Introduction to Robotics

Chapter IX

By Prof. Song LIU

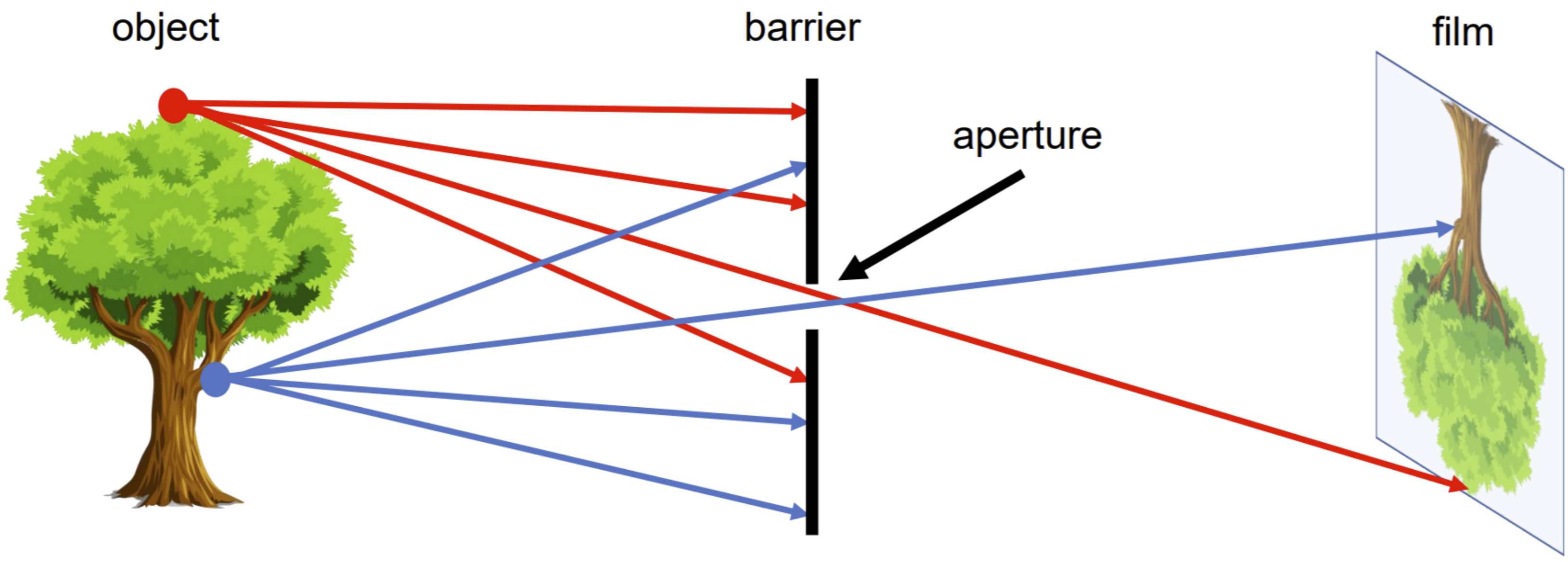
AMNR Lab Director

liusong@shanghaitech.edu.cn

Visual Measurement

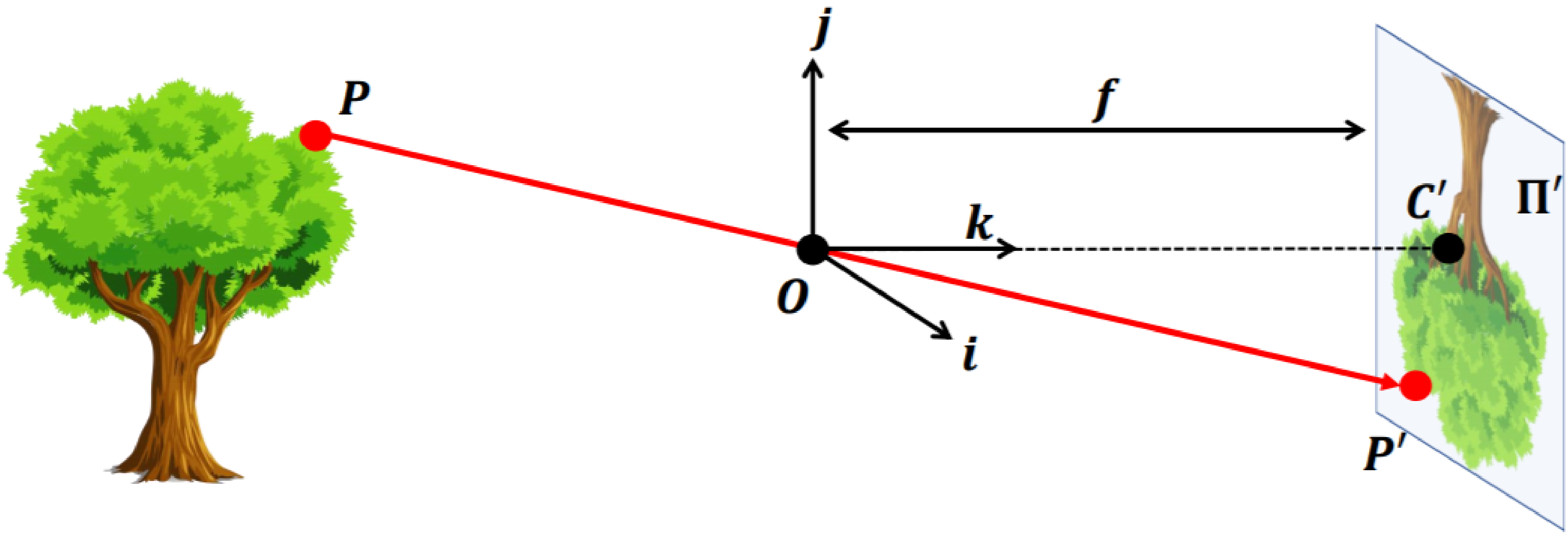
- Pinhole Camera Model
- Constraint Conditions of Visual Measurement
- Position Measurement based on Monocular Vision
- Position Measurement based on Stereo Vision
- Pose Measurement based on Rectangular Object Constraint
- Pose Measurement based on PnP Problem
- Visual Measurement base on Vanishing Point

