Rigid Body Motion and
Image Formation

Jana Kosecka, CS 482

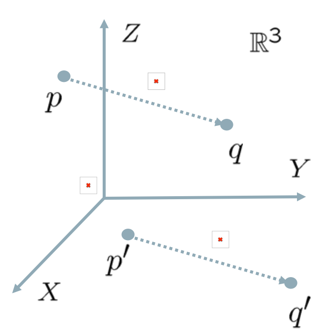
## 3-D Euclidean Space - Vectors

A "free" vector is defined by a pair of points $(p, q)$

$$\boldsymbol{X}_p = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \in \mathbb{R}^3, \ \boldsymbol{X}_q = \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} \in \mathbb{R}^3,$$

Coordinates of the vector :

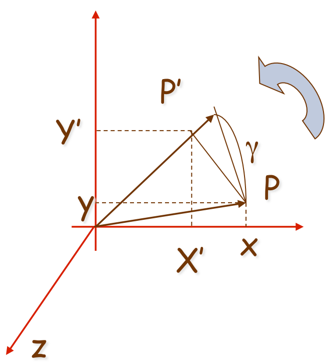
$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} X_2 - X_1 \\ Y_2 - Y_1 \\ Z_2 - Z_1 \end{bmatrix} \in \mathbb{R}^3$$



## 3D Rotation of Points – Euler angles

Rotation around the coordinate axes, counter-clockwise:

$$\begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$$




$$R = R_z(\gamma)R_y(\beta)R_x(\alpha)$$

## Rotation Matrices in 3D

- 3 by 3 matrices
- 9 parameters – only three degrees of freedom
- Representations – either three Euler angles
- or axis and angle representation

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

- Properties of rotation matrices (constraints between the elements)

$$RR^T = I$$

$$det(R) = I$$

## Rotation Matrices in 3D

- 3 by 3 matrices
- 9 parameters – only three degrees of freedom
- Representations – either three Euler angles
- or axis and angle representation

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

- Properties of rotation matrices (constraints between the elements)

$$R.R^T = I \qquad r_i^T r_j = \delta_{ij} \doteq \begin{cases} 1 & \text{for } i = j, \\ 0 & \text{for } i \neq j, \end{cases} \qquad \forall i, j \in \{1, 2, 3\}.$$

$det(R) = I$ Columns are orthonormal

## Canonical Coordinates for Rotation

Property of R $\quad R(t)R^T(t) = I$

Taking derivative

$$\dot{R}(t)R^T(t) + R(t)\dot{R}^T(t) = 0 \quad \Rightarrow \quad \dot{R}(t)R^T(t) = -(\dot{R}(t)R^T(t))^T$$

Skew symmetric matrix property

$$\dot{R}(t)R^T(t) = \widehat{\omega}(t)$$

By algebra

$$\dot{R}(t) = \widehat{\omega}R(t)$$

By solution to ODE

$$R(t) = e^{\widehat{\omega}t}$$

## 3D Rotation (axis & angle)

Solution to the ODE

$$R(t) = e^{\widehat{\omega}t}$$

$$R = I + \widehat{\omega} sin(\theta) + \widehat{\omega}^2(1 - cos(\theta))$$

with $\|\omega\| = 1$ $\omega = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} \in \mathbb{R}^3$

or

$$R = I + \frac{\widehat{\omega}}{\|\omega\|}\sin(\|\omega\|) + \frac{\widehat{\omega}^2}{\|\omega\|^2}(1 - \cos(\|\omega\|))$$

## Rotation Matrices

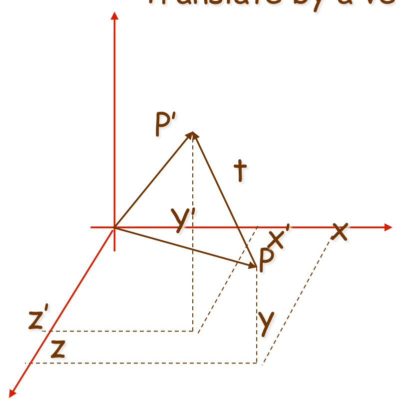
Given $R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix},$

