Camera Projection Model

Tomas Pajdla Center of Machine Perception
Department of Cybernetices
Czech Technical University in Prague Karlovo nam. 13, CZ-12135, Prague
pajdla@cmp.felk.cvut.cz

Figure 1: ...

Perspective projection is the most common image formation model [**?**, **?**, **?**].

1. Notation

Affine space [**?**, page 69] consists of a set of points and a mapping that assignes vectors $\vec{a}$ from the underlying linear space to ordered pairs $\overline{AB}$ of points. $\overline{AB}$ is called *affine vector*. $\mathcal{S} = (O; \Lambda)$ denotes the *coordinate system* of an affine space with origin O and basis Λ. $\vec{a} \approx \mathbf{a}_\Omega$ says that $\vec{a}$ has coordinates $\mathbf{a}_\Omega$ with respect to basis Ω. Assuming $\mathcal{S} = (O; \Lambda)$, $\mathbf{a}_\mathcal{S}$ means $\mathbf{a}_\Lambda$. We write $\overline{AB} \approx \vec{a}_\mathcal{S}$ and $\overline{AB} \approx \mathbf{a}_\mathcal{S}$ to say that vector $\vec{a}_\mathcal{S}$ represents affine vector $\overline{AB}$ in the coordinate system $\mathcal{S}$ and that vector with coordinates $\mathbf{a}_\mathcal{S}$ w.r.t. the basis of $\mathcal{S}$ represents $\overline{AB}$. **B** means linear basis, **OGB** means right-handed orthogonal linear basis, **ONB** means right-handed orthonormal linear basis.

2. Perspective projection model

Let us first review how are points in perspective images related to their projection half-rays and to their preimage points in space. We assume central projection with a finite projection center C for simplicity. Figure **??** shows a diagram of perspective image formation.

There are six important coordinate systems: (1) the scene Cartesian coordinte system $\mathcal{S} = (O; \vec{a}_1, \vec{a}_2, \vec{a}_3)$ with **ONB** $(\vec{a}_1, \vec{a}_2, \vec{a}_3)$, (2) the camera Cartesian coordinate system $\mathcal{W} = (C; \vec{i}, \vec{j}, \vec{k})$ with **ONB** $(\vec{i}, \vec{j}, \vec{k})$, (3) the camera orthogonal coordinate system $\mathcal{O} = (C; \vec{i}, \vec{j}, \vec{e})$ with **OGB** $(\vec{i}, \vec{j}, \vec{e})$, (4) the camera affine coordinate system $\mathcal{C} = (C; \vec{b}_1, \vec{b}_2, \vec{b}_3)$ with **B** $(\vec{b}_1, \vec{b}_2, \vec{b}_3)$, (5) the image orthogonal coordinate system $\mathcal{P} = (p; \vec{i}, \vec{j}, \vec{e})$ with **ONB** $(\vec{i}, \vec{j})$, and (6) the image affine coordinate system $\mathcal{I} = (o; \vec{b}_1, \vec{b}_2, \vec{b}_3)$ with **B** $(\vec{b}_1, \vec{b}_2, \vec{b}_3)$.

The coordinates system $\mathcal{I}$ is defined partially by the given coordinate system $(o; \vec{b}_1, \vec{b}_2)$ of the digital image, in which we measure pixel coordinates, and partially by vector $\vec{b}_3 \approx \overline{oC}$, which forms altogether with $\vec{b}_1, \vec{b}_2$ a linear basis. Such a coordinate system exists since we assumed to measure linear pixel coordinates. Vectors $\vec{b}_1, \vec{b}_2$ form a linear basis, i.e. we allow for a 'general non-square and non-orhogonal pixel'.