Pinhole Camera Model



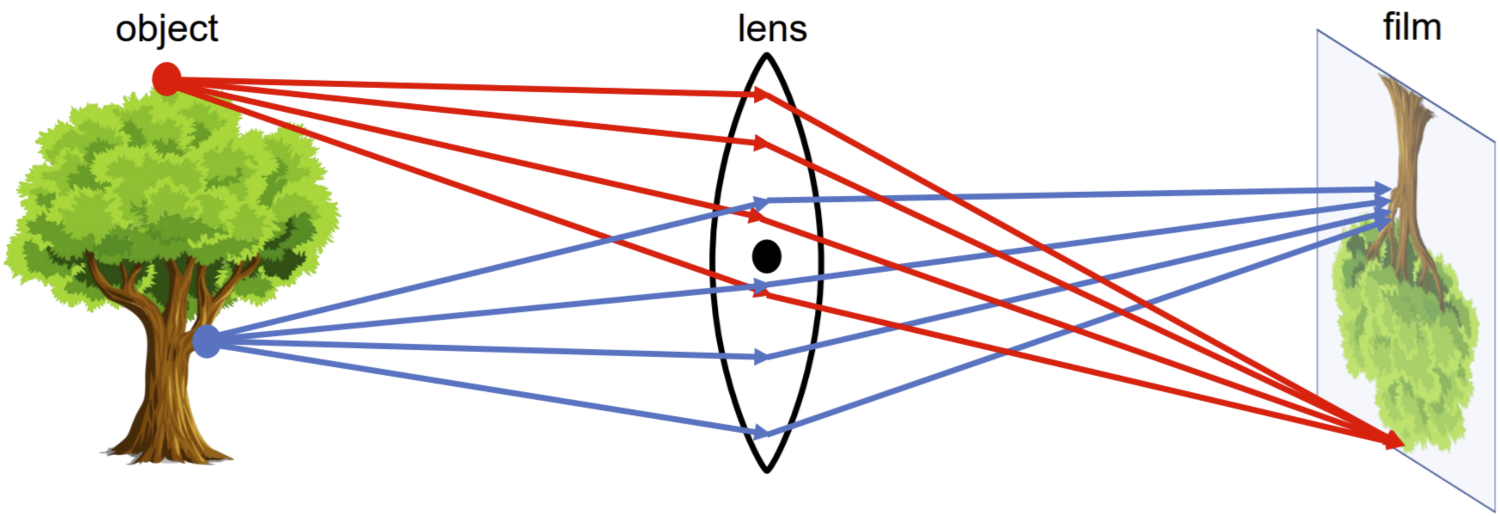
- Camera – a system that can record an image of an object or scene in the 3D world, which essentially is a 3D-to-2D mapping system.
- This camera system can be designed by placing a barrier with a small **aperture** between the 3D object and a photographic **film** or sensor.

Pinhole Camera Model



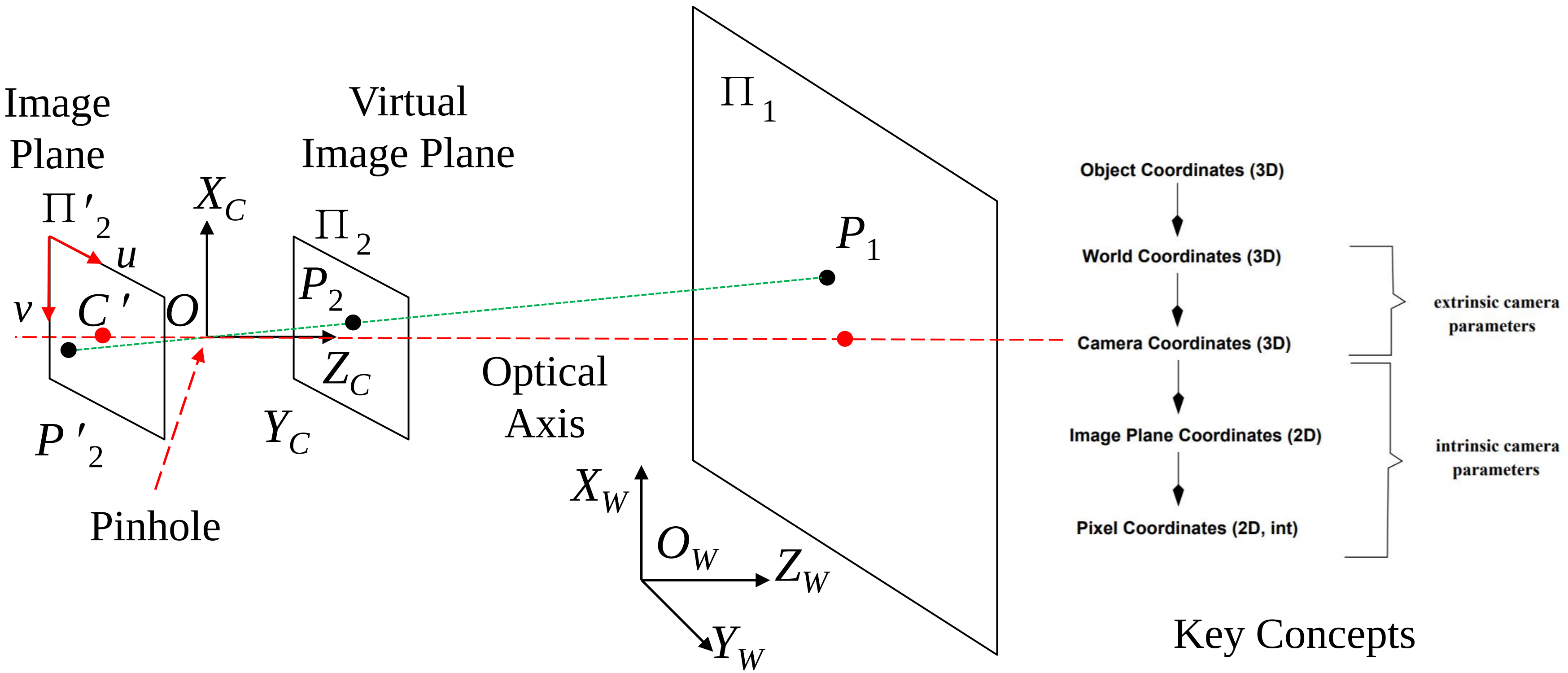
- A more formal construction of the pinhole.
- The film is commonly called the **image or retinal plane**.
- The aperture is referred to as the **pinhole *O*** or center of the camera.
- The distance between the image plane and the pinhole O is the **focal length *f***.
- Sometimes, the retinal plane is placed between ***O*** and the 3D object at a distance ***f*** from ***O***. In this case, it is called **virtual image or virtual retinal plane**.
- The line defined by ***C`*** and ***O*** is called the **optical axis** of the camera system.
- Notice that one large assumption we make in this pinhole model is that the aperture is a single point.

Cameras and Lens



- The strong assumption of pinhole model is mitigated by lenses to make sure enough light passing through to reach the image plane.
- The Concept of "**Depth of Field**": in photography and computer graphics, DOF is the effective range at which cameras can take clear photos.
- In this case, the **focal length** has different meanings compared with Pinhole model.
- Optical lens induces optical **distortion**.

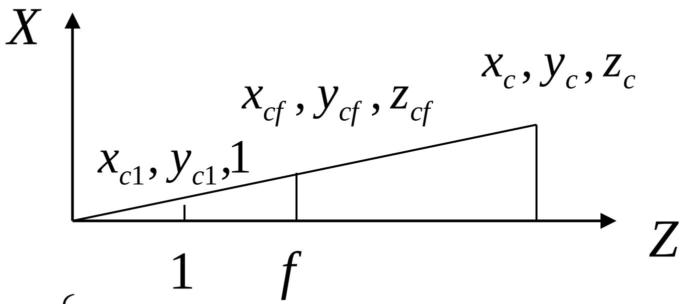
Pinhole Camera Model



- The $X_C Y_C Z_C$-O_C is called **camera reference system** or **camera coordinate system**.
- The $X_W Y_W Z_W$-O_W is the world coordinate system.
- *uv* is the image coordinate system.

