Survey on Camera calibration

Yuji Oyamada

[1]HVRL, Keio University

[2]Chair for Computer Aided Medical Procedure (CAMP), Technische Universität München

April 3, 2012

Camera calibration

- Necessary to recover 3D metric from image(s).
  - 3D reconstruction,
  - Object/camera localization, and
  - *etc.*
- Computes 3D (real world)–2D (camera image) relationship.

For quick literature review

Several review/survey papers and book chapters

- [Salvi et al., 2002]
- [Zhang, 2005]
- [Remondino and Fraser, 2006]

For camera calibration

REFORMAT: Important issues for calibration should be synchronized with the following text.

- Camera modeling:
- Parameters estimation:

A point in camera geometry

A point is expressed with several coordinate system.

3D points in world coordinate

A point $\mathbf{X}_w = (X_w, Y_w, Z_w)^\mathsf{T}$ in a world coordinate.

3D points in camera coordinate

A point $\mathbf{X}_c = (X_c, Y_c, Z_c)^\mathsf{T}$ in a camera coordinate.

2D points in image coordinate

A point $\mathbf{x} = (x, y)^\mathsf{T}$ in an image plane.

Projection matrix

A 3×4 projection matrix $\mathbf{P}$ denotes relationship between $\mathbf{X}_w$ and $\mathbf{x}$ as

$$\mathbf{x} = \mathbf{P}\mathbf{X}_w, \tag{1}$$

$$\rightarrow s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \tag{2}$$

Intrinsic and extrinsic parameters

A projection matrix can be decomposed into two components, intrinsic and extrinsic parameters, as

$$\mathbf{x} = \mathbf{P}\mathbf{X}_w = \mathbf{A}\left[\mathbf{R}|\mathbf{t}\right]\mathbf{X}_w, \tag{3}$$

$$\rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}, \tag{4}$$

where

- Intrinsic: 3×3 calibration matrix $\mathbf{A}$.
- Extrinsic: 3×3 Rotation matrix $\mathbf{R}$ and 3×1 translation vector $\mathbf{t}$.

Extrinsic parameters

Denotes transformation between $\mathbf{X}_w$ and $\mathbf{X}_c$ as

$$\mathbf{X}_c = [\mathbf{R}|\mathbf{t}]\,\mathbf{X}_w, \tag{5}$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \tag{6}$$

Intrinsic parameters

Project a 3D point $\mathbf{X}_c$ to image plane as

$$\mathbf{x} = \mathbf{A}\left[\mathbf{R}|\mathbf{t}\right]\mathbf{X}_w = \mathbf{A}\mathbf{X}_c, \tag{7}$$

$$\rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}, \tag{8}$$

where

- α_x and α_y are focal lengths in pixel unit.
- x_0 and y_0 are image center in pixel unit.
- s is skew parameter.

4 steps projecting a 3D world point to a 2D image point

A 3×4 projection matrix $\mathbf{P}$ denotes relationship between ${}^{W}\mathbf{X}_w$ and ${}^{I}\mathbf{x}$ as

$$
{}^{I}\mathbf{x} = \mathbf{P}\,{}^{W}\mathbf{X}_w, \tag{9}
$$

$$
\rightarrow s \begin{bmatrix} {}^{I}x_d \\ {}^{I}y_d \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} {}^{W}X_w \\ {}^{W}Y_w \\ {}^{W}Z_w \\ 1 \end{bmatrix}. \tag{10}
$$

1/4: A 3D world point to a 3D camera point

Change the world coordinate system to the camera one.