How to compute angle and axis

$$\|\omega\| = \cos^{-1}\left(\frac{\text{trace}(R) - 1}{2}\right), \quad \frac{\omega}{\|\omega\|} = \frac{1}{2\sin(\|\omega\|)}\begin{bmatrix} r_{32} - r_{23} \\ r_{13} - r_{31} \\ r_{21} - r_{12} \end{bmatrix}.$$

## 3D Translation of Points

Translate by a vector $t = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \in \mathbb{R}^3$




$$\mathrm{x}' = x + t = \begin{bmatrix} X + t_x \\ Y + t_y \\ Z + t_z \end{bmatrix}$$

## Rigid Body Motion – Homogeneous Coordinates

3-D coordinates are related by: $\boldsymbol{X}_c = R\boldsymbol{X}_w + T,$

Homogeneous coordinates:

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \rightarrow \quad \boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \in \mathbb{R}^4,$$

Homogeneous coordinates are related by:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

## Rigid Body Motion – Homogeneous Coordinates

3-D coordinates are related by: $\boldsymbol{X}_c = R\boldsymbol{X}_w + T,$

Homogeneous coordinates:

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \to \quad \boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \in \mathbb{R}^4,$$

Homogeneous coordinates are related by:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

## Properties of Rigid Body Motions

Rigid body motion composition

$$\bar{g}_1\bar{g}_2 = \begin{bmatrix} R_1 & T_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} R_2 & T_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} R_1R_2 & R_1T_2 + T_1 \\ 0 & 1 \end{bmatrix} \quad \in SE(3)$$

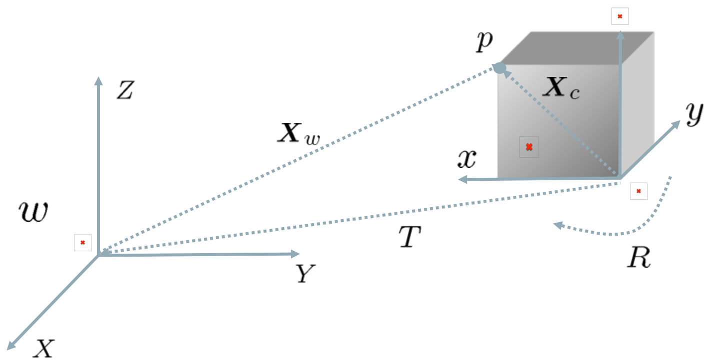
Rigid body motion inverse

$$\bar{g}^{-1} = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} R^T & -R^TT \\ 0 & 1 \end{bmatrix} \quad \in SE(3).$$

Rigid body motion acting on vectors

Vectors are only affected by rotation – 4th homogeneous coordinate is zero

## Rigid Body Transformation




Coordinates are related by: $\boldsymbol{X}_c = R\boldsymbol{X}_w + T,$

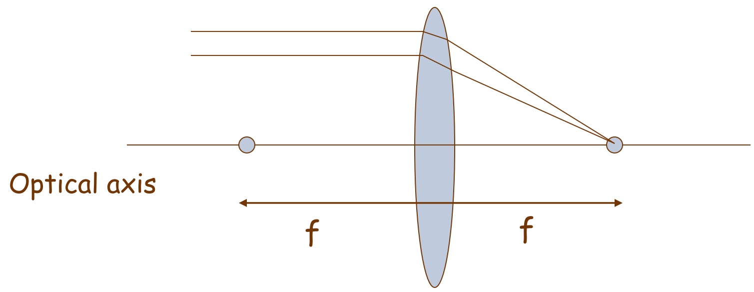
Camera pose is specified by: $g = (R, T) \in SE(3)$