Pin-hole Model: Intrinsic Parameter

The parameters necessary to link the pixel coordinates of an image point with the corresponding coordinates in the camera reference frame.



$$\begin{cases} \dfrac{x_{cf}}{f} = \dfrac{x_c}{z_c} \\[2ex] \dfrac{y_{cf}}{f} = \dfrac{y_c}{z_c} \end{cases} \qquad \begin{cases} x_{c1} = \dfrac{x_c}{z_c} \\[2ex] y_{c1} = \dfrac{y_c}{z_c} \end{cases}$$

$$\begin{cases} u - u_0 = \alpha_x x_{cf} = \alpha_x f \dfrac{x_c}{z_c} = k_x x_{c1} \\[2ex] v - v_0 = \alpha_y y_{cf} = \alpha_y f \dfrac{y_c}{z_c} = k_y y_{c1} \end{cases} \Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{c1} \\ y_{c1} \\ 1 \end{bmatrix}$$

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}$$

Four Parameter Pin-hole Model:

$$M_{in} = \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \qquad M_{in} = \begin{bmatrix} k_x & \gamma & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

We can easily calibrate M_{in} with four known 3D point. A skew parameter γ sometimes is introduced to model the CCD fabrication error, where u-axis is not seriously orthogonal to v-aixs.

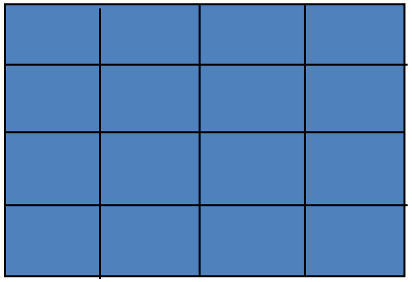
Pin-Hole Model – Lens Distortion

- Second-order Radial Distortion

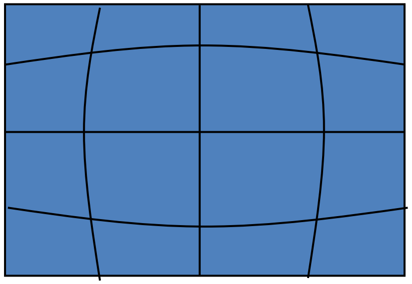
$$\begin{cases} u - u_0 = (u' - u_0)(1 + k_u' r^2) \\ v - v_0 = (v' - v_0)(1 + k_v' r^2) \end{cases}$$

$[u', v']$ is ideal image coordinate without distortion, $[u, v]$ is distorted image coordinate, $[u_0, v_0]$ is image coordinate of optical center (reference point), $[k'_u, k'_v]$ are distortion coefficients along u、v-axis, r is distance from image point to reference point.

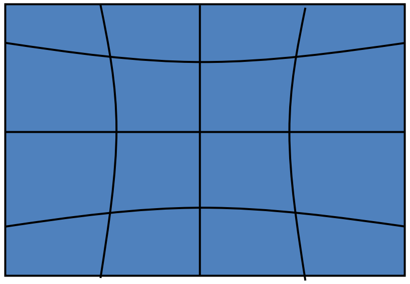
$$r = \sqrt{(u' - u_0)^2 + (v' - v_0)^2}$$



Ideal Image



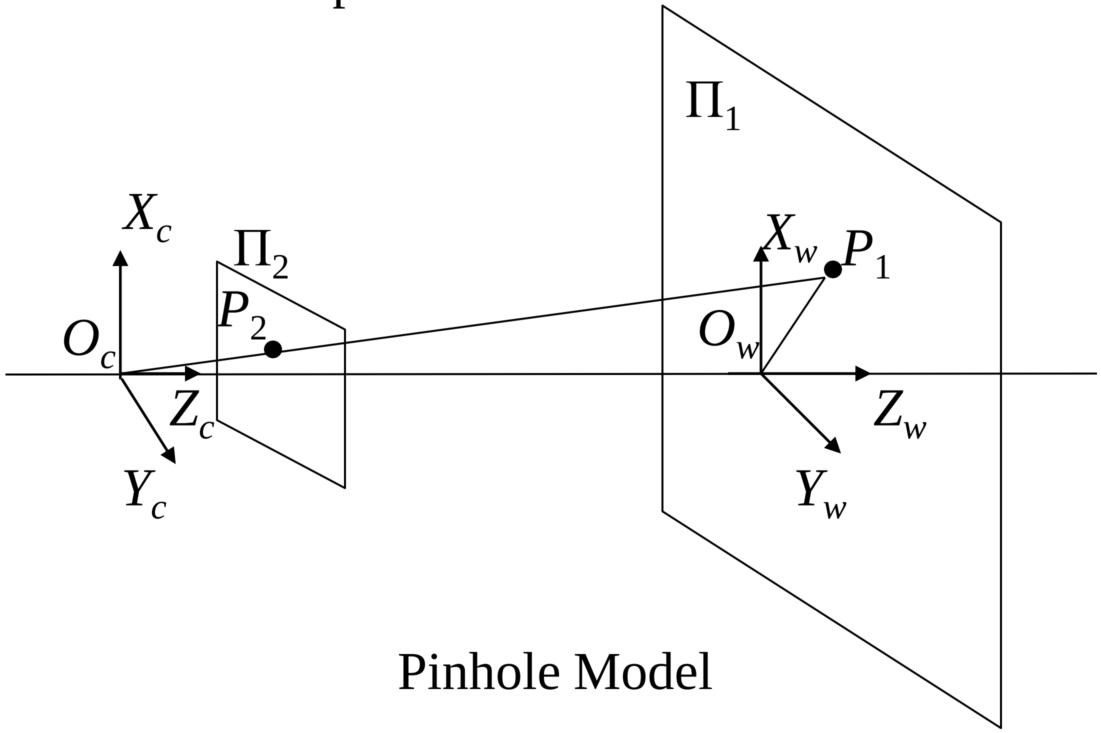
Distortion Coefficient < 0
Barrel Distortion



Distortion Coefficient > 0
Pincushion Distortion

Pin-hole Model – Extrinsic Parameter

The parameters that define the location and orientation of the camera reference frame with respect to a known world reference frame.