- From a 3D point ${}^{W}\mathbf{X}_w$ in metric system w.r.t. the world coordinate
- To a 3D point ${}^{C}\mathbf{X}_w$ in metric system w.r.t. the camera coordinate

$$
{}^{C}\mathbf{X}_w = [{}^{C}\mathbf{R}_w | {}^{C}\mathbf{T}_w] {}^{W}\mathbf{X}_w, \tag{11}
$$

$$
\rightarrow s \begin{bmatrix} {}^{C}X_w \\ {}^{C}Y_w \\ {}^{C}Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & t_{14} \\ R_{21} & R_{22} & R_{23} & t_{24} \\ R_{31} & R_{32} & R_{33} & t_{34} \end{bmatrix} \begin{bmatrix} {}^{W}X_w \\ {}^{W}Y_w \\ {}^{W}Z_w \\ 1 \end{bmatrix}. \tag{12}
$$

2/4: A 3D camera point to a 2D camera point

Change the 3D camera coordinate system to the 2D camera one.

- From a 3D point ${}^{C}\mathbf{X}_w$ in metric system w.r.t. the camera coordinate
- To a 2D point ${}^{C}\mathbf{X}_u$ in metric system w.r.t. the camera coordinate

$$
{}^{C}\mathbf{X}_u = [{}^{C}\mathbf{R}_w | {}^{C}\mathbf{T}_w] {}^{W}\mathbf{X}_w, \tag{13}
$$

$$
\rightarrow s \begin{bmatrix} {}^{C}X_u \\ {}^{C}Y_u \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} {}^{W}X_w \\ {}^{W}Y_w \\ {}^{W}Z_w \\ 1 \end{bmatrix}, \tag{14}
$$

$$
{}^{C}X_u = \frac{f}{{}^{W}Z_w} {}^{W}X_w \qquad {}^{C}Y_u = \frac{f}{{}^{W}Z_w} {}^{W}Y_w,
$$

where f denotes focal length in metric system.

3/4: Lens distortion

Practical lens distort the previous 3D→2D projection.

$$ {}^{C}X_u = {}^{C}X_d + \delta_x \qquad {}^{C}Y_u = {}^{C}Y_d + \delta_y, \tag{15} $$

where δ_x and δ_y denote distortion parameter along with each axis. In the case of no lens distortion,

$$ \delta_x = 0 \qquad \delta_y = 0 \tag{16} $$

- Radial distortion δ_{xr} and δ_{yr},
- Decentering distortoin δ_{xd} and δ_{yd},
- Thin prism distortion δ_{xp} and δ_{yp}.

3/4: Lens distortion: Radial distortion

- Caused by flawed radial curvature of lens.
- Modeled by [Tsai, 1987] [1].

$$\delta_{xr} = k_1 {}^C X_d ({}^C X_d^2 + {}^C Y_d^2) \qquad \delta_{yr} = k_1 {}^C Y_d ({}^C X_d^2 + {}^C Y_d^2) \tag{17}$$

[1] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Transactions on Robotics and Automation*, 3(4):323–344, 1987

3/4: Lens distortion: Weng's model

Model the lens distortion from the undistorted image point ${}^{C}\mathbf{X}_u$ instead of the distorted one ${}^{C}\mathbf{X}_d$.

$$ {}^{C}X_d = {}^{C}X_u + \delta_x \qquad {}^{C}Y_d = {}^{C}Y_u + \delta_y \tag{18} $$

3/4: Lens distortion: Decentering distortion

- Since the optical center of the lens is not correctly aligned with the center of the camera.
- Modeled by [Weng et al., 1992] [2],

$$\delta_{xd} = p_1(3\,{}^{C}X_u^2 + {}^{C}Y_u^2) + 2p_2\,{}^{C}X_u\,{}^{C}Y_u \tag{19}$$

$$\delta_{yd} = 2p_1\,{}^{C}X_u\,{}^{C}Y_u + p_2({}^{C}X_u^2 + 3\,{}^{C}Y_u^2) \tag{20}$$

[2] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965–980, 1992. ISSN 0162-8828. doi: 10.1109/34.159901. URL `http://dx.doi.org/10.1109/34.159901`

3/4: Lens distortion: Thin prism distortion

- From imperfection in lens design and manufacturing as well as camera assembly.
- Modeled by adding a thin prism to the optic system, causing radial and tangential distortions [Weng et al., 1992].

