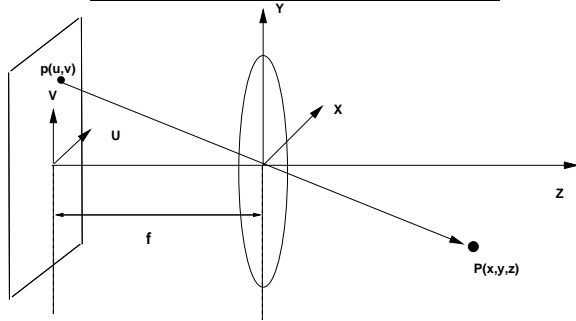


The Pinhole Camera Model



Based on simply trigonometry (or using 3D line equations), we can derive

$$u = \frac{-fx}{z} \quad v = \frac{-fy}{z}$$

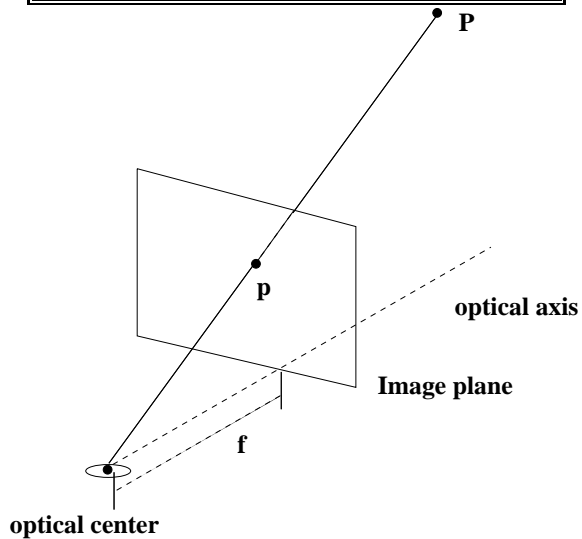
Slide 1

$$u = \frac{fx}{z} \quad v = \frac{fy}{z}$$

where $\frac{f}{z}$ is referred to as isotropic scaling. The full perspective projection is non-linear.

Slide 3

The Computer Vision Camera Model



Slide 2

Weak Perspective Projection

If the relative distance δz (scene depth) between two points of a 3D object along the optical axis is much smaller than the average distance \bar{z} ($\delta z < \frac{\bar{z}}{20}$),

then

$$u = f \frac{x}{z} \approx \frac{fx}{\bar{z}}$$

$$v = f \frac{y}{z} \approx \frac{fy}{\bar{z}}$$

We have linear equations since all projections have the same scaling factor.

Slide 4

Orthographic Projection

As a special case of the weak perspective projection, when $\frac{f}{z}$ factor equals 1, we have $u = x$ and $v = y$, i.e., the lines (rays) of projection are parallel to the optical axis. This leads to the sizes of image and the object are the same. This is called orthographic projection.

Notations

Let $P = (x \ y \ z)^t$ be a 3D point in object frame and $U = (u \ v)^t$ the corresponding image point in the image frame before digitization. Let $X_c = (x_c \ y_c \ z_c)^t$ be the coordinates of P in the camera frame and $p = (c \ r)^t$ be the coordinates of U in the row-column frame after digitization.