Pinhole Model

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = {}^{c}M_{w} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

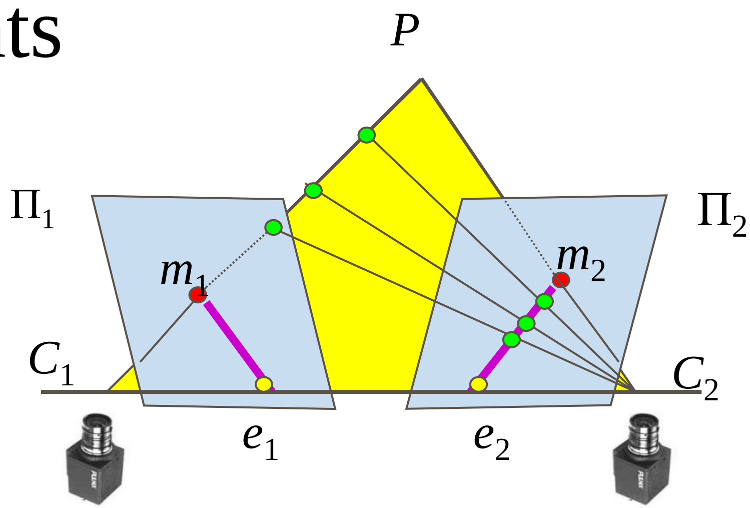
H – Homography Matrix

$$\begin{bmatrix} x_{c1} \\ y_{c1} \\ z_{c1} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_{c2} \\ y_{c2} \\ z_{c2} \end{bmatrix} \Rightarrow \frac{z_{c1}}{z_{c2}} \begin{bmatrix} x_{1c1} \\ y_{1c1} \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_{1c2} \\ y_{1c2} \\ 1 \end{bmatrix} = H \begin{bmatrix} x_{1c2} \\ y_{1c2} \\ 1 \end{bmatrix}$$

H relates points on the two plane of $z_c = 1$ of two cameras.

Constraint Conditions of Visual Measurement

- Stereo Matching Constraints
  - Epipolar line constraint



Epipolar Geometry

Epipolar Plane: the plane formed by object point P and camera center C_1, C_2.

Epipolar Line: the intersection line between epipolar plane and imaging plane.

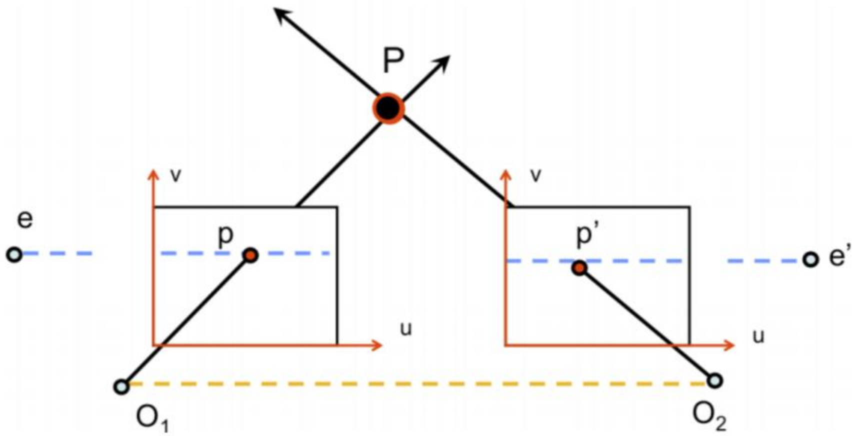
Epipole: the intersection points between the line (**baseline**) connecting C_1 ,C_2 and the imaging plane.

Stereo Matching: the projection points m_1 and m_2 (stereo pair) of object point P on imaging plane are conjugate points. Stereo matching is to find the conjugate points.

m_1 and m_2 are also called feature points. With m_1 determined, and pre-calibrated relationship between cameras, the epipolar line connecting m_2 and e_2 can be determined. Then m_2 must be on the epipolar line.

Constraint Conditions of Visual Measurement

- Stereo Matching Constraints
  - Epipolar Geometry in Real Applications



- In practice, the two image planes are parallel to each other.
- When the image planes are parallel to each other, then the epipoles and will be located at infinity since the baseline is parallel to the image planes.
- Another important byproduct of this case is that the epipolar lines are parallel to an axis of each image plane.

The BIG question is how to mathematically describe the epipolar line constraint?

Stereo Matching

➢Basic Constraint

$$Z_c\begin{bmatrix}u\\v\\1\end{bmatrix}=\begin{bmatrix}k_x & 0 & u_0 & 0\\ 0 & k_y & v_0 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}\begin{bmatrix}R & t\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_w\\y_w\\z_w\\1\end{bmatrix}=M_{in}M_{ex}\begin{bmatrix}x_w\\y_w\\z_w\\1\end{bmatrix}=M\begin{bmatrix}x_w\\y_w\\z_w\\1\end{bmatrix}$$

Projective Transformation

So we have: $M=\begin{bmatrix}m_{11} & m_{12} & m_{13} & m_{14}\\ m_{21} & m_{22} & m_{23} & m_{24}\\ m_{31} & m_{32} & m_{33} & m_{34}\end{bmatrix}$

Regarding the same object point, for the two cameras, we have:

$$\begin{cases}z_{c1}u_1=M_1X_p=\begin{bmatrix}M_{11} & M_{12}\end{bmatrix}X_p\\ z_{c2}u_2=M_2X_p=\begin{bmatrix}M_{21} & M_{22}\end{bmatrix}X_p\end{cases}\qquad X_p=\begin{bmatrix}x_w & y_w & z_w & 1\end{bmatrix}^T=\begin{bmatrix}X & 1\end{bmatrix}^T\qquad u_i=\begin{bmatrix}u_{i1} & v_{i1} & 1\end{bmatrix}^T$$

$$\Downarrow$$

$$\begin{cases}z_{c1}u_1=M_{11}X+M_{12}\\ z_{c2}u_2=M_{21}X+M_{22}\end{cases}\Rightarrow\quad\begin{aligned}z_{c2}u_2&=M_{21}M_{11}^{-1}(z_{c1}u_1-M_{12})+M_{22}\\ z_{c2}u_2-M_{21}M_{11}^{-1}z_{c1}u_1&=M_{22}-M_{21}M_{11}^{-1}M_{12}\end{aligned}$$

Stereo Matching

Antisymmetric Matrix

$$[t]_\times = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}$$

For Vector Cross Product：

$$t \times r = [t]_\times r$$

if $[t]_\times r = 0$, then $r = kt$

Define $m = M_{22} - M_{21} M_{11}^{-1} M_{12}$

$$[m]_\times (z_{c2} u_2 - M_{21} M_{11}^{-1} z_{c1} u_1) = [m]_\times m = 0$$

Both sides divided by z_{c2}

$$[m]_\times z_c M_{21} M_{11}^{-1} u_1 = [m]_\times u_2, \quad z_c = \frac{z_{c1}}{z_{c2}}$$

Both sides time u_2^T

$$u_2^T [m]_\times z_c M_{21} M_{11}^{-1} u_1 = u_2^T [m]_\times u_2 = 0$$

$$\boxed{u_2^T [m]_\times M_{21} M_{11}^{-1} u_1 = 0}$$

This is Epipolar Constraint, where u_1, u_2 are homogeneous coordinates

$u_2^T F u_1 = 0$ Fundamental Matrix: $F = [m]_\times M_{21} M_{11}^{-1}$

Essential Matrix: $E = M_{in}^T F M_{in}$

With $[u_{11}, v_{11}]^T$ known, u_{21} and v_{21} can be calculated.