$$\delta_{xp} = s_1({}^C X_u^2 + {}^C Y_u^2) \qquad \delta_{yp} = s_2({}^C X_u^2 + {}^C Y_u^2) \tag{21}$$

3/4: Lens distortion: Merged distortion

Merge the all distortion as

$$\delta_{xp} = (g_1 + g_3)^C X_u^2 + g_4 {}^C X_u {}^C Y_u + g_1 {}^C Y_u^2 + k_1 {}^C X_u ({}^C X_u^2 + {}^C Y_u^2), \quad (22)$$
$$\delta_{yp} = g_2 {}^C X_u^2 + g_3 {}^C X_u {}^C Y_u + (g_2 + g_4)^C Y_u^2 + k_1 {}^C Y_u ({}^C X_u^2 + {}^C Y_u^2), \quad (23)$$

where

$$g_1 = s_1 + p_1$$
$$g_2 = s_2 + p_2$$
$$g_3 = 2p_1$$
$$g_4 = 2p_2$$

4/4: A 2D camera point to a 2D image point

Change the 2D camera coordinate system to the 2D image one.

- From a 2D point ${}^{C}\mathbf{X}_d$ in metric system w.r.t. the camera coordinate
- To a 2D point ${}^{I}\mathbf{X}_d$ in pixel system w.r.t. the camera coordinate

$$s \begin{bmatrix} {}^{I}X_d \\ {}^{I}Y_d \\ 1 \end{bmatrix} = \begin{bmatrix} -k_u & 0 & u_0 \\ 0 & -k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^{C}X_d \\ {}^{C}Y_d \\ 1 \end{bmatrix}, \tag{24}$$

$${}^{I}X_d = -k_u {}^{C}X_d + u_0 \qquad {}^{I}Y_d = -k_v {}^{C}Y_d + v_0,$$

where

- parameters (k_u, k_v) transform from metric measures to pixel.
- (u_0, v_0) define the projection of the focal point in the plain.

Camera calibration: General idea

Task

Compute camera parameters:

- Packed parameters $\mathbf{P}$.
- Each components $\mathbf{A}$, $\mathbf{R}$, and $\mathbf{t}$.

Given

- Known 3D points $\{\mathbf{X}_i | i = 1, \ldots, N\}$.
- Observed 2D points $\{\mathbf{x}_i | i = 1, \ldots, N\}$.

Camera calibration: Projective matrix estimation

Setting $p_{34} = 1$, i-th image point $\mathbf{x}_i$ is written as

$$x_i = \frac{X_i p_{11} + Y_i p_{12} + Z_i p_{13} + p_{14}}{X_i p_{31} + Y_i p_{32} + Z_i p_{33} + 1} \tag{25}$$

$$y_i = \frac{X_i p_{21} + Y_i p_{22} + Z_i p_{23} + p_{24}}{X_i p_{31} + Y_i p_{32} + Z_i p_{33} + 1} \tag{26}$$

Solve as an optimization problem w.r.t. $\mathbf{P}$ such as

1. Linear method 1 solves as $\mathbf{Ax} = \mathbf{b}$.
2. Linear method 2 solves as $\mathbf{Ax} = \mathbf{0}$.
3. Non-linear method solves non-linearly.

Camera calibration: Linear method 1

Proposed by [Hall et al., 1982] [3].
Eq. (25) and Eq. (26) is rewritten as

$$X_i p_{11} + Y_i p_{12} + Z_i p_{13} + p_{14} - x_i X_i p_{31} - x_i Y_i p_{32} - x_i Z_i p_{33} = x_i \tag{27}$$
$$X_i p_{21} + Y_i p_{22} + Z_i p_{23} + p_{24} - y_i X_i p_{31} - y_i Y_i p_{32} - y_i Z_i p_{33} = y_i \tag{28}$$

Given N corresponding points $\{\mathbf{X}_i\}$ and $\{\mathbf{x}_i\}$, generate following equation:

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1 X_1 & -x_1 Y_1 & -x_1 Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1 X_1 & -y_1 Y_1 & -y_1 Z_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & 0 & -x_N X_N & -x_N Y_N & -x_N Z_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -y_N X_N & -y_N Y_N & -y_N Z_N \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ \vdots \\ p_{32} \\ p_{33} \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ \vdots \\ x_N \\ y_N \end{bmatrix} \tag{29}$$

$$\rightarrow \mathbf{A}\mathbf{p} = \mathbf{b}$$

where $\mathbf{A} \in \mathbb{R}^{2N \times 11}$, $\mathbf{p} \in \mathbb{R}^{11}$, and $\mathbf{b} \in \mathbb{R}^{2N}$.