Perspective projection geometry

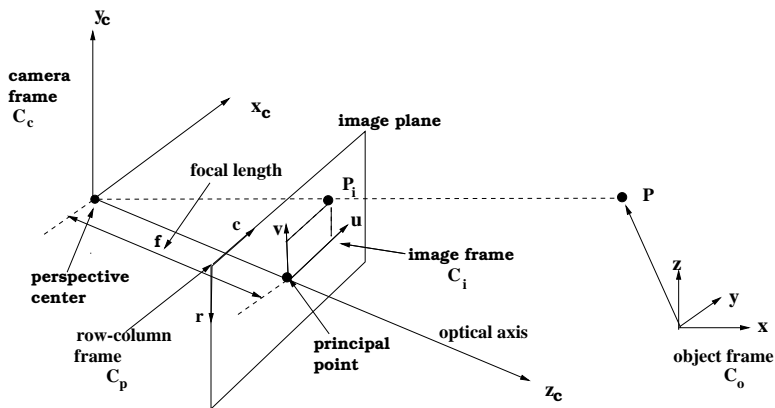


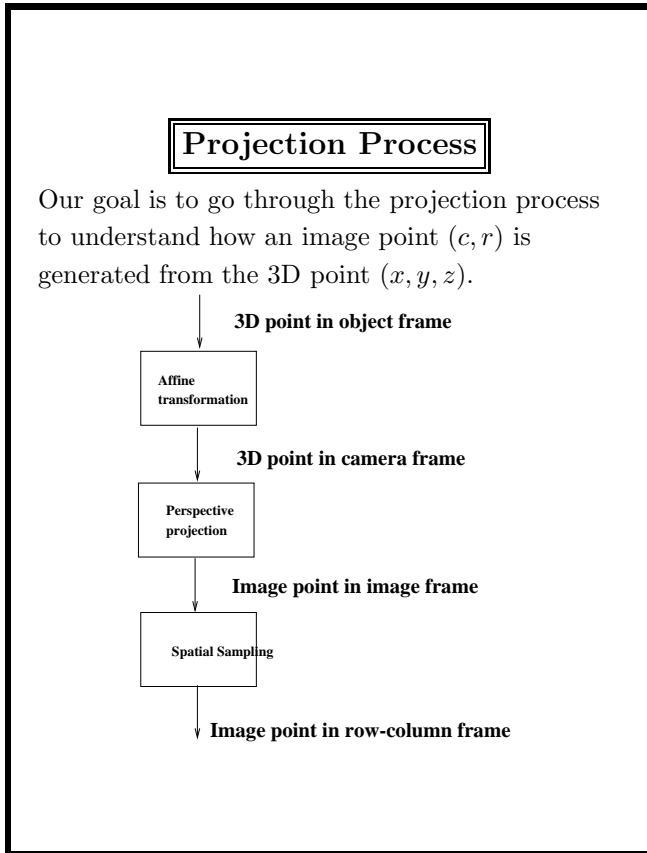
Figure 1: Perspective projection geometry

Slide 5

Slide 7

Slide 6

Slide 8



Slide 9

Relationships between different frames

Between camera frame (C_c) and object frame (C_o)

$$X_c = RX + T \tag{1}$$

X is the 3D coordinates of P w.r.t the object frame. R is the rotation matrix and T is the translation vector. R and T specify the orientation and position of the object frame relative to the camera frame.

Slide 10

R and T can be parameterized as

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \quad T = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

Slide 11

Substituting the parameterized T and R into equation 1 yields

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (2)$$

Slide 13

Relationships between different frames (cont'd)

- Between image frame (C_i) and row-col frame (C_p) (spatial quantization process)

$$\begin{pmatrix} c \\ r \end{pmatrix} = \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} c_0 \\ r_0 \end{pmatrix} \quad (4)$$

where s_x and s_y are scale factors (*pixels/mm*) due to spatial quantization. c_0 and r_0 are the coordinates of the principal point in pixels relative to C_p

Slide 12

- Between image frame (C_i) and camera frame (C_c)

Perspective Projection:

$$u = \frac{fx_c}{z_c}$$

$$v = \frac{fy_c}{z_c}$$

Hence,

$$X_c = \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \\ f \end{pmatrix} \quad (3)$$

where $\lambda = \frac{z_c}{f}$ is a scalar and f is the camera focal length.