## Image Formation

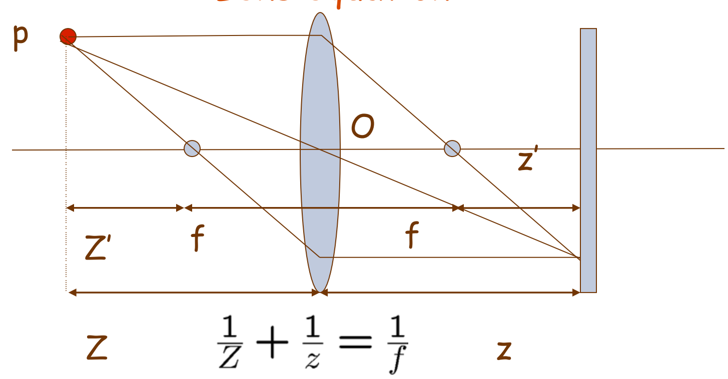
- If the object is our lens the refracted light causes the images
- How to integrate the information from all the rays being reflected from the single point on the surface ?
- Depending in their angle of incidence, some are more refracted then others – refracted rays all meet at the point – basic principles of lenses
- Also light from different surface points may hit the same lens point but they are refracted differently - Kepler's retinal theory

## Thin lens equation

- Idea – all the rays entering the lens parallel to the optical axis on one side, intersect on the other side at the point.




## Lens equation

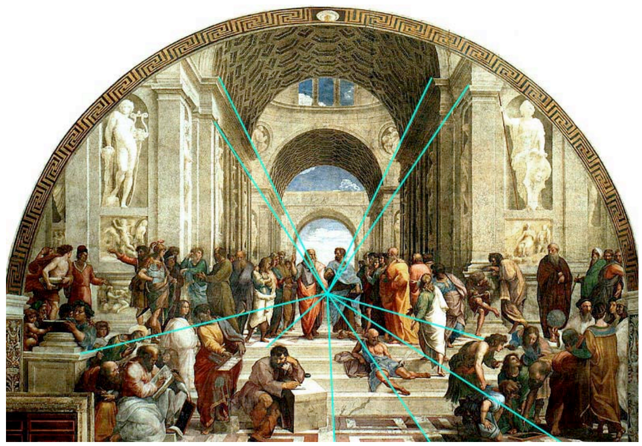



$$\frac{1}{Z}+\frac{1}{z}=\frac{1}{f}$$

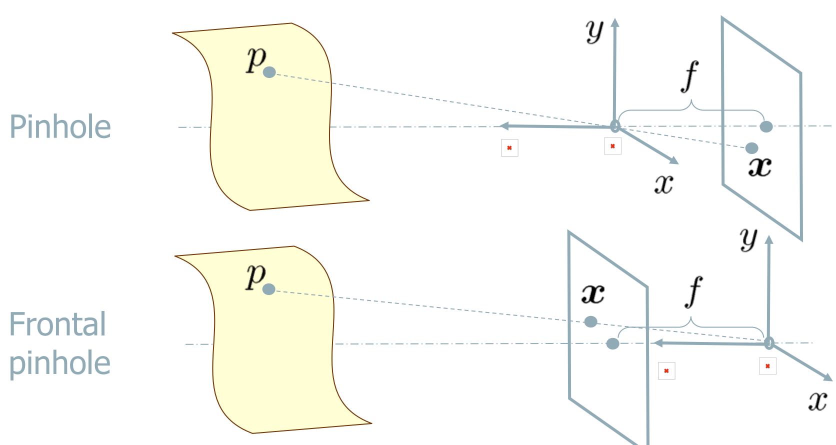
- distance behind the lens at which points becomes in focus depends on the distance of the point from the lens
- in real camera lenses, there is a range of points which are brought into focus at the same distance
- depth of field of the lens , as Z gets large – z' approaches f
- human eye – power of accommodation – changing f

## Image Formation – Perspective Projection

"The School of Athens," Raphael, 1518



## Pinhole Camera Model




$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow \boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

## More on homogeneous coordinates

In homogenous coordinates – these represent the Same point in 3D

$$[X, Y, Z, 1]^T, \quad [XW, YW, ZW, W]^T \quad \in \mathbb{R}^4$$

The first coordinates can be obtained from the second by division by W

What if W is zero ?
Special point – point at infinity – more later

In homogeneous coordinates – there is a difference between point and vector

## Pinhole Camera Model

- Image coordinates are nonlinear function of world coordinates
- Relationship between coordinates in the camera frame and sensor plane

2-D coordinates $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$

Homogeneous coordinates

$$\boldsymbol{x} \to \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}, \qquad \boldsymbol{X} \to \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix},$$