[3] E. L. Hall, J. B. K. Tio, C. A. McPherson, and F. A. Sadjadi. Measuring curved surfaces for robot vision. *Computer*, 15 (12):42–54, 1982. ISSN 0018-9162. doi: 10.1109/MC.1982.1653915. URL http://dx.doi.org/10.1109/MC.1982.1653915

Camera calibration: Linear method 1 cont.

Considering an energy function $E_1 = \|\mathbf{A}\mathbf{p} - \mathbf{b}\|^2$, projection matrix is obtained by minimizing E_1 as

$$\hat{\mathbf{p}} = \arg\min_{p} E_1 = \arg\min_{p} (\mathbf{A}\mathbf{p} - \mathbf{b})^{\mathsf{T}} (\mathbf{A}\mathbf{p} - \mathbf{b}) \tag{30}$$

Differentiating E_1 w.r.t. $\mathbf{p}$,

$$\begin{aligned} &\frac{\partial E_1}{\partial \mathbf{p}} = 0 \\ &\rightarrow \mathbf{A}^{\mathsf{T}}(\mathbf{A}\hat{\mathbf{p}} - \mathbf{b}) = 0 \\ &\rightarrow \mathbf{A}^{\mathsf{T}}\mathbf{A}\hat{\mathbf{p}} = \mathbf{A}^{\mathsf{T}}\mathbf{b} \\ &\rightarrow \hat{\mathbf{p}} = (\mathbf{A}^{\mathsf{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{T}}\mathbf{b} \end{aligned} \tag{31}$$

$\mathbf{p}$ can be estimated if $\mathbf{A}^{\mathsf{T}}\mathbf{A}$ is invertible.

Camera calibration: Linear method 1 cont.

This method heavily relies on whether the matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$ is invertible or not. Alternatively, we solve the problem by solving $\mathbf{A}\mathbf{x} = \mathbf{0}$ as Linear method 2 does.

Camera calibration: Linear method 2

Eq. (25) and Eq. (26) is rewritten as

$$X_i p_{11} + Y_i p_{12} + Z_i p_{13} + p_{14} - x_i X_i p_{31} - x_i Y_i p_{32} - x_i Z_i p_{33} - x_i p_{34} = 0 \tag{32}$$

$$X_i p_{21} + Y_i p_{22} + Z_i p_{23} + p_{24} - y_i X_i p_{31} - y_i Y_i p_{32} - y_i Z_i p_{33} - y_i p_{34} = 0 \tag{33}$$

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1 X_1 & -x_1 Y_1 & -x_1 Z_1 & -x_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1 X_1 & -y_1 Y_1 & -y_1 Z_1 & -y_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & 0 & -x_N X_N & -x_N Y_N & -x_N Z_N & -x_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -y_N X_N & -y_N Y_N & -y_N Z_N & -y_N \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ \vdots \\ p_{33} \\ p_{34} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \tag{34}$$

$\rightarrow \mathbf{Ap} = \mathbf{0}$

where $\mathbf{A} \in \mathbb{R}^{2N \times 12}$ is points matrix, $\mathbf{p} \in \mathbb{R}^{12}$ is unknown projection matrix parameters vector, and $\mathbf{b} \in \mathbb{R}^{2N}$ is 2D points vector

Camera calibration: Linear method 2 cont.