Slide 14

Collinearity Equations

Combining equations 1 to 4 yields

$$c = s_x f \frac{r_{11}x + r_{12}y + r_{13}z + t_x}{r_{31}x + r_{32}y + r_{33}z + t_z} + c_0$$

$$r = s_y f \frac{r_{21}x + r_{22}y + r_{23}z + t_y}{r_{31}x + r_{32}y + r_{33}z + t_z} + r_0$$

Homogeneous system: perspective projection

In homogeneous coordinate system, equation 3 may be rewritten as

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} \quad (5)$$

Note $\lambda = z_c$.

Slide 15

Homogeneous system: quantization + projection

Substituting equation 5 into equation 6 yields

$$\lambda \begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = \begin{pmatrix} s_x f & 0 & c_0 & 0 \\ 0 & s_y f & r_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} \quad (7)$$

where $\lambda = z_c$.

Slide 17

Homogeneous System: Spatial Quantization

Similarly, in homogeneous system, equation 4 may be rewritten as

$$\begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = \begin{pmatrix} s_x & 0 & c_0 \\ 0 & s_y & r_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad (6)$$

Slide 16

Homogeneous system: Affine Transformation

In homogeneous coordinate system, equation 2 can be expressed as

$$\begin{pmatrix} x_c \\ y_c \\ z_c \\ 1 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (8)$$

Slide 18

Homogeneous system: full perspective

Combining equation 8 with equation 7 yields

$$\lambda \begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = \underbrace{\begin{pmatrix} s_x f r_1 + c_0 r_3 & s_x f t_x + c_0 t_z \\ s_y f r_2 + r_0 r_3 & s_y f t_y + r_0 t_z \\ r_3 & t_z \end{pmatrix}}_P \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (9)$$

where $r_1, r_2,$ and r_3 are the row vectors of the rotation matrix R , $\lambda = z_c$ is a scalar and matrix P is called the homogeneous projection matrix.

$$P = WM$$

Slide 19

Weak Perspective Camera Model

For weak perspective projection, we have $z_c \approx \bar{z}_c$, i.e., $\bar{z}_c \approx r_3^t \bar{X} + t_z$ Hence,

$$u = \frac{f x_c}{\bar{z}_c}$$

$$v = \frac{f y_c}{\bar{z}_c}$$

Hence,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{f}{\bar{z}_c} \begin{pmatrix} x_c \\ y_c \end{pmatrix}$$

Or

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{f}{\bar{z}_c} \begin{pmatrix} x_c \\ y_c \\ \frac{\bar{z}_c}{f} \end{pmatrix}$$

Slide 21

where

$$W = \begin{pmatrix} f s_x & 0 & c_0 \\ 0 & f s_y & r_0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$M = \begin{pmatrix} R & T \end{pmatrix}$$

W is often referred to as the intrinsic matrix and M as exterior matrix.

Since $P = WM = [WR \quad WT]$, for P to be a projection matrix, $Det(WR) \neq 0$, i.e., $Det(W) \neq 0$.

Slide 20

Since

$$\begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = \begin{pmatrix} s_x & 0 & c_0 \\ 0 & s_y & r_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

We have

$$\begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = \frac{f}{\bar{z}_c} \begin{pmatrix} s_x & 0 & c_0 \\ 0 & s_y & r_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_c \\ y_c \\ \frac{\bar{z}_c}{f} \end{pmatrix}$$

Since

$$\begin{pmatrix} x_c \\ y_c \end{pmatrix} = [R_2 \quad T_2] \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

where R_2 is the first two rows of R and T_2 is the first two elements

Slide 22

of T. Or

$$\begin{pmatrix} x_c \\ y_c \\ \frac{\bar{z}_c}{f} \end{pmatrix} = \begin{pmatrix} R_2 & & T_2 \\ 0 & 0 & 0 \\ & & \frac{\bar{z}_c}{f} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Hence,

$$\begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = \frac{f}{\bar{z}_c} \begin{pmatrix} s_x & 0 & c_0 \\ 0 & s_y & r_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R_2 & & T_2 \\ 0 & 0 & 0 \\ & & \frac{\bar{z}_c}{f} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Slide 23

Weak Perspective Camera Model

The weak perspective projection matrix is

$$P_{weak} = \begin{pmatrix} f s_x r_1 & f s_x t_x + c_0 \bar{z}_c \\ f s_y r_2 & f s_y t_y + r_0 \bar{z}_c \\ 0^{1 \times 3} & \bar{z}_c \end{pmatrix} \quad (10)$$

where r_1 and r_2 are the first two rows of R_2 and $\bar{z}_c = r_3 \bar{X} + t_z$.