The coordinate system $\mathcal{P}$ is derived from $\mathcal{I}$. The origin p is taken so that the line cP is perpendicular to the image plane π. Vector $\vec{i}$ is a unit vector in the direction of vector $\vec{b}_1$. Vector $\vec{j}$ is a linear combination of $\vec{b}_1$, $\vec{b}_2$ and has a unit length. Orientation of $\vec{j}$ is chosen such that $\vec{i}, \vec{j}, \vec{b}_3$ form a right-handed basis. Vector $\vec{k}$ forms altogether with $\vec{i}$, $\vec{j}$ **ONB**. Vector $\vec{e}$ represents point C in the coordinate system $(p; \vec{i}, \vec{j}, \vec{k})$. We may write

$$\begin{aligned} \vec{i} &= a_{11}\vec{b}_1 \\ \vec{j} &= a_{12}\vec{b}_1 + a_{22}\vec{b}_2 \\ \vec{e} &= a_{13}\vec{b}_1 + a_{23}\vec{b}_2 + \vec{b}_3 \end{aligned} \tag{1}$$

with a_{11}, a_{12}, and a_{22} such that $\vec{i}, \vec{j}$ are lineraly independent.

There are seven important affine vectors that are represented by vectors w.r.t. different coordinate systems: (1) $\overline{OC} \approx \mathbf{C}_\mathcal{S}$, (2) $\overline{OX} \approx \mathbf{X}_\mathcal{S}$, (3) $\overline{CX} \approx \mathbf{X}_\mathcal{W}, \mathbf{X}_\mathcal{O}, \mathbf{X}_\mathcal{C}$, (4) $\overline{Cx} \approx \mathbf{x}_\mathcal{O}$, (5) $\overline{uC} \approx \vec{x} \approx \mathbf{x}_\mathcal{W}, \mathbf{x}_\mathcal{C}, \mathbf{x}_\mathcal{P}, \mathbf{x}_\mathcal{I}$, (6) $\overline{ou} \approx \mathbf{u}_\mathcal{I}$, and (7) $\overline{pu} \approx \vec{u}_\mathcal{P} \approx \mathbf{u}_\mathcal{P}$.

Let us see how a scene point X, expressed as $\mathbf{X}_\mathcal{S}$, is projected into image point with pixel coordinates $(u, v)^\top$. Point $\mathbf{X}_\mathcal{S}$ is transformed to the camera Cartesian coordinate system $\mathcal{W}$ as

$$X_\mathcal{W} = \mathtt{R}(\mathbf{X}_\mathcal{S} - \mathbf{C}_\mathcal{S}),$$

and further to the camera orthogonal coordinates system $\mathcal{O}$ as

$$X_\mathcal{O} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & 1 & \\ & & a_{33} \end{pmatrix} X_\mathcal{W}$$

for some $a_{33} = \|\vec{e}\| \neq 0$. We may write

$$X_{\mathcal{O}} = \left\langle \begin{array}{ll} y_3 \neq 0: & y_3 \begin{pmatrix} y_1/y_3 \\ y_2/y_3 \\ 1 \end{pmatrix} \\ \\ y_3 = 0: & 1 \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix} \end{array} \right. . \tag{2}$$

Perspective projection of a point X is realized by passing from point $X_{\mathcal{O}}$ to the the following special direction vector $\mathbf{x}_{\mathcal{O}}$ of the projection ray CX

$$X_{\mathcal{O}} \rightarrow \mathbf{x}_{\mathcal{O}} = \left\langle \begin{array}{ll} y_3 \neq 0: & \begin{pmatrix} y_1/y_3 \\ y_2/y_3 \\ 1 \end{pmatrix} \\ \\ y_3 = 0: & \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix} \end{array} \right. .$$

Let us introduce a new notation for the coordinates of

$$\mathbf{x}_{\mathcal{O}} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \text{ with } \quad x_3 = \left\langle \begin{array}{l} 1 \\ \\ 0 \end{array} \right. .$$

Vector $\mathbf{x}_{\mathcal{O}}$ can be represented in coordinate sysytem $\mathcal{O}$ by affine vector $\overline{Cx}$, i.e. $\mathbf{x}_{\mathcal{O}} \approx \overline{Cx}$. It can be, however, also represented by the affine vector $\overline{uC}$ in coordinates system $\mathcal{P}$, i.e. $\overline{uC} \approx \vec{x} \approx \mathbf{x}_{\mathcal{O}} = \mathbf{x}_{\mathcal{P}}$. Vector $\vec{x}$ can, using (**??**), be writen as

$$\begin{aligned} \vec{x} &= x_1\vec{i} + x_2\vec{j} + x_3\vec{e} = \\ &= (a_{11}x_1 + a_{12}x_2 + a_{13}x_3)\,\vec{b}_1 + \\ &+ \ (a_{22}x_2 + a_{23}x_3)\,\vec{b}_2 + \\ &+ \ x_3\,\vec{b}_3 \end{aligned}$$

but also as

$$\vec{x} = -(u_1\vec{b}_1 + u_2\vec{b}_2) + u_3\vec{b}_3.$$

when introducing an explicit notation for the coordinates of $\mathbf{x}_{\mathcal{I}}$ as

$$\mathbf{x}_{\mathcal{I}} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}.$$