To obtain the non-trivial solution of homogeneous system $\mathbf{Ap} = \mathbf{0}$, apply constrained optimization.
Considering an energy function $E_2 = \|\mathbf{Ap}\|^2$ subject to the constraint $\|\mathbf{p}\|^2 - 1 = 0$, prevents $\mathbf{p}$ from becoming a zero vector.
With a Lagrange multiplier $\lambda > 0$, we obtain the following energy function

$$\begin{aligned} E_2(\mathbf{p}, \lambda) &= \|\mathbf{Ap}\|^2 - \lambda(\|\mathbf{p}\|^2 - 1) \\ &= (\mathbf{Ap})^{\mathbf{T}}(\mathbf{Ap}) - \lambda(\mathbf{p}^{\mathbf{T}}\mathbf{p} - 1). \end{aligned} \tag{35}$$

$$\hat{\mathbf{p}} = \arg\min_{\mathbf{p}} E_2(\mathbf{p}, \lambda) = \arg\min_{\mathbf{p}} (\mathbf{Ap})^{\mathbf{T}}(\mathbf{Ap}) - \lambda(\mathbf{p}^{\mathbf{T}}\mathbf{p} - 1) \tag{36}$$

Camera calibration: Linear method 2 cont.

Differentiating E_2 w.r.t. $\mathbf{p}$

$$\begin{aligned} &\frac{\partial E_2}{\partial \mathbf{p}} = 0 \\ &\rightarrow \mathbf{A}^\mathsf{T}\mathbf{A}\hat{\mathbf{p}} - \lambda\hat{\mathbf{p}} = 0 \\ &\rightarrow \mathbf{A}^\mathsf{T}\mathbf{A}\hat{\mathbf{p}} = \lambda\hat{\mathbf{p}} \end{aligned} \tag{37}$$

Differentiating E_2 w.r.t. λ

$$\begin{aligned} &\frac{\partial E_2}{\partial \lambda} = 0 \\ &\rightarrow \hat{\mathbf{p}}^\mathsf{T}\hat{\mathbf{p}} - 1 = 0 \\ &\rightarrow \|\hat{\mathbf{p}}\|^2 = 1 \end{aligned} \tag{38}$$

Camera calibration: Linear method 2 cont.

Pre-multiplying both sides of Eq. (37) by $\hat{\mathbf{p}}^{\mathsf{T}}$ gives

$$
\begin{aligned}
&\hat{\mathbf{p}}^{\mathsf{T}}\mathbf{A}^{\mathsf{T}}\mathbf{A}\hat{\mathbf{p}} = \lambda\hat{\mathbf{p}}^{\mathsf{T}}\hat{\mathbf{p}} \\
&\rightarrow (\mathbf{A}\hat{\mathbf{p}})^{\mathsf{T}}(\mathbf{A}\hat{\mathbf{p}}) = \lambda 1 \\
&\rightarrow \|\mathbf{A}\hat{\mathbf{p}}\|^2 = \lambda
\end{aligned} \tag{39}
$$

Eq. (39) is the same expression that $E_2 = \|\mathbf{Ap}\|^2$. This means that minimizing $\|\mathbf{Ap}\|^2$ is to minimize λ.

Camera calibration: Linear method 2 cont.

Differencing the energy function E_2 tells that

- Since Eq. (37) forms like $\mathbf{Ax} = \lambda\mathbf{x}$, $\hat{\mathbf{p}}$ should be an eigenvector of the matrix $\mathbf{A}^\mathbf{T}\mathbf{A}$ whose corresponding eigenvalue is λ.
- Eq. (38) minimizes λ as much as possible (ideally 0)

Thus, $\hat{\mathbf{p}}$ should be the eigenvector corresponding to the smallest eigenvalue of the matrix $\mathbf{A}^\mathbf{T}\mathbf{A}$.

Camera calibration: Non-linear method

Consider the following reprojection error:

$$E_{\mathrm{NL}} = \sum_{i=1}^{N} \left\| x_i - \frac{X_i p_{11} + Y_i p_{12} + Z_i p_{13} + p_{14}}{X_i p_{31} + Y_i p_{32} + Z_i p_{33} + p_{34}} \right\|^2 + \left\| y_i - \frac{X_i p_{21} + Y_i p_{22} + Z_i p_{23} + p_{24}}{X_i p_{31} + Y_i p_{32} + Z_i p_{33} + p_{34}} \right\|^2 . \tag{40}$$

- Need good initial guess (use linear method's result).