Slide 25

$$= \frac{1}{\bar{z}_c} \begin{pmatrix} s_x & 0 & c_0 \\ 0 & s_y & r_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f R_2 & & f T_2 \\ 0 & 0 & 0 \\ & & \bar{z}_c \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Slide 24

Affine Camera Model

A further simplification from weak perspective camera model is the affine camera model, which is often assumed by computer vision researchers due to its simplicity. The affine camera model assumes that the object frame is located on the centroid of the object being observed. As a result, we have $\bar{z}_c \approx t_z$, the affine perspective projection matrix is

$$P_{affine} = \begin{pmatrix} s_x f r_1 & s_x f t_x + c_0 t_z \\ s_y f r_2 & s_y f t_y + r_0 t_z \\ 0 & t_z \end{pmatrix} \quad (11)$$

Affine camera model represents the first order approximation of the full perspective projection camera model. It still only gives an approximation and is no longer useful when the object is close to

Slide 26

Slide 27

the camera or the camera has a wide angle of view.

Slide 29

$$v = y_c$$

which can be approximated by $\frac{f}{z_c} \approx 1$.

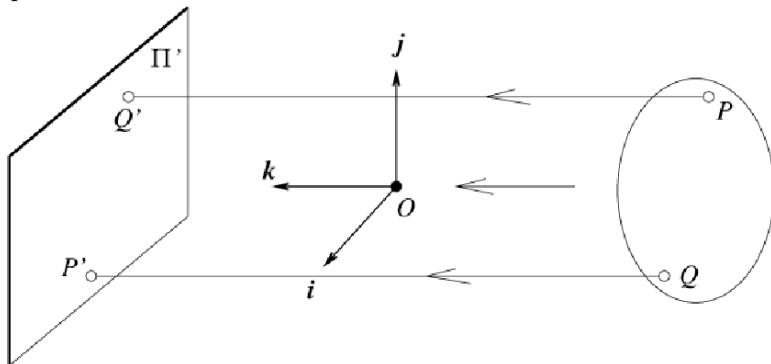
The orthographic projection matrix can therefore be obtained as

$$P_{orth} = \begin{pmatrix} s_x r_1 & s_x t_x + c_0 \\ s_y r_2 & s_y t_y + r_0 \\ 0 & 1 \end{pmatrix} \quad (12)$$

Slide 28

Orthographic Projection Camera Model

Under orthographic projection, projection is parallel to the camera optical axis.



therefore we have

$$u = x_c$$

Slide 30

Non-perspective Projection Camera Model

The weak perspective projection, affine, and orthographic camera model can be collectively classified as *non-perspective projection* camera model. In general, the projection matrix for the non-perspective projection camera model

$$\lambda \begin{pmatrix} c \\ r \\ 1 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & p_{34} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Dividing both sides by p_{34} (note $\lambda = p_{34}$) yields

$$\begin{pmatrix} c \\ r \end{pmatrix} = M_{2 \times 3} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

where $m_{ij} = p_{ij}/p_{34}$ and $v_x = p_{14}/p_{34}$, $v_y = p_{24}/p_{34}$

For any given reference point (c_r, r_r) in image and (x_0, y_0, z_0) in space, the relative coordinates (\bar{c}, \bar{r}) in image and $(\bar{x}, \bar{y}, \bar{z})$ in space are

$$\begin{pmatrix} \bar{c} \\ \bar{r} \end{pmatrix} = \begin{pmatrix} c - c_r \\ r - r_r \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix} = \begin{pmatrix} x - x_r \\ y - y_r \\ z - z_r \end{pmatrix}$$