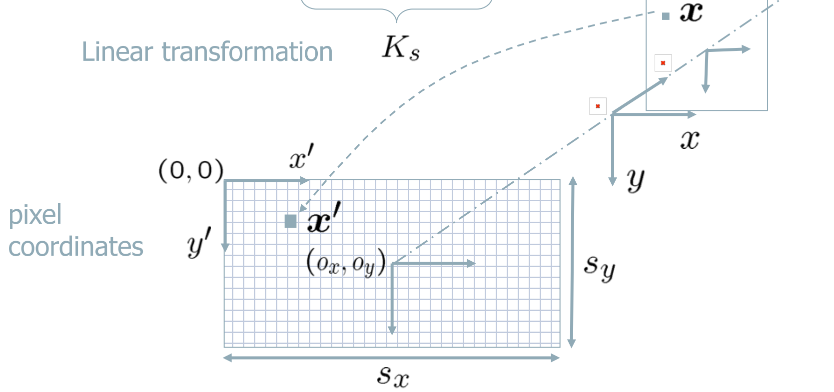
$$Z \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{K_f} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\Pi_0} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

# Image Coordinates

- Relationship between coordinates in the sensor plane and image

$$\boldsymbol{x}' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{K_s} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Linear transformation




# Calibration Matrix and Camera Model

- Relationship between coordinates in the camera frame and image

Pinhole camera

$$\lambda \boldsymbol{x} = K_f \Pi_0 \boldsymbol{X}$$

Pixel coordinates

$$\boldsymbol{x}' = K_s \boldsymbol{x}$$

$$\lambda \boldsymbol{x}' = K_s K_f \Pi_0 \mathbf{X} = \underbrace{\begin{bmatrix} f s_x & f s_\theta & o_x \\ 0 & f s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}\ \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Calibration matrix (intrinsic parameters)

$$\underbrace{K = K_s K_f \qquad \Pi_0}$$

Projection matrix

$$\Pi = [K, 0] \in \mathbb{R}^{3\times 4}$$

Camera model

$$\lambda \boldsymbol{x}' = K \Pi_0 \boldsymbol{X} = \Pi \boldsymbol{X}$$

## Calibration Matrix and Camera Model

- Relationship between coordinates in the world frame and image

Pinhole camera: $\lambda \boldsymbol{x} = K_f \Pi_0 \boldsymbol{X}$

Pixel coordinates: $\boldsymbol{x}' = K_s \boldsymbol{x}$

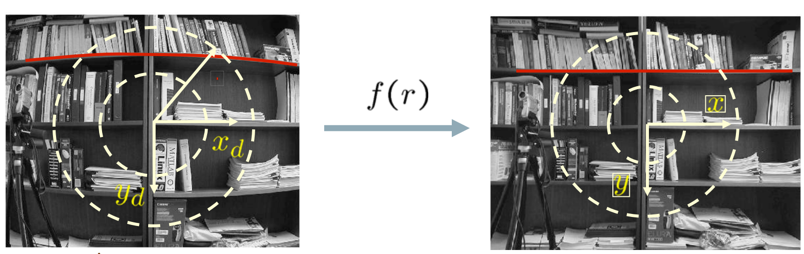
$$\lambda \boldsymbol{x}' = \begin{bmatrix} fs_x & fs_\theta & o_x \\ 0 & fs_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

More compactly $\lambda \boldsymbol{x} = K_f \Pi_0 g \boldsymbol{X} = \Pi \boldsymbol{X}$

Transformation between camera coordinate
Systems and world coordinate system

## Radial Distortion

Nonlinear transformation along the radial direction




New coordinates

$$\begin{aligned} \boldsymbol{x} &= c + f(r)(\boldsymbol{x}_d - c), \quad r = \|\boldsymbol{x}_d - c\| \\ f(r) &= 1 + a_1 r + a_2 r^2 + a_3 r^3 + a_4 r^4 + \cdots \end{aligned}$$

