in units of pixel width and pixel height.

If the elements of $\mathtt{K}$ are known then the angle φ between $\vec{b}_1$ and $\vec{b}_2$ can be calculated from the equation $l/k = -\text{cotan}\,\varphi$. And the ratio of basis vector lengths is $\|\vec{b}_2\|/\|\vec{b}_1\| = \sqrt{k^2+l^2}\,/n$.

Sometimes we can work with a little modification of equation (3) in the form:

$$\mathtt{A} = \left(\frac{1}{f}\mathtt{K}\right)\mathtt{R} = \mathtt{K}_\beta\,\mathtt{R}$$

The interpretation of such decomposition is: first only rotation without scaling is done and then $\mathtt{K}_\beta$ does orthogonalization together with scaling. The matrix $\mathtt{K}_\beta$ is named *image calibration matrix* and it has the form

$$\mathtt{K}_\beta = \begin{bmatrix} \frac{1}{\|\vec{b}_1\|} & -\frac{\cos\varphi}{\|\vec{b}_1\|\sin\varphi} & \frac{m}{f} \\ 0 & \frac{1}{\|\vec{b}_2\|\sin\varphi} & \frac{p}{f} \\ 0 & 0 & \frac{1}{f} \end{bmatrix}$$

Note that $\mathtt{K}_\beta$ is equal to the matrix $\mathtt{U}$ which is direct output from the RQ decomposition process mentioned at the end of the previous subsection.

4 Camera projection matrix vs. Image projection matrix

Let $\mathtt{P}_\beta$ and $t\,\mathtt{P}_\beta$ (where $t > 0$) be two image projection matrices. Then the projection result using both matrices is the same. This fact should be reformulated more lapidary: the projection result is unchanged when f and pixel sizes are scaled by the same scalar. We will show this fact more exactly.

Let X be a point with coordinates $\vec{x}_\delta$ w.r.t the world coordinate system δ. Let $\mathtt{P}_\beta = [\mathtt{A}\,|\,\vec{c}\,]$ be an image projection matrix. Then,

$$\mathtt{P}_\beta \begin{bmatrix} \vec{X}_\delta \\ 1 \end{bmatrix} = \mathtt{A}\vec{x}_\delta + \vec{c} = \vec{X}_\beta = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

are the coordinates of the point X w.r.t camera coordinate system β. Finally we do the projection:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \sim \begin{bmatrix} x/z \\ y/z \\ 1 \end{bmatrix}, \qquad \text{i.e.} \qquad \frac{1}{z}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x/z \\ y/z \\ 1 \end{bmatrix},$$

so $[x/z, y/z]^\top$ are coordinates of the projection of point X. Now, do the same calculation with the matrix $t\mathtt{P}_\beta = [t\mathtt{A}\,|\,t\vec{c}\,]$:

$$t\mathtt{P}_\beta \begin{bmatrix} \vec{x}_\delta \\ 1 \end{bmatrix} = t\mathtt{A}\vec{x}_\delta + t\vec{c} = \begin{bmatrix} tx \\ ty \\ tz \end{bmatrix} \quad \text{and the projection is} \quad \begin{bmatrix} tx \\ ty \\ tz \end{bmatrix} \sim \begin{bmatrix} x/z \\ y/z \\ 1 \end{bmatrix}$$

The coordinates of the projection point are the same. Note that $[tx, ty, tz]^\top$ are coordinates of the point X w.r.t a scaled camera coordinate system with new focal length $f' = f/t$ and with new pixel sizes $\|\vec{b}'_i\| = \|\vec{b}_i\|/t$. The depth of the point X is z or tz, respectively, and it is measured in multiples of f or in multiples of f', respectively.

Remember that the focal length f can be calculated from image projection matrix $\mathtt{P}_\beta = [\mathtt{A}\,|\,\vec{c}\,]$ by $(1/f) = \|\vec{a}_3\|$ where $\vec{a}_3$ is the third row of matrix $\mathtt{A}$. But there exist situations when we have no access to physical dimensions of the camera and no knowledge about f. In such case the *camera projection matrix* $\mathtt{P} = [\mathtt{A}'\,|\,\vec{c}\,']$ with the property $\|\vec{a}'_3\| = 1$ is introduced. The decomposition $\mathtt{A}' = \mathtt{K}\,\mathtt{R}$ should be done directly without knowledge of f. The distance from camera center C to the projection plane is 1 unit (w.r.t world coordinate system) and the element k_{11} of the matrix $\mathtt{K}$ (for example) says how many pixel widths are in one unit. The camera projection matrix is an image projection matrix with $f = 1$.

The results of a projection using camera projection matrix $\mathtt{P}$ is the same as when a real image projection matrix $\mathtt{P}_\beta$ is used iff $\mathtt{P} = f\,\mathtt{P}_\beta$.