Stereo Matching

- Both Essential Matrix and Fundamental Matrix relate corresponding points in two images to mathematically describe the epipolar line constraint for stereo pair.
- The difference is that in the case of the Fundamental matrix, the points are in pixel coordinates, while in the case of the Essential matrix, the points are in 3D physical points in camera reference frame.
- Normalized image coordinates have the origin at the optical center of the image, and the x and y coordinates are normalized by Fx and Fy respectively.
- The two matrices are related as follows: E = ${M_{in}}^T$ * F * M_{in}, where M_{in} is the intrinsic matrix of the camera.
- The Essential Matrix is only concerned with the relative pose between the two camera coordinate systems. ***T*** is translation vector, while ***R*** is relative orientation.

$$E = [T_\times]R$$

Stereo Vision based Measurement (General Perspetive)

Constraint between stereo pair (Epipolar Constraint):

$$\boldsymbol{p}^{\mathrm{T}} F \boldsymbol{p}^{*} = 0$$

F is fundamental matrix, $\boldsymbol{p}$ is homogeneous coordinate in image coordinate, and further:

$$\mathrm{E} = M_{in}{}^{T} * F * M_{in} \qquad E = [T_{\times}]R$$

- There are 8 free parameters in $\boldsymbol{F}$, 3×3 matrix (why?).
- With 8 stereo pairs, we can calibrate $\boldsymbol{F}$.
- With the calibrated $\boldsymbol{M_{in}}$, $\boldsymbol{E}$ can be calculated.
- From E, by singular value decomposition (SVD), we can calculate $\boldsymbol{R}$ and $\boldsymbol{T}_{\mathrm{e}}$. $\boldsymbol{T}_{\mathrm{e}}$ and $\boldsymbol{T}$ is parallel to each other by a scalar.
- With $\boldsymbol{R}$ and $\boldsymbol{T}$, we can estimate the scene depth and finally the 3D world.

Stereo Matching Constraint

✓Consistency Constraint: the stereo pair of the same object point probability have significantly different intensities. So the images should be pre-processed before stereo matching

$$\bar{f}_0(i,j) = (f_0(i,j) - \mu_0)/\sigma_0$$
$$\bar{f}_k(i,j) = (f_k(i,j) - \mu_k)/\sigma_k$$

μ is the mean intensity of pixels within window size of $m \times n$, σ is the intensity variance.

$$\sigma = \frac{1}{mn}\sum_{j=1}^{n}\sum_{i=1}^{m}(f(i,j) - \mu)^2$$

Similarity Evaluation Function (one possible):

$$\varepsilon_k = \frac{1}{mn}\sum_{j=1}^{n}\sum_{i=1}^{m}\left|\bar{f}_0(i,j) - \bar{f}_k(i,j)\right|$$

✓Uniqueness Constraint: For one feature point from one image, there is a unique feature point in another counterpart image to form a stereo pair.

✓Continuity Constraint: The surface of an object is generally smooth. Therefore, the projection of points on the surface of an object on the image is also continuous, and the parallax is also continuous. At the boundary of the object, the continuity constraint is not valid .

Constraints for Visual Measurement

- Invariance Constraints Determined by Projective Geometry:

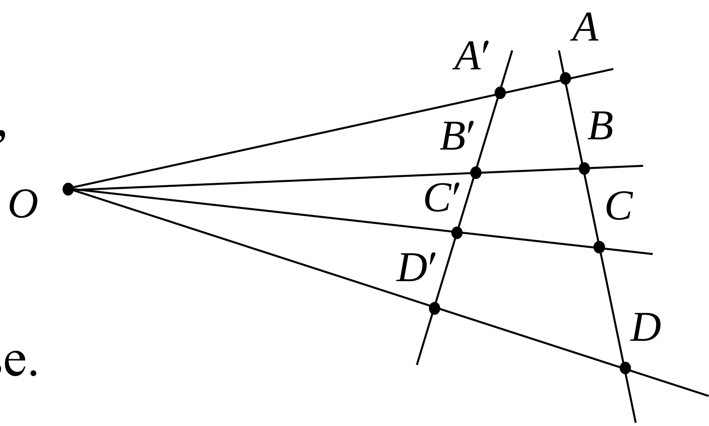
➢ Define Cross-Ratio: The cross-ratio of a quadruple of distinct points on the real line with coordinates z_1, z_2, z_3, z_4 is given by

$$(z_1, z_2; z_3, z_4) = \frac{(z_3 - z_1)(z_4 - z_2)}{(z_3 - z_2)(z_4 - z_1)}.$$

➢ Regarding projective transformation, the cross-ratio of line keeps constant.

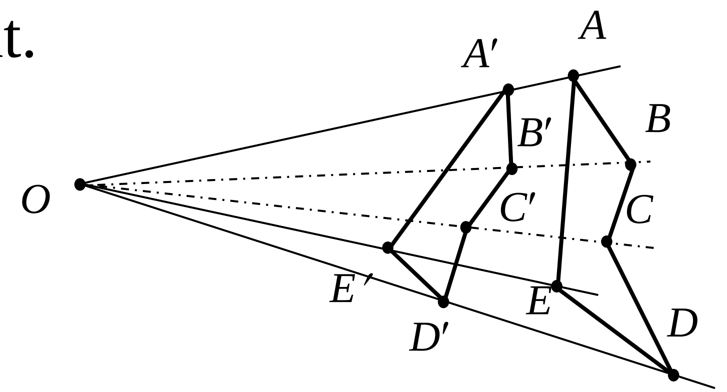
$$R(AC, BC, AD, BD) = \frac{AC / BC}{AD / BD} = \frac{A'C' / B'C'}{A'D' / B'D'}$$

Cross-ration of projections of segments *AC*, *BC*, *AD*, *BD* on image, which is: $(A'C'/B'C')/(A'D'/B'D')=R(AC, BC, AD, BD)$ keeps constant and independent of camera's pose.



Constraints for Visual Measurement

- Cross-ratio of polygon keeps constant.



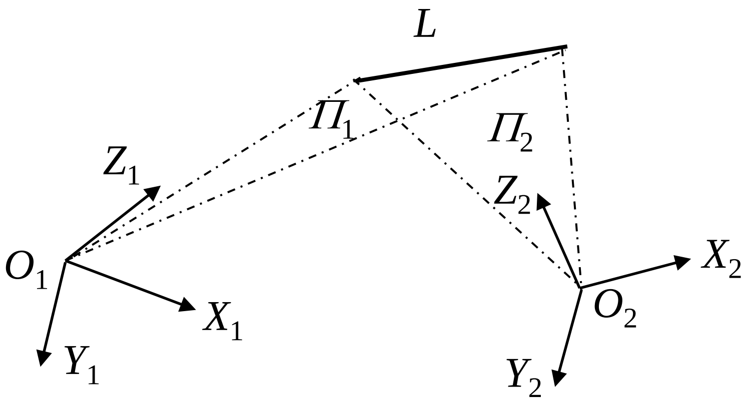
$$R(AB, AC, AD, AE) = \frac{\sin\angle BAD / \sin\angle CAD}{\sin\angle BAE / \sin\angle CAE}$$

$$= \frac{\sin\angle B'A'D' / \sin\angle C'A'D'}{\sin\angle B'A'E' / \sin\angle C'A'E'}$$

We can use this property to identify polygons.

