Computer Vision for HCI

Camera Models and Calibration

1

Overview

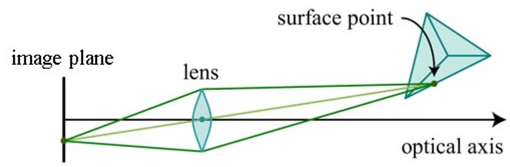
- Thin Lens Model
- Projective Camera Model
  - Intrinsic and Extrinsic Camera Parameters
  - Incorporating Radial Distortion
- Additional Camera Models and Effects

Some slides adapted from M. Pollefeys, K. Grauman, S. Seitz, T. Darrell, L. Sigal, D. Fleet, G. Bebis, and A. Hertzmann

2

Thin Lens Model

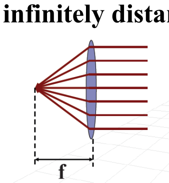
- Lenses are used to focus light onto image plane
  - Enables capturing enough light in sufficiently short duration (as determined by the shutter) so that
    - Objects do not move appreciably during capture
    - Image is bright enough to show detail
- Models of real lenses are <u>complex</u>
  - Approximate with "thin lens" model



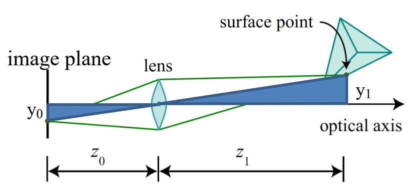
3

"Thin Lens" Model

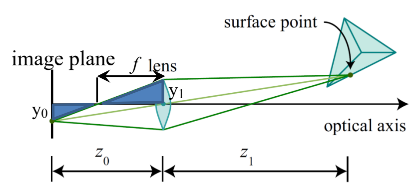
- Focal length f is the distance behind lens where <u>parallel rays converge</u> (i.e., point where rays from **infinitely distant source** converge)



- Can derive thin lens model using rays that are not parallel



$$\frac{y_0}{z_0} = \frac{y_1}{z_1} \rightarrow \frac{y_0}{y_1} = \frac{z_0}{z_1}$$



$$\frac{y_0}{z_0 - f} = \frac{y_1}{f} \rightarrow \frac{y_0}{y_1} = \frac{z_0 - f}{f}$$

4

Thin Lens Model

- Solve for focal length f:

$$\frac{y_0}{y_1} = \frac{z_0}{z_1} \qquad \frac{y_0}{y_1} = \frac{z_0 - f}{f} \longrightarrow \frac{z_0}{z_1} = \frac{z_0 - f}{f}$$

Cross multiply

$$z_0 f = z_0 z_1 - z_1 f \longrightarrow z_0 f + z_1 f = z_0 z_1$$

Solve for f

$$f = \frac{z_0 z_1}{z_0 + z_1} \longrightarrow \frac{1}{f} = \frac{z_0 + z_1}{z_0 z_1} \longrightarrow \boxed{\frac{1}{f} = \frac{1}{z_0} + \frac{1}{z_1}}$$

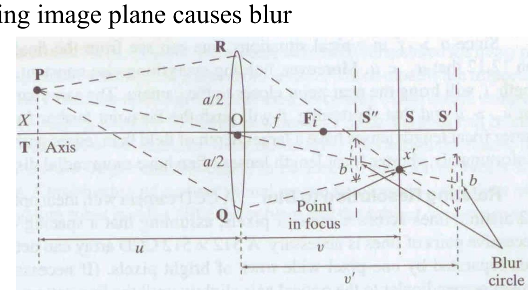
Thin lens equation

<u>Points that satisfy the thin lens equation are "in focus"</u>
(Given f and z_0, the object should be z_1 away)

5

Focus and "Depth of Field"

- For the thin lens model, points from different depths come in focus at different image planes
  - Moving image plane causes blur



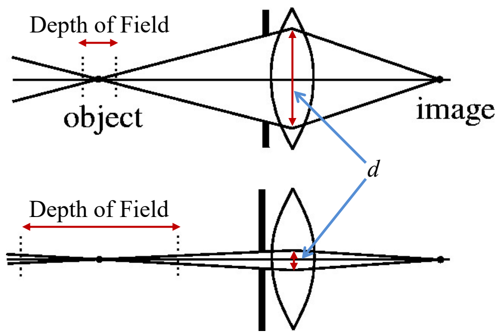
Shapiro and Stockman

- **Depth of field** = the distance between the nearest and farthest points that produce "tolerable" amounts of blur

6

Depth of Field and Lens Aperture Diameter

- Depth of field <u>increases</u> as lens **aperture** diameter *d* <u>decreases</u>