The method of Faugeras with radial distortion

Let's derive the Formulation

$$\begin{bmatrix} {}^{C}X_u & {}^{C}Y_u \end{bmatrix}^{\mathsf{T}} = \frac{f}{{}^{C}Z_w} \begin{bmatrix} {}^{C}X_w & {}^{C}Y_w \end{bmatrix}^{\mathsf{T}} \tag{41}$$

$$\begin{cases} {}^{C}X_w = r_{11}{}^{W}X_w + r_{12}{}^{W}Y_w + r_{13}{}^{W}Z_w + t_x \\ {}^{C}Y_w = r_{21}{}^{W}X_w + r_{22}{}^{W}Y_w + r_{23}{}^{W}Z_w + t_y \\ {}^{C}Z_w = r_{31}{}^{W}X_w + r_{32}{}^{W}Y_w + r_{33}{}^{W}Z_w + t_z \end{cases} \tag{42}$$

$$\begin{bmatrix} {}^{C}X_u & {}^{C}Y_u \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} {}^{C}X_d + \delta_x & {}^{C}Y_d + \delta_y \end{bmatrix}^{\mathsf{T}} \tag{43}$$

$$\begin{cases} \delta_{xr} = k_1{}^{C}X_d({}^{C}X_d^2 + {}^{C}Y_d^2) \\ \delta_{yr} = k_1{}^{C}Y_d({}^{C}X_d^2 + {}^{C}Y_d^2) \end{cases} \tag{44}$$

$$\begin{bmatrix} {}^{I}X_d & {}^{I}Y_d \end{bmatrix}^{\mathsf{T}} = \begin{bmatrix} -k_u{}^{C}X_d + u_0 & -k_v{}^{C}Y_d + v_0 \end{bmatrix}^{\mathsf{T}} \tag{45}$$

The method of Faugeras with radial distortion cont.

Since we only know $\begin{bmatrix} {}^W X_w & {}^W Y_w & {}^W Z_w \end{bmatrix}^\mathbf{T}$ and $\begin{bmatrix} {}^I X_d & {}^I Y_d \end{bmatrix}^\mathbf{T}$,

$$\begin{cases} f \frac{{}^C X_w}{{}^C Z_w} - {}^C X_d + k_1 {}^C X_d ({}^C X_d^2 + {}^C Y_d^2) = 0 \\ f \frac{{}^C Y_w}{{}^C Z_w} - {}^C Y_d + k_1 {}^C Y_d ({}^C X_d^2 + {}^C Y_d^2) = 0 \end{cases} \tag{46}$$

where

$$\begin{cases} {}^C X_w = r_{11} {}^W X_w + r_{12} {}^W Y_w + r_{13} {}^W Z_w + t_x \\ {}^C Y_w = r_{21} {}^W X_w + r_{22} {}^W Y_w + r_{23} {}^W Z_w + t_y \\ {}^C Z_w = r_{31} {}^W X_w + r_{32} {}^W Y_w + r_{33} {}^W Z_w + t_z \\ {}^C X_d = \frac{{}^I X_d - u_0}{-k_u} \\ {}^C Y_d = \frac{{}^I Y_d - v_0}{-k_v} \end{cases} \tag{47}$$

The method of Faugeras with radial distortion cont.