- Line Projection Constraint

Without considering image distortion, the projection of a line in 3D space on the image is still a line, which can be verified by the epipolar geometry.



Monocular Vision based Visual Measurement

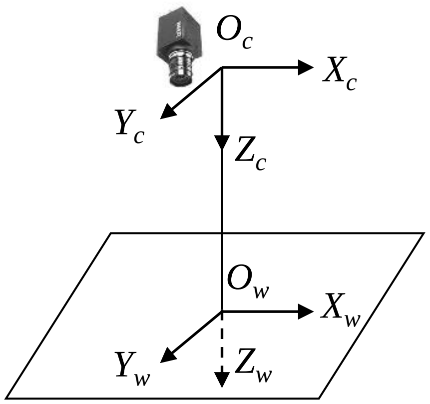
- **Camera's Pose of Monocular Vision System**
  - Optical axis is orthogonal to working plane, and object coordinate is described as:

$$[x_w \quad y_w \quad 0]^T$$

  - Camera's pose is fixed with fixed focal length. The world coordinates (object coordinate) is established accordingly so that $R=I$, $p=[0\ 0\ d]^T$, d is the distance from camera optical center to working plane。

- **Camera Calibration (Extrinsic)**

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_w \\ y_w \\ d \\ 1 \end{bmatrix}$$

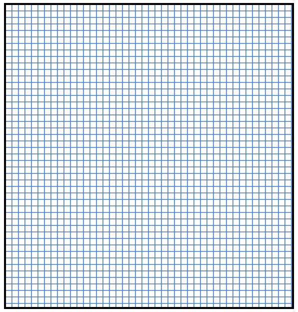


2D Monocular Visual Measurement System

Monocular Vision based Visual Measurement

- Camera Calibration (Intrinsic)

$$z_c\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \Rightarrow \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_0 \\ 0 & k_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{x_w}{d} \\ \frac{y_w}{d} \\ 1 \end{bmatrix}$$



$$\begin{cases} u - u_0 = \frac{k_x}{d} x_w \\ v - v_0 = \frac{k_y}{d} y_w \end{cases} \Rightarrow \begin{cases} u_2 - u_1 = \frac{k_x}{d}(x_{w2} - x_{w1}) \\ v_2 - v_1 = \frac{k_y}{d}(y_{w2} - y_{w1}) \end{cases} \Rightarrow \begin{cases} \frac{k_x}{d} = \frac{u_2 - u_1}{x_{w2} - x_{w1}} \\ \frac{k_y}{d} = \frac{v_2 - v_1}{y_{w2} - y_{w1}} \end{cases}$$

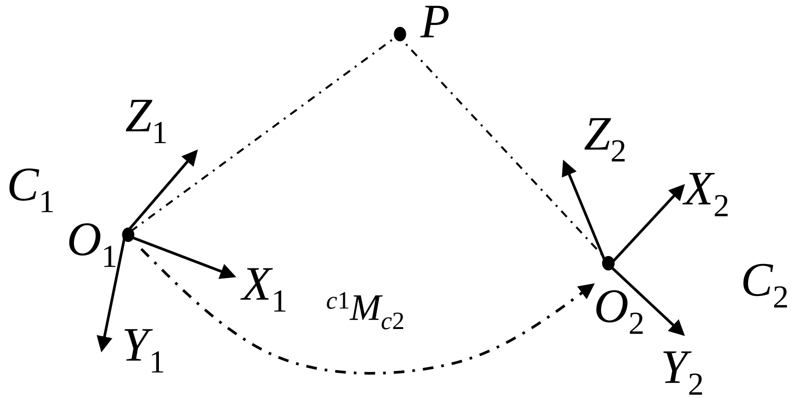
- 2D Measurement

$$\begin{cases} x_{w2} - x_{w1} = \frac{d}{k_x}(u_2 - u_1) \\ y_{w2} - y_{w1} = \frac{d}{k_y}(v_2 - v_1) \end{cases} \Rightarrow \begin{cases} x_{w2} - x_{w1} = k_{dx}(u_2 - u_1) \\ y_{w2} - y_{w1} = k_{dy}(v_2 - v_1) \end{cases}$$

Position Measurement based on Stereo Vision

- Binocular Vision System

(1) With calibrated intrinsic parameters of two cameras, we can get P_{1c1} and P_{1c2}. (2) With P_{1c1} and P_{1c2}, we get two lines O_1P_{1c1} and O_2P_{1c2}. represented in $\{C_1\}$ and $\{C_2\}$ (3) With Essential Matrix, we can represent *the two lines* both in $\{C_1\}$. (3) We then can get the position the ***P***.

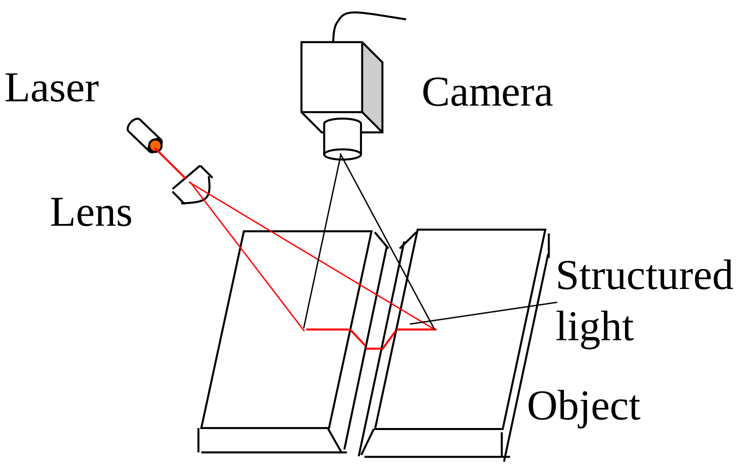


Binocular Stereo Vision System

- Structured Light Vision

It follows triangulation principles. The core is the light is structured; the camera is well calibrated; the patten of the structure light in image plane is calibrated. And we can have different structured patterns, like line-structured pattern or lattice structured pattern. For line light, the 3D position is calculated as:

$$\begin{cases} x = \dfrac{-x_{1c1}}{ax_{1c1}+by_{1c1}+c} \\ y = \dfrac{-y_{1c1}}{ax_{1c1}+by_{1c1}+c} \\ z = \dfrac{-1}{ax_{1c1}+by_{1c1}+c} \end{cases}$$