For constant focal length *f*

$f/d = 5.6$

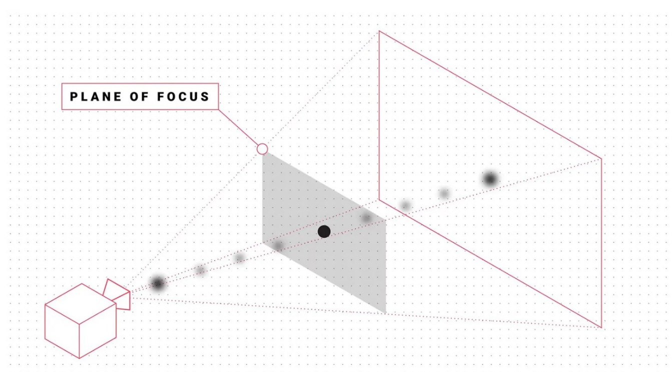
$f/d = 32$

Reduced light transmission and increased diffraction place practical limits on minimum aperture diameter

http://en.wikipedia.org/wiki/Depth_of_field

7

"Portrait" Mode (on phone)

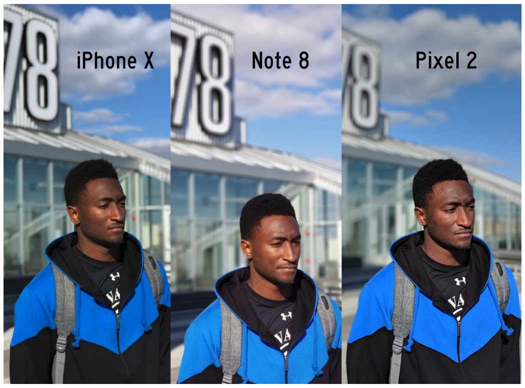


Because a smartphone's sensor is so small and the field of view is so wide, most of the time everything in a normal photo taken with the phone's camera will be in focus. Portrait mode simulates a shallow depth of field by using edge detection and/or depth mapping to differentiate between the foreground and background. It then blurs the background, simulating that shallow depth of field and making the foreground pop.

https://petapixel.com/2017/12/11/portrait-mode-works-compares-8000-camera/

8

“Portrait” Mode (on phone)



The iPhone X and Note 8 use depth mapping to figure out what is in the foreground of the image. These smartphones use data from the wide angle and telephoto lenses to create a depth map, and then artificially blur objects depending on how far they are from the in-focus subject.

The Pixel 2 takes a different approach — it utilizes pixel splitting to create a depth map and machine learning helps to identify the subject and create a mask. Because it doesn’t rely on two distinct lenses like the iPhone X and Note 8, the Pixel 2 is able to take Portrait Mode shots from the front-facing camera.

https://petapixel.com/2017/12/11/portrait-mode-works-compares-8000-camera/

9

Overview

- Thin Lens Model
- Projective Camera Model
  - Intrinsic and Extrinsic Camera Parameters
  - Incorporating Radial Distortion
- Additional Camera Models and Effects

10

Pinhole Camera Model

- Pinhole camera model results from reducing aperture diameter to infinitesimally small point
  - Simple model and all objects in focus
- Rays connect image plane to object through pinhole
  - Object is imaged upside-down on image plane

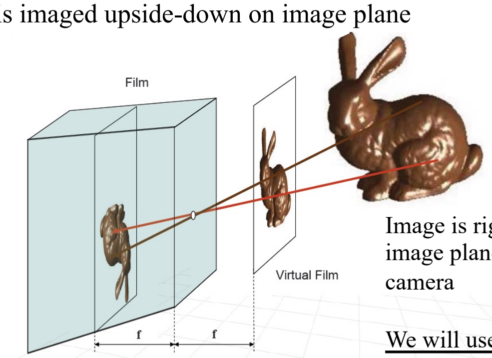
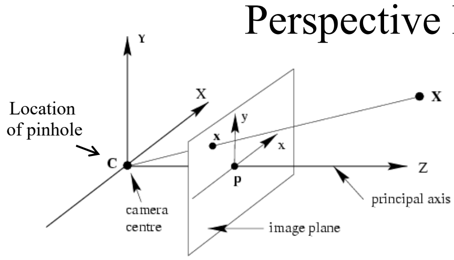


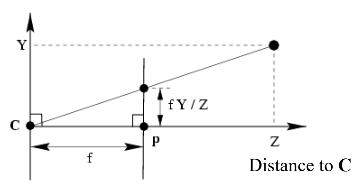
Image is rightside-up if "virtual image plane" is in front of the camera

We will use **this** model

11

Pinhole Camera Model
Perspective Projection





Side View

- Relationship between 3-D point **X** = $(X,Y,Z)^T$ and 2-D point **x** = $(x,y)^T$:

$$\frac{X}{Z} = \frac{x}{f} \rightarrow x = f\frac{X}{Z} \qquad \frac{Y}{Z} = \frac{y}{f} \rightarrow y = f\frac{Y}{Z}$$

- Is it possible to write this nonlinear transformation as a linear transformation???