Finally, we derive the two energy function

$$U(x) = f\frac{r_{11}{}^{W}X_w + r_{12}{}^{W}Y_w + r_{13}{}^{W}Z_w + t_x}{r_{31}{}^{W}X_w + r_{32}{}^{W}Y_w + r_{33}{}^{W}Z_w + t_z} - \frac{{}^{I}X_d - u_0}{-k_u} + k_1\frac{{}^{I}X_d - u_0}{-k_u}r^2 \tag{48}$$

$$V(x) = f\frac{r_{21}{}^{W}X_w + r_{22}{}^{W}Y_w + r_{23}{}^{W}Z_w + t_y}{r_{31}{}^{W}X_w + r_{32}{}^{W}Y_w + r_{33}{}^{W}Z_w + t_z} - \frac{{}^{I}Y_d - v_0}{-k_v} + k_1\frac{{}^{I}Y_d - v_0}{-k_v}r^2 \tag{49}$$

where x packs all unknown parameters and

$$r^2 = \sqrt{\left(\frac{{}^{I}X_d - u_0}{-k_u}\right)^2 + \left(\frac{{}^{I}Y_d - v_0}{-k_v}\right)^2}. \tag{50}$$

The method of Faugeras with radial distortion cont.

Since the rotation matrix can be described by 3 rotation angles, instead of r_{ij}, we use 3 parameters α, β, and γ as

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \tag{51}$$

$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \tag{52}$$

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{53}$$

The method of Faugeras with radial distortion cont.

Since our energy functions, Eq. (48) and Eq. (49), are non-linear, the method of Faugeras solves the problem with iterative non-linear optimization as

$$G(x_k) \approx G(x_{k-1}) + J\Delta x_k, \tag{54}$$

where x_k denotes the k-th estimated parameters, $G(x)$ is the minimization function, and J denotes its Jacobian matrix as

$$G(x_{k-1}) = \begin{bmatrix} U_1(x_{k-1}) \\ V_1(x_{k-1}) \\ \vdots \\ U_n(x_{k-1}) \\ V_n(x_{k-1}) \end{bmatrix} J = \begin{bmatrix} \frac{\partial U_1(x_{k-1})}{\partial \alpha} & \frac{\partial U_1(x_{k-1})}{\partial \beta} & \cdots & \frac{\partial U_1(x_{k-1})}{\partial k_1} \\ \frac{\partial V_1(x_{k-1})}{\partial \alpha} & \frac{\partial V_1(x_{k-1})}{\partial \beta} & \cdots & \frac{\partial V_1(x_{k-1})}{\partial k_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial U_n(x_{k-1})}{\partial \alpha} & \frac{\partial U_n(x_{k-1})}{\partial \beta} & \cdots & \frac{\partial U_n(x_{k-1})}{\partial k_1} \\ \frac{\partial V_n(x_{k-1})}{\partial \alpha} & \frac{\partial V_n(x_{k-1})}{\partial \beta} & \cdots & \frac{\partial V_n(x_{k-1})}{\partial k_1} \end{bmatrix} \tag{55}$$

The method of Faugeras with radial distortion cont.

Then, we compute Δx_k using J and $G(x_{k-1})$ as

$$\Delta x_k = -(J^{\mathsf{T}} J)^{-1} J^{\mathsf{T}} G(x_{k-1}) \tag{56}$$

and update the current estimate as $x_k = x_{k-1} + \Delta x_k$.

How to choose the points set

Simply speaking, 6 corresponding points are required because

- we have 12 or 11 unknowns.
- One corresponding points give two equations.

Any points are fine for optimization?
The answer is NO.

Rank issues

For a $n \times m$ matrix $\mathbf{A}$, $n \leq m$,

$$\operatorname{rank}(\mathbf{A}) + \operatorname{null}(\mathbf{A}) = \min(n, m), \tag{57}$$

where $\operatorname{null}(\mathbf{A})$ represents the dimension of the null space of $\mathbf{A}$.
Since $n \leq m = 12$ in our case, we have to consider three cases:

1. $\operatorname{rank}(\mathbf{P}) = 12$: The null space has 0 dimension means there is only one solution, namely $\mathbf{p} = \mathbf{0}$.
2. $\operatorname{rank}(\mathbf{P}) = 11$: The null space has 1 dimension and there is a unique solution up to s scale factor.
3. $\operatorname{rank}(\mathbf{P}) < 11$: The null space has 2 or more dimension. The null vector $\mathbf{p}$ can be any vector in the dimensional space. This means that there exist infinite number of solutions.