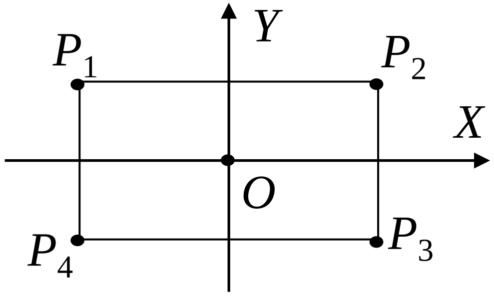
Line-structured light vision system

Pose Measurement with Rectangular Constraint

- Pose measurement of rectangle based on binocular vision system

With binocular vision system, coordinates of the four corners of a rectangle P_1~P_4 in 3D space can be extracted. Using the geometrical center (centroid) of P_1~P_4 as origin, establish the world coordinates.

$$p = \frac{1}{4}\sum_{i=1}^{4} P_i = \frac{1}{4}\sum_{i=1}^{4}\begin{bmatrix} x_{wi} & y_{wi} & z_{wi} \end{bmatrix}^T$$



Rectangle and Coordinates

The X-axis is along the line P_1P_2, while the Y-axis is along the line P_3P_2.

$$\begin{cases} n' = \frac{1}{2}(P_2 + P_3) - \frac{1}{2}(P_1 + P_4) = \frac{1}{2}\begin{bmatrix} x_{w2} + x_{w3} - x_{w1} - x_{w4} \\ y_{w2} + y_{w3} - y_{w1} - y_{w4} \\ z_{w2} + z_{w3} - z_{w1} - z_{w4} \end{bmatrix}, \; n = \frac{n'}{\|n'\|} \\ o' = \frac{1}{2}(P_1 + P_2) - \frac{1}{2}(P_3 + P_4) = \frac{1}{2}\begin{bmatrix} x_{w1} + x_{w2} - x_{w3} - x_{w4} \\ y_{w1} + y_{w2} - y_{w3} - y_{w4} \\ z_{w1} + z_{w2} - z_{w3} - z_{w4} \end{bmatrix}, \; o = \frac{o'}{\|o'\|} \end{cases}$$

This approach works in case of corner points being available on image.

Pose Measurement based on PnP Problem

- Definition of PnP Problem

➢ Perspective-n-Point

Given a set of *n* 3D points in a world reference frame and their corresponding 2D image projections as well as the calibrated intrinsic camera parameters, determine the 6 DOF pose of the camera in the form of its rotation and translation with respect to the world. This follows the perspective project model for cameras:

$$s\,p_c = K\,[\,R\,|\,T\,]\,p_w.$$

Where

$p_w = [\,x \quad y \quad z \quad 1\,]^T$ is the homogeneous world point

$p_c = [\,u \quad v \quad 1\,]^T$ is the corresponding homogeneous image point

K is the matrix of Intrinsic Camera Parameters
R and ***T*** are Extrinsic Camera Parameters.

Pose Measurement based on PnP Problem

- Definition of PnP Problem

➢ Perspective-n-Point

- In PnP problem, a basic assumption is the camara is already calibrated.
- In PnP problem, the chosen point correspondences can not be colinear.
- PnP Problem has multiple solutions; choosing a particular solution would require post-processing of the solution set.
- RANSAC is commonly used with a PnP method to make the solution robust to outliers in the set of point correspondences.
- P3P methods assume that the data is noise free, most PnP methods assume Gaussian noise on the inlier set.
- When n = 3, the PnP problem is in its minimal form of P3P and can be solved with three point correspondences. However, with just three point correspondences, P3P yields up to four real, geometrically feasible solutions.
- The oldest published solution dates to 1841. A recent algorithm for solving the problem as well as a solution classification for it is given in the 2003 IEEE Transactions on Pattern Analysis and Machine Intelligence paper by Gao, et al. An open source implementation of Gao's P3P solver can be found in OpenCV's calib3d module in the solvePnP function.

Gao, Xiao-Shan; Hou, Xiao-Rong; Tang, Jianliang; Cheng, Hang-Fei (2003). "Complete Solution Classification for the Perspective-Three-Point Problem". IEEE Transactions on Pattern Analysis and Machine Intelligence. ***25*** *(8): 930–943. doi:10.1109/tpami.2003.1217599.*

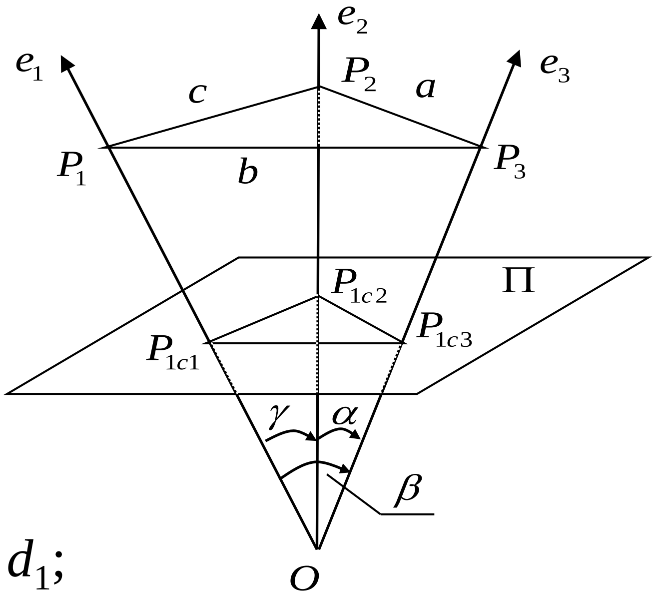
Pose Measurement based on PnP Problem

- Solutions to P3P

Determine Unit Vector e_i:

$$e_i = \frac{1}{\sqrt{x_{1ci}^2 + y_{1ci}^2 + 1}} \begin{bmatrix} x_{1ci} \\ y_{1ci} \\ 1 \end{bmatrix}$$

$$\begin{cases} \cos\alpha = e_2^T e_3 \\ \cos\beta = e_1^T e_3 \\ \cos\gamma = e_1^T e_2 \end{cases}$$



P3P Projection

$$d_2^2 + d_3^2 - 2d_2 d_3 \cos\alpha = a^2$$
$$d_1^2 + d_3^2 - 2d_1 d_3 \cos\beta = b^2$$
$$d_1^2 + d_2^2 - 2d_1 d_2 \cos\gamma = c^2$$

Distance from P_1 to O is d_1;
Distance from P_2 to O is d_2;
Distance from P_3 to O is d_3;

Define: $\begin{cases} d_2 = x d_1 \\ d_3 = y d_1 \end{cases}$

We have:

$$d_1^2 = \frac{a^2}{x^2 + y^2 - 2xy\cos\alpha} \quad (1)$$

$$d_1^2 = \frac{b^2}{1 + y^2 - 2y\cos\beta} \quad (2)$$

$$d_1^2 = \frac{c^2}{x^2 + 1 - 2x\cos\gamma} \quad (3)$$

J. G. Juang, "Parameter Estimation in the Three-Point Perspective Projection Problem in Computer Vision," Proceedings of the IEEE International Symposium on Industrial Electronics, vol. 3, pp. 1065-1070, Jul. 1997.