Distortion correction: make lines straight

Coordinates of distorted points

## Image of a point

Homogeneous coordinates of a 3-D point $p$

$$\boldsymbol{X} = [X, Y, Z, W]^T \in \mathbb{R}^4, \quad (W = 1)$$

Homogeneous coordinates of its 2-D image

$$\boldsymbol{x} = [x, y, z]^T \in \mathbb{R}^3, \quad (z = 1)$$

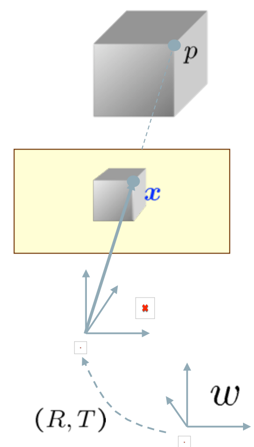
Projection of a 3-D point to an image plane

$$\boxed{\lambda \boldsymbol{x} = \Pi \boldsymbol{X}}$$

$$\lambda \in \mathbb{R}, \ \Pi = [R, T] \in \mathbb{R}^{3\times 4}$$

$$\boxed{\lambda \boldsymbol{x}' = \Pi \boldsymbol{X}}$$

$$\lambda \in \mathbb{R}, \ \Pi = [KR, KT] \in \mathbb{R}^{3\times 4}$$




## Image of a line – homogeneous representation

Homogeneous representation of a 3-D line $L$

$$\boldsymbol{X} = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} X_o \\ Y_o \\ Z_o \\ 1 \end{bmatrix} + \mu \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ 0 \end{bmatrix}, \quad \mu \in \mathbb{R}$$

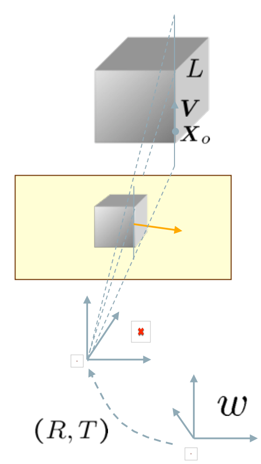
Homogeneous representation of its 2-D image

$$\boldsymbol{l} = [a, b, c]^T \in \mathbb{R}^3$$

Projection of a 3-D line to an image plane

$$\boxed{\boldsymbol{l}^T \boldsymbol{x} = \boldsymbol{l}^T \Pi \boldsymbol{X} = 0}$$

$$\Pi = [KR, KT] \in \mathbb{R}^{3\times 4}$$



## Image of a line – 2D representations

Representation of a 3-D line

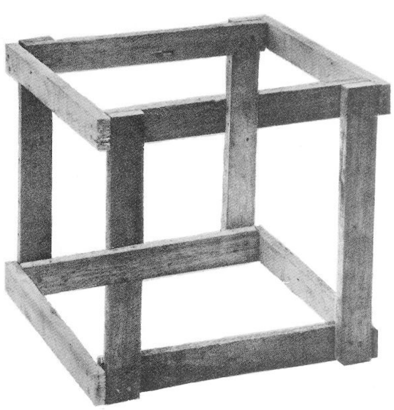
$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} X_o \\ Y_o \\ Z_o \end{bmatrix} + \mu \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix}, \quad \mu \in \mathbb{R}$$

Projection of a line - line in the image plane

$$\begin{aligned} x &= \frac{X_o + \lambda V_1}{Z_o + \lambda V_3} \\ y &= \frac{Y_o + \lambda V_2}{Z_o + \lambda V_3} \end{aligned}$$

Special cases – parallel to the image plane, perpendicular
When $\lambda$ -> infinity - vanishing points
In art – 1-point perspective, 2-point perspective, 3-point perspective

## Visual Illusions, Wrong Perspective

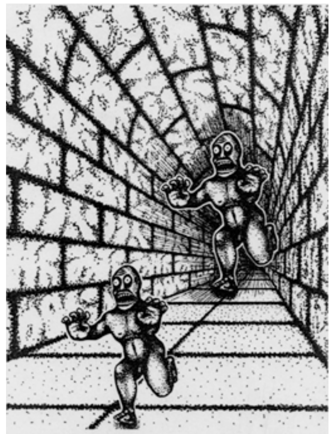


## Vanishing points

Different sets of parallel lines in a plane intersect
at vanishing points, vanishing points form a horizon line

## Ames Room Illusions

## More Illusions



Which of the two monsters is bigger ?