We see that

$$\mathbf{x}_{\mathcal{I}} = \begin{pmatrix} u_1 \\ u_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -a_{11} & -a_{12} & -a_{13} \\ & -a_{22} & -a_{23} \\ & & 1 \end{pmatrix} \mathbf{x}_{\mathcal{O}}$$

We have arrived at a relationship between coordinates $\mathbf{X}_{\mathcal{S}}$ of the scene point X and coordinates $\mathbf{u}_{\mathcal{I}} = (u_1, u_2)^\top$ of its projection u.

Let us rewrite the relationship in a more standard and concise way. Let us introduce matrices $\mathtt{A}, \mathtt{B}$ and vector $\mathbf{t}$ by the following equation

$$\mathtt{B} = \begin{pmatrix} -a_{11} & -a_{12} & -a_{13} \\ & -a_{22} & -a_{23} \\ & & 1 \end{pmatrix} = \begin{pmatrix} -\mathtt{A}^{-1} & -\mathtt{A}^{-1}\mathbf{t} \\ & 1 \end{pmatrix} \tag{3}$$

and, using (**??**), write

$$\mathbf{x}_{\mathcal{I}} \quad = \quad \left\langle \begin{array}{ll} y_3 \neq 0: & \mathtt{B}\dfrac{\mathbf{X}_{\mathcal{O}}}{y_3} \\ \\ y_3 = 0: & \mathtt{B}\dfrac{\mathbf{X}_{\mathcal{O}}}{1} \end{array} \right. .$$

Therefore, we get

$$\alpha \begin{pmatrix} u_1 \\ u_2 \\ x_3 \end{pmatrix} = \mathtt{B}\,\mathbf{X}_{\mathcal{O}} \quad \text{with} \quad \alpha = \left\langle \begin{array}{ll} y_3 \neq 0: & y_3 \\ y_3 = 0: & 1 \end{array} \right. .$$

Thus, we finally arrive to the celebrated pojection equation

$$\alpha \begin{pmatrix} \mathbf{u}_{\mathcal{I}} \\ x_3 \end{pmatrix} = \alpha \mathbf{x}_{\mathcal{I}} = \mathtt{P} \begin{pmatrix} \mathbf{X}_{\mathcal{S}} \\ 1 \end{pmatrix} \tag{4}$$

with

$$\mathtt{P} = \underbrace{\mathtt{B} \begin{pmatrix} 1 & & \\ & 1 & \\ & & a_{33} \end{pmatrix}}_{\mathtt{K}} \mathtt{R}(\mathtt{I} \ -\mathbf{C}_{\mathcal{S}}),$$

which relates a scene point $\mathbf{X}_{\mathcal{S}}$ to image point u with pixel coordinates $\mathbf{u}_{\mathcal{I}} = (u_1, u_2)^\top$.

Assume that we measured a point u in an uncalibrated perspective image with image pixel coordinates $\mathbf{u}_{\mathcal{I}} = (u_1, u_2)^\top$ w.r.t. affine coordinate system. If the point u is finite, then $\mathbf{u}_{\mathcal{I}}$ corresponds to $\overline{ou}$. If the point u is at infinity, then $\mathbf{u}_{\mathcal{I}}$ is a direction vector pointing towards that point at infinity. Points at infinity can't be measured directly but they can be constructed, e.g., as intersections of paralel lines. Let us construct coordinates $\mathbf{x}$ of the direction vector $\vec{x}$ of the projection half-ray associated with u in an affine camera coordinate system.