Pose Measurement based on PnP Problem

- Solutions to P3P

Further Processing yield (Get rid of d_1):

$$x^2 = 2xy\cos\alpha - \frac{2a^2}{b^2}y\cos\beta + \frac{a^2-b^2}{b^2}y^2 + \frac{a^2}{b^2}$$ Combine (1) and (2)

$$x^2 = 2x\cos\gamma - \frac{2c^2}{b^2}y\cos\beta + \frac{c^2}{b^2}y^2 + \frac{c^2-b^2}{b^2}$$ Combine (2) and (3)

Finally:

$$x = \frac{\frac{a^2-b^2-c^2}{b^2}y^2 + 2\cos\beta\frac{a^2-c^2}{b^2}y + \frac{a^2+b^2-c^2}{b^2}}{2(\cos\gamma - y\cos\alpha)}$$

Replace x into equation combining (1) and (3)

$$a_4y^4 + a_3y^3 + a_2y^2 + a_1y + a_0 = 0$$ Solve $y \to x \to d_1 \to d_2 \to d_3$

Where:

$$a_0 = \left(\frac{a^2+b^2-c^2}{b^2}\right)^2 - \frac{4a^2}{b^2}\cos^2\gamma$$

$$a_1 = 4\left[-\left(\frac{a^2+b^2-c^2}{b^2}\right)\left(\frac{a^2-c^2}{b^2}\right)\cos\beta + \frac{2a^2}{b^2}\cos^2\gamma\cos\beta + \left(\frac{a^2-b^2+c^2}{b^2}\right)\cos\alpha\cos\gamma\right]$$

$$a_2 = 2\left[\left(\frac{a^2-c^2}{b^2}\right)^2 - 1 - 4\left(\frac{a^2+c^2}{b^2}\right)\cos\alpha\cos\beta\cos\gamma + 2\left(\frac{b^2-c^2}{b^2}\right)\cos^2\alpha + 2\left(\frac{a^2-c^2}{b^2}\right)^2\cos^2\beta + 2\left(\frac{b^2-a^2}{b^2}\right)\cos^2\gamma\right]$$

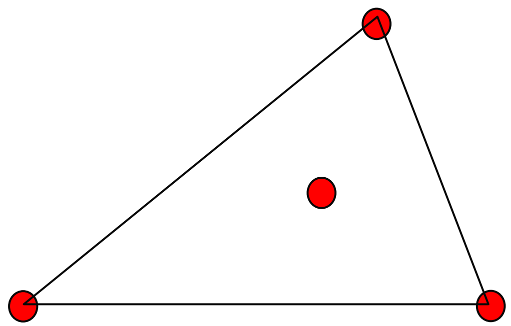
$$a_3 = 4\left[-\left(\frac{a^2-b^2-c^2}{b^2}\right)\left(\frac{a^2-c^2}{b^2}\right)\cos\beta + \frac{2c^2}{b^2}\cos^2\alpha\cos\beta + \left(\frac{a^2-b^2+c^2}{b^2}\right)\cos\alpha\cos\gamma\right]$$

$$a_4 = \left(\frac{a^2-b^2-c^2}{b^2}\right)^2 - \frac{4c^2}{b^2}\cos^2\alpha$$

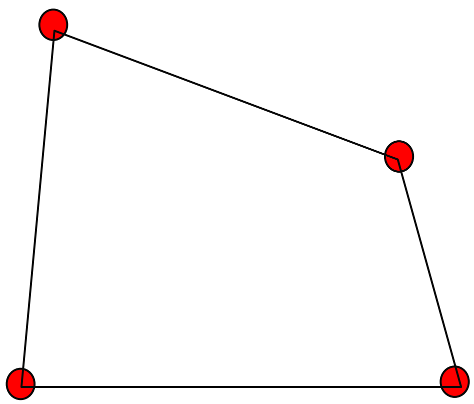
Conclusion: P3P has four solutions.

Pose Measurement based on PnP Problem

- Unique solution pattern design
- For four coplanar points, in which any three points are non-collinear, the solution for the PnP problem is unique.
- More given points can improve positioning accuracy
- Planar Patterns



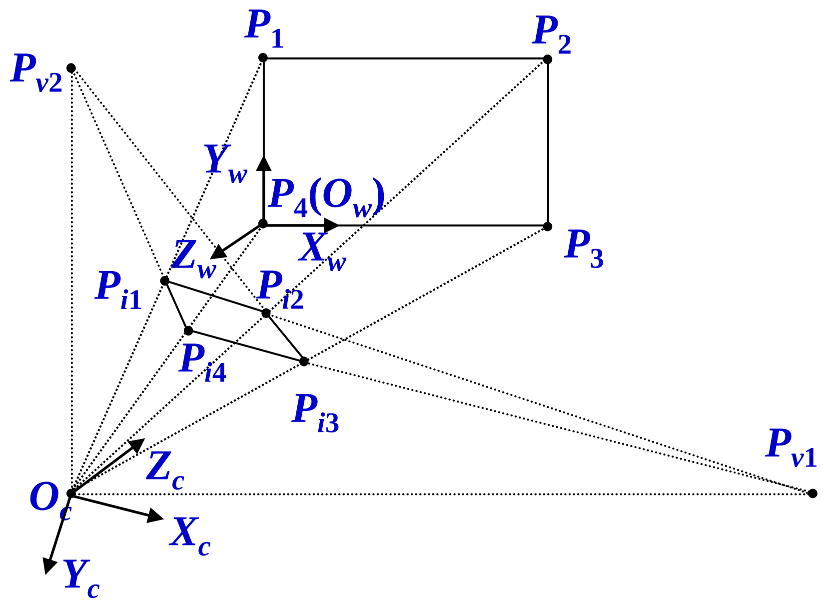
Triangle pattern



Quadrangle pattern

Pose Measurement base on Vanishing Point

- A vanishing point is a point on the image plane of a perspective drawing where the two-dimensional perspective projections (or drawings) of mutually parallel lines in three-dimensional space appear to converge.



$\boldsymbol{P}_{v1}$ and $\boldsymbol{P}_{v2}$ are vanishing points

O_cP_{v1} // P_1P_2 // P_4P_3 or X-axis;

O_cP_{v2} // P_1P_4 // P_2P_3 or Y-axis;

$$ {}^{c}n_w = \frac{1}{\sqrt{x_{1c1}^2 + y_{1c1}^2 + 1}} \begin{bmatrix} x_{1c1} \\ y_{1c1} \\ 1 \end{bmatrix} \qquad {}^{c}o_w = \frac{1}{\sqrt{x_{1c2}^2 + y_{1c2}^2 + 1}} \begin{bmatrix} x_{1c2} \\ y_{1c2} \\ 1 \end{bmatrix} \qquad {}^{c}a_w = {}^{c}n_w \times {}^{c}o_w $$

E. Guillou, D. Meneveaux, E. Maisel, K. Bouatouch, Using vanishing points for camera calibration and coarse 3D reconstruction from a single image, Visual Computer, v 16, n 7, p 396-410, 2000.

Conclusion

- Using monocular vision system for pose measurement, basically we need some pre-setting for the working environment, or we need priori knowledge regarding the working environment, otherwise it does not work at all.
- With binocular with system for pose measurement, we need stereo pair matching.