12

Homogeneous Coordinates

- Add one more coordinate to form **homogeneous** coordinates

Converting *from* inhomogeneous coordinates *to* homogeneous coordinates

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous image coordinates

homogeneous scene coordinates

Converting *from* homogeneous coordinates *to* inhomogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w) \qquad \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

13

Pinhole Camera Model
Central Projection

- Projection is linear matrix multiplication when using homogeneous coordinates

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad \Rightarrow \left(f\frac{X}{Z}, f\frac{Y}{Z}\right)$$

divide by the third coordinate to convert back to inhomogeneous coordinates

$$\mathbf{x} = P\mathbf{X}$$

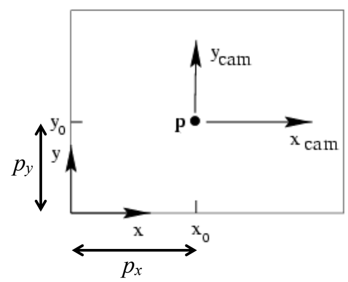
This assumes the image plane origin is at the principal point (the point on image plane that is on the principal axis)

14

Pinhole Camera Model
Principal Point Offset

- Image plane origin is generally not at the principal point
  - Need to account for offset of principal point (p_x, p_y)

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{bmatrix} \Rightarrow (f\frac{X}{Z} + p_x, f\frac{Y}{Z} + p_y)$$



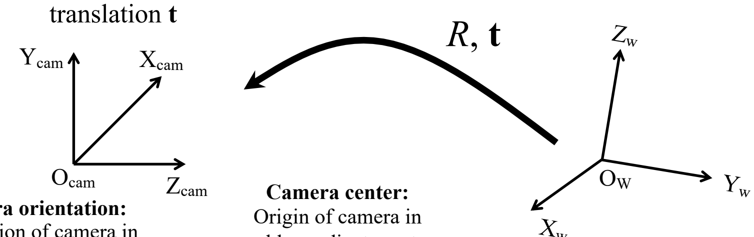
$$\mathbf{x} = K[I|\mathbf{0}]\mathbf{X}_{\text{cam}} = P\mathbf{X}_{\text{cam}}$$

Camera calibration matrix $K = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}$

15

Pinhole Camera Model
Camera Rotation and Translation

- Points in space often expressed in terms of world coordinate frame (not camera coordinate frame)
  - Two Euclidean coordinate systems are related by rotation R and translation $\mathbf{t}$



Camera orientation: Orientation of camera in world coordinate system w.r.t. the camera

Camera center: Origin of camera in world coordinate system

$\mathbf{t} = -R\mathbf{O'}_{\text{cam}}$

$\mathbf{X}_{\text{cam}} = R(\mathbf{X}_W - \mathbf{O'}_{\text{cam}}) = R\mathbf{X}_W + \mathbf{t}$ Inhomogeneous coordinates

$\mathbf{x} = KR[I| - \mathbf{O'}_{\text{cam}}]\mathbf{X}_W = K[R|\mathbf{t}]\mathbf{X}_W = P\mathbf{X}_W$ Homogeneous coordinates

16

Camera

- Pinhole camera model assumes image coordinates are Euclidean and equally scaled in both directions
- Cameras can have non-square pixels
  - Number pixels per unit distance in image coordinates are m_x and m_y
- Need to adjust camera calibration matrix

$$K = \begin{bmatrix} m_x & 0 & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

α_x and α_y represent the scale factors in the x and y dimensions

$\alpha_y/\alpha_x = m_y/m_x$ is the pixel aspect ratio

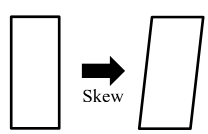
(x_0, y_0) is the principal point in terms of pixel dimensions

17

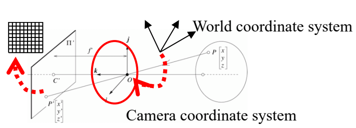
General Projective Camera

- Add generality by incorporating a skew parameter s (parallel to x-axis) into the camera calibration matrix
  - In most normal cameras, $s = 0$

$$K = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$



$$\mathbf{x} = K[R|\mathbf{t}]\mathbf{X}_W = P\mathbf{X}_W$$



DoF = 5 + 3 + 3 = 11

18

Camera Parameters

Extrinsic parameters (**outside** camera):

Intrinsic parameters (**inside** camera):