It follows that the basic projection equation for the affine and weak perspective model in terms of relative coordinates is

The M matrix is

$$\begin{aligned} M &= \begin{pmatrix} \frac{f s_x r_1}{\bar{z}_c} \\ \frac{f s_y r_2}{\bar{z}_c} \end{pmatrix} \\ &= \frac{f}{\bar{z}_c} \begin{pmatrix} s_x r_1 \\ s_y r_2 \end{pmatrix} \\ &= \frac{f}{\bar{z}_c} \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \end{aligned}$$

For affine projection, $\bar{z}_c = t_z$, for orthographic projection, $\frac{f}{\bar{z}_c} = 1$.
If we assume $s_x = s_y$, then

$$M = \frac{f s_x}{\bar{s}} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

Then, we have only four parameters: three rotation angles and a

Slide 31

Slide 33

$$\begin{pmatrix} \bar{c} \\ \bar{r} \end{pmatrix} = M_{2 \times 3} \begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix}$$

An non-perspective projection camera $M_{2 \times 3}$ has 3 independent parameters. The reference point is often chosen as the centroid since centroid is preserved under either affine or weak perspective projection.

Given the weak projection matrix P ,

$$P = \begin{pmatrix} f s_x r_1 & f s_x t_x + c_0 \bar{z}_c \\ f s_y r_2 & f s_y t_y + r_0 \bar{z}_c \\ 0 & \bar{z}_c \end{pmatrix}$$

Slide 32

Slide 34

scale factor.

Rotation Matrix Representation: Euler angles

Assume rotation matrix R results from successive Euler rotations of the camera frame around its X axis by ω , its once rotated Y axis by ϕ , and its twice rotated Z axis by κ , then

$$R(\omega, \phi, \kappa) = R_X(\omega)R_Y(\phi)R_Z(\kappa)$$

where ω , ϕ , and κ are often referred to as pan, tilt, and swing angles respectively.

Slide 35

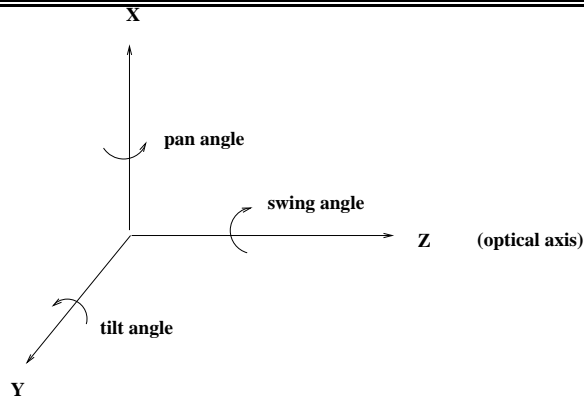
$$R_x(\omega) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \omega & \sin \omega \\ 0 & -\sin \omega & \cos \omega \end{pmatrix}$$

$$R_y(\phi) = \begin{pmatrix} \cos \phi & 0 & -\sin \phi \\ 0 & 1 & 0 \\ \sin \phi & 0 & \cos \phi \end{pmatrix}$$

$$R_z(\kappa) = \begin{pmatrix} \cos \kappa & \sin \kappa & 0 \\ -\sin \kappa & \cos \kappa & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Slide 37

Rotation Matrix Representation: Euler angles



Slide 36

Quaternion Representation of R

The relationship between a quaternion $q = [q_0, q_1, q_2, q_3]$ and the equivalent rotation matrix is

$$R = \begin{pmatrix} q_0q_0 + q_1q_1 - q_2q_2 - q_3q_3 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_2q_1 + q_0q_3) & q_0q_0 - q_1q_1 + q_2q_2 - q_3q_3 & 2(q_2q_3 - q_0q_1) \\ 2(q_3q_1 - q_0q_2) & 2(q_3q_2 + q_0q_1) & q_0q_0 - q_1q_1 - q_2q_2 + q_3q_3 \end{pmatrix}$$

Here the quaternion is assumed to have been scaled to unit length, i.e., $|q| = 1$.

