Computer Vision
and
Machine Learning
(Camera calibration and correction)

Bhabatosh Chanda
chanda@isical.ac.in

Position and orientation of camera

- Camera may be placed anywhere in a scene.
- Same object may appear differently while captured by same camera at different position and orientation.
- It would be easier to handle if the relation between their geometry is known.

Coordinate transformation

- World coordinate system
  - World (object) centric scene coordinates
- Camera coordinate system
  - Camera centric scene coordinates
- Transforming world (scene) coordinate to camera coordinate
  - Rotation, translation and (optional) scaling
  - Linear transformation is advantageous.

Projection matrix for image formation

In practice: lots of coordinate transformations…

$$\begin{pmatrix}\text{2D point (3x1)}\end{pmatrix} = \begin{pmatrix}\text{Camera to pixel coord. trans. matrix (3x3)}\end{pmatrix}\begin{pmatrix}\text{Perspective projection matrix (3x4)}\end{pmatrix}\begin{pmatrix}\text{World to camera coord. trans. matrix (4x4)}\end{pmatrix}\begin{pmatrix}\text{3D point (4x1)}\end{pmatrix}$$

$$\begin{bmatrix} u^c\varphi_x + w^c d_x/f \\ v^c\varphi_y + w^c d_y/f \\ w^c/f \end{bmatrix} = \begin{bmatrix} \varphi_x & 0 & d_x \\ 0 & \varphi_y & d_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_u \\ r_{32} & r_{22} & r_{23} & t_v \\ r_{31} & r_{32} & r_{33} & t_w \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} u^c\varphi_x + w^c d_x/f \\ v^c\varphi_y + w^c d_y/f \\ w^c/f \end{bmatrix} = \begin{bmatrix} \varphi_x & 0 & d_x/f & 0 \\ 0 & \varphi_y & d_y/f & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_u \\ r_{32} & r_{22} & r_{23} & t_v \\ r_{31} & r_{32} & r_{33} & t_w \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix}$$

Projection matrix for image formation

- Combining the matrices of

$$\begin{bmatrix} u^c\varphi_x + w^c d_x/f \\ v^c\varphi_y + w^c d_y/f \\ w^c/f \end{bmatrix} = \begin{bmatrix} \varphi_x & 0 & d_x/f \\ 0 & \varphi_y & d_y/f \\ 0 & 0 & 1/f \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_u \\ r_{32} & r_{22} & r_{23} & t_v \\ r_{31} & r_{32} & r_{33} & t_w \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix}$$

- We obtain $\begin{bmatrix} u^c\varphi_x + w^c d_x/f \\ v^c\varphi_y + w^c d_y/f \\ w^c/f \end{bmatrix} = K[R \quad t] \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix} = A \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix}$

- Finally, $\begin{bmatrix} x' \\ y' \\ w^c/f \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{32} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix}$

Camera calibration

Finally, $\begin{bmatrix} x' \\ y' \\ w_c/f \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix}$

$$\begin{bmatrix} \lambda x \\ \lambda y \\ \lambda \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ 1 \end{bmatrix}$$

where $x = \frac{x'}{\lambda}$; $y = \frac{y'}{\lambda}$; i.e., (x, y) is image coordinate.

Camera calibration

- For i-th world point (u_i, v_i, w_i) and corresponding image point (x_i, y_i) we have

$$\begin{bmatrix} \lambda_i x_i \\ \lambda_i y_i \\ \lambda_i \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} u_i \\ v_i \\ w_i \\ 1 \end{bmatrix}$$

- Determining the elements a_{ij} of matrix A is called ***camera calibration.***
- Note that ***matrix A*** contains ***both intrinsic and extrinsic parameters***.

Camera calibration

- Matrix A has 12 elements: 3 rows each with 4 elements.

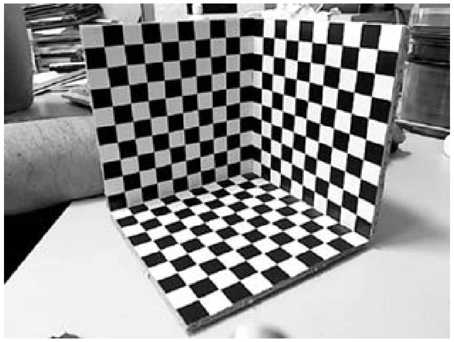
$$p_i^I = \begin{bmatrix} \lambda_i x_i \\ \lambda_i y_i \\ \lambda_i \end{bmatrix} = Ap_i = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \begin{bmatrix} u_i \\ v_i \\ w_i \\ 1 \end{bmatrix}$$

- Each image point provides two linear equations.

$$x_i = \frac{\lambda_i x_i}{\lambda_i} = \frac{r_1 \cdot p_i}{r_3 \cdot p_i} \quad \text{and} \quad y_i = \frac{\lambda_i y_i}{\lambda_i} = \frac{r_2 \cdot p_i}{r_3 \cdot p_i}$$

- 6 world points and corresponding image points give solution.

Camera calibration grid



Origin may be assumed at the junction of three faces.

Camera calibration

- World points are located on *calibration grid* with known *geometric structure*.
- Much more than six points are taken and solution is obtained by least square estimation.
- Solutions are correct up to a scale.

Camera calibration

- For i-th world point $(u_i, v_i, w_i)$and corresponding image point (x_i, y_i) we have

$$\begin{bmatrix} \lambda_i x_i \\ \lambda_i y_i \\ \lambda_i \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} u_i \\ v_i \\ w_i \\ 1 \end{bmatrix}$$

- There are 12 unknowns a_{ij}.
- So to solve we need at least 12 linear equations relating image points and world points.
- Solution is correct up to a scale.

Camera calibration

- Once we determine the camera projection matrix A

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = K[R \quad t]$$

- Thus we can write

$$\hat{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = KR \text{ and } \begin{bmatrix} a_{14} \\ a_{24} \\ a_{34} \end{bmatrix} = Kt$$

($\hat{A} = KR$ can be solved by RQ-decomposition, where R is an upper-triangular matrix and Q is an orthonormal matrix.)

Stereo: Anaglyph

Numerical problems

- Suppose the tip of a lamppost forms an image on the image plane of a digital camera having focal length 1.5 cm and image plane of size 4 cm x 3 cm. The tip of the lamppost is located at (2 m., 5 m. 10 m.) with respect to the centre of the lens of the camera. Determine the image coordinate in terms of pixels, if the photoreceptor density is 1200/cm in both horizontal and vertical direction.
- A digital camera has the intrinsic parameters as: focal length = 0.8 cm, photoreceptor density 1000 per cm in both directions, image plane height = 2 cm and width = 3 cm. When imaged a particular object point appears at (1580, 1320) in pixels. The camera is shifted by 5m along horizontal axis. Now the image of the same object point appears at (1380, 1320) in pixels. Calculate the location of the object point with respect to the initial position of the camera.

Thank you!

Any question?