The third case happens when all the points $\{X_i\}$ are located on a plane.

Camera calibration: Projective matrix decomposition

Now, we have

- An estimate of projective matrix $\mathbf{P}$.
- A set of corresponding points $\{\mathbf{X}_i\}$ and $\{\mathbf{x}_i\}$.

Next task is to decompose $\mathbf{P}$ into $\mathbf{A}$, $\mathbf{R}$, and $\mathbf{t}$.
Basically, we use constraint on matrix form

Projective matrix decomposition cont.

Consider

$$
\begin{aligned}
\mathbf{P} &= \mathbf{A}[\mathbf{R}|\mathbf{t}] \\
&= \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \\
&= \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{r}_1^{\mathbf{T}} & t_1 \\ \mathbf{r}_2^{\mathbf{T}} & t_2 \\ \mathbf{r}_3^{\mathbf{T}} & t_3 \end{bmatrix} \\
&= \begin{bmatrix} \alpha_x \mathbf{r}_1^{\mathbf{T}} + s\mathbf{r}_2^{\mathbf{T}} + x_0 \mathbf{r}_3^{\mathbf{T}} & \alpha_x t_x + s \cdot t_2 + x_0 t_z \\ \alpha_y \mathbf{r}_2^{\mathbf{T}} + y_0 \mathbf{r}_3^{\mathbf{T}} & \alpha_x t_y + y_0 t_z \\ \mathbf{r}_3^{\mathbf{T}} & t_z \end{bmatrix}
\end{aligned} \tag{58}
$$

where $\mathbf{r}_j^{\mathbf{T}} = (r_{j1}, r_{j2}, r_{j3})$, for $j = 1, 2, 3$.

Projective matrix decomposition cont.

If the camera parameters follows Eq. (58) format, following conditions should be satisfied:

1. $\|\mathbf{p}_3\| = 1$, and
   - Since $\mathbf{p}_3 = \mathbf{r}_3$ and $\|\mathbf{r}_3\| = 1$.
2. $(\mathbf{p}_1 \wedge \mathbf{p}_3) \cdot (\mathbf{p}_2 \wedge \mathbf{p}_3) = 0$.
   - Using $\mathbf{r}_i$ and $\mathbf{r}_j$, $i \neq j$, is orhothogonal.

Camera calibration methods

1. [Hall et al., 1982]
2. [Faugeras and Toscani, 1986]
3. [Salvi et al., 2002]
4. [Tsai, 1987]
5. [Weng et al., 1992]

Camera calibration type

- Lens distortion: Linear vs. Non-linear
- Camera parameters: Intrinsic vs. Extrinsic
- Computing camera parameters: Implicit vs. Explicit
- Calibration pattern: known 3D points vs. a reduced set of 3D points

References I

O. D. Faugeras and G. Toscani. The calibration problem for stereo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE Publ.86CH2290-5, pages 15–20. IEEE, 1986.

E. L. Hall, J. B. K. Tio, C. A. McPherson, and F. A. Sadjadi. Measuring curved surfaces for robot vision. *Computer*, 15(12):42–54, 1982. ISSN 0018-9162. doi: 10.1109/MC.1982.1653915. URL `http://dx.doi.org/10.1109/MC.1982.1653915`.

F. Remondino and C. Fraser. Digital camera calibration methods: Considerations and comparisons. In *ISPRS Commision V Symposium*, volume 6, pages 266–272, 2006.

J. Salvi, X. Armangue, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35:1617–1635, 2002.

R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Transactions on Robotics and Automation*, 3(4):323–344, 1987.

J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10): 965–980, 1992. ISSN 0162-8828. doi: 10.1109/34.159901. URL `http://dx.doi.org/10.1109/34.159901`.

Z. Zhang. Camera calibration. In G. Medioni and S. B. Kang, editors, *Emerging Topics in Computer Vision*, chapter 2, pages 4–43. Prentice Hall, 2005.