The axis/angle representation \mathbf{a}/θ is strongly related to a quaternion, according to the formula

$$\begin{pmatrix} \cos(\theta/2) \\ a_x \sin(\theta/2) \\ a_y \sin(\theta/2) \\ a_z \sin(\theta/2) \end{pmatrix}$$

Slide 38

Slide 39

where $\mathbf{a} = (\mathbf{a}_x, \mathbf{a}_y, \mathbf{a}_z)$ and $|\mathbf{a}| = 1$.

Slide 41

Interior Camera Parameters

Parameters (c_0, r_0) , s_x , s_y , and f are collectively referred to as *interior camera parameters*. They do not depend on the position and orientation of the camera. Interior camera parameters allow us to perform metric measurements, i.e., to convert pixel measurements to inch or mm.

Slide 40

R's Orthnormality

The rotation matrix is an orthnormal matrix, which means its rows (columns) are normalized to one and they are orthogonal to each other. The orthnormality property produces

$$R^t = R^{-1}$$

Slide 42

Exterior Camera Parameters

Parameters like Euler angles ω , ϕ , κ , t_x , t_y , and t_z are collectively referred to as *exterior camera parameters*. They determine the position and orientation of the camera.

Camera Calibration and Pose Estimation

The purpose of camera calibration is to determine intrinsic camera parameters: c_0, r_0, s_x, s_y , and f . Camera calibration is also referred to as interior orientation problem in photogrammetry.

The goal of pose estimation is to determine exterior camera parameters: $\omega, \phi, \kappa, t_x, t_y$, and t_z . In other words, pose estimation is to determine the position and orientation of the object coordinate frame relative to the camera coordinate frame.

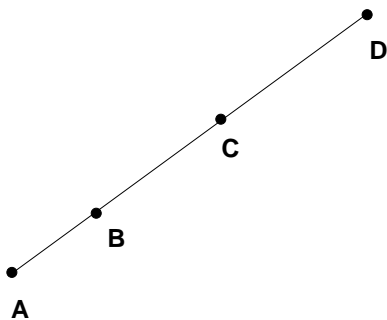
Slide 43

Slide 45

Cross-ratio is preserved under perspective projection.

Perspective Projection Invariants

Distances and angles are invariant with respect to Euclidian transformation (rotation and translation). The most important invariant with respect to perspective projection is called *cross ratio*. It is defined as follows:



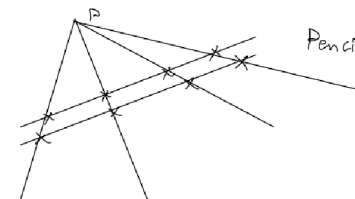
$$\tau(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}) = \frac{AC}{BC} / \frac{AD}{BD}$$

Slide 44

Slide 46

Projective Invariant for non-collinear points

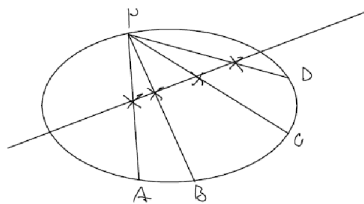
Cross ratio for a set of pencil of 4 lines are only function of the angles among the lines, independent of the points on the lines. The use of cross-ratio for grand plane detection from multiple image frames.



Chasles' theorem:

Let A, B, C, D be distinct points on a (non-singular) conic (ellipse, circle, ..). If P is another point on the conic then the cross-ratio of the pencil PA, PB, PC, PD does not depend on the point P.

This theorem may be used for ellipse detection.



See section 19.3 and 19.4 of Daves book.