There is a matrix $\mathtt{A}$ and a vector $\mathbf{t}$ as in (**??**) that take $\mathbf{u}_{\mathcal{I}}$ to $\mathbf{u}_{\mathcal{P}}$ as $\mathbf{u}_{\mathcal{P}} = \mathtt{A}\mathbf{u}_{\mathcal{I}} + \mathbf{t}$ if $\mathbf{u}_{\mathcal{I}}$ represents a finite point, or as $\mathbf{u}_{\mathcal{P}} = \mathtt{A}\mathbf{u}_{\mathcal{I}}$ if $\mathbf{u}_{\mathcal{I}}$ represents a point at infinity. If u is finite, then, Figure **??**, $\vec{x} = \vec{e} - \vec{u}$, or $\vec{x} = -\vec{u}$ if u is at infinity. Thus

$$\mathbf{x} = \mathbf{x}_{\mathcal{O}} = \mathbf{x}_{\mathcal{P}} = \left\langle \begin{array}{ll} \mathbf{u}_{\mathcal{I}} \text{ finite}: & \begin{pmatrix} -\mathbf{u}_{\mathcal{P}} \\ 1 \end{pmatrix} \\ \\ \mathbf{u}_{\mathcal{I}} \text{ infinite}: & \begin{pmatrix} -\mathbf{u}_{\mathcal{P}} \\ 0 \end{pmatrix} \end{array} \right. . \tag{5}$$

We are done here. Nevertheless, let us make one more remark. We have chosen coordinates $\mathbf{x}_{\mathcal{O}}$ of $\vec{x}$ in the coordinate system $\mathcal{O}$. However, if we choose the camera affine coordinate system $\mathcal{C} = (C; \vec{b}_1, \vec{b}_2, \vec{b}_3)$ instead, then finite u has coordinates

$$\mathbf{x}_{\mathcal{C}} = \mathtt{B}\,\mathbf{x}_{\mathcal{O}} = \begin{pmatrix} -\mathtt{A}^{-1} & -\mathtt{A}^{-1}\mathbf{t} \\ & 1 \end{pmatrix} \begin{pmatrix} -\mathtt{A}\mathbf{u}_{\mathcal{I}} - \mathbf{t} \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{u}_{\mathcal{I}} \\ 1 \end{pmatrix}.$$

There exists a camera affine coordinate system in which $\mathbf{u}_{\mathcal{P}}$ augmented by 1 has the meaning of a projection ray direction vector.

References

[1] M. Kemp. *The Science of Art: Optical themes in western art from Brunelleschi to Seurat*. Yale University Press, New Haven, USA, 1990.

[2] M. M. Thompson, editor. *Manual of Photogrammetry*. Nedra Moskva, 1970. In Russian.

[3] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[4] O. Faugeras and Q. T. Luong and T. Papadopoulo. The Geometry of Multiple Images. MIT Press, 2001.

[5] S. K. Nayar. Omnidirectional Vision. *Eight International Symposium on Robotics Research*, Shonan, Japan, October 1997.

[6] S. Baker and S. K. Nayar. A theory of single-viewpoint catadioptric image formation. *IJCV*, 35(2): 175–196, Kluwer 1999.

[7] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems. *ECCV 2000*, pp. 445–461, LNCS 1843, Springer 2000.

[8] S. F. Ray. *Applied photographic optics: Lenses and optical systems for photography, film, video, electronic and digital imaging*. Focal Press, Oxford, 2002.

[9] M. M. Fleck. Perspective projection: the wrong imaging model. Technical Report TR 95-01, Comp. Sci., Univerity of Iowa, 1995. http://www.cs.hmc.edu/ fleck/stereo-tr2.ps.Z

[10] G. P. Stein. Lens distortion calibrating using point correspondences. *CVPR 1997*, pp. 602–609, IEEE Computer Society 1997.

[11] S. B. Kang. Catadioptric self-calibration. *CVPR 2000*, vol.$\tilde{1}$, pp. 201–207, IEEE Computer Society 2000.

[12] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. *CVPR 2001*, vol. 1, pp. 125–132, IEEE Computer Society 2001.

[13] B. Micusik and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. *CVPR 2003*, vol. 1, pp. 485–490, IEEE Computer Society 2003.