- *Extrinsic* params: rotation matrix and translation vector
- *Intrinsic* params: focal length, pixel sizes (mm), image center point, radial distortion parameters

19

Estimating Camera Matrix *P*

- Camera matrix *P* can be computed from corresponding 2-D and 3-D points

$$\mathbf{x} = P\mathbf{X}_{\mathrm{W}}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} \leftarrow \begin{bmatrix} x \\ y \\ w \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} \qquad (X_W, Y_W, Z_W) \leftrightarrow \left(\frac{x}{w}, \frac{y}{w}\right) \leftrightarrow (x', y')$$

$$x' = \frac{x}{w} = \frac{p_{11}X_W + p_{12}Y_W + p_{13}Z_W + p_{14}}{p_{31}X_W + p_{32}Y_W + p_{33}Z_W + p_{34}} \qquad y' = \frac{y}{w} = \frac{p_{21}X_W + p_{22}Y_W + p_{23}Z_W + p_{24}}{p_{31}X_W + p_{32}Y_W + p_{33}Z_W + p_{34}}$$

20

Degrees of Freedom

- 12 unknowns → 12 degrees of freedom?
- Multiplying P by nonzero constant k yields same equations → 11 degrees of freedom
  - Hence we could set p_{34} equal to some constant value, but what if it is 0? (We will come back to ensuring 11 DoF in a few slides)

$$x' = \frac{kp_{11}X_W + kp_{12}Y_W + kp_{13}Z_W + kp_{14}}{kp_{31}X_W + kp_{32}Y_W + kp_{33}Z_W + kp_{34}} \qquad y' = \frac{kp_{21}X_W + kp_{22}Y_W + kp_{23}Z_W + kp_{24}}{kp_{31}X_W + kp_{32}Y_W + kp_{33}Z_W + kp_{34}}$$

$$\Rightarrow$$

$$x' = \frac{p_{11}X_W + p_{12}Y_W + p_{13}Z_W + p_{14}}{p_{31}X_W + p_{32}Y_W + p_{33}Z_W + p_{34}} \qquad y' = \frac{p_{21}X_W + p_{22}Y_W + p_{23}Z_W + p_{24}}{p_{31}X_W + p_{32}Y_W + p_{33}Z_W + p_{34}}$$

21

Solving For Camera Matrix P

$$x' = \frac{x}{w} = \frac{p_{11}X_W + p_{12}Y_W + p_{13}Z_W + p_{14}}{p_{31}X_W + p_{32}Y_W + p_{33}Z_W + p_{34}} \qquad y' = \frac{y}{w} = \frac{p_{21}X_W + p_{22}Y_W + p_{23}Z_W + p_{24}}{p_{31}X_W + p_{32}Y_W + p_{33}Z_W + p_{34}}$$

- Multiply by denominators

$$(p_{31}X_W + p_{32}Y_W + p_{33}Z_W + p_{34})x' = p_{11}X_W + p_{12}Y_W + p_{13}Z_W + p_{14}$$
$$(p_{31}X_W + p_{32}Y_W + p_{33}Z_W + p_{34})y' = p_{21}X_W + p_{22}Y_W + p_{23}Z_W + p_{24}$$

- Distribute and set equal to zero

$$p_{11}X_W + p_{12}Y_W + p_{13}Z_W + p_{14} - p_{31}X_W x' - p_{32}Y_W x' - p_{33}Z_W x' - p_{34}x' = 0$$
$$p_{21}X_W + p_{22}Y_W + p_{23}Z_W + p_{24} - p_{31}X_W y' - p_{32}Y_W y' - p_{33}Z_W y' - p_{34}y' = 0$$

22

Solving for Camera Matrix P (continued)

- Write in matrix form

$$\underbrace{\begin{bmatrix} X_{W_1} & Y_{W_1} & Z_{W_1} & 1 & 0 & 0 & 0 & 0 & -X_{W_1}x_i' & -Y_{W_1}x_i' & -Z_{W_1}x_i' & -x_i' \\ 0 & 0 & 0 & 0 & X_{W_1} & Y_{W_1} & Z_{W_1} & 1 & -X_{W_1}y_i' & -Y_{W_1}y_i' & -Z_{W_1}y_i' & -y_i' \\ & & & & & & \vdots & & & & & \end{bmatrix}}_{2N \text{ x } 12} \underbrace{\begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ p_{14} \\ p_{21} \\ p_{22} \\ p_{23} \\ p_{24} \\ p_{31} \\ p_{32} \\ p_{33} \\ p_{34} \end{bmatrix}}_{12 \text{ x } 1} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \end{bmatrix}}_{2N \text{ x } 1}$$

$$A\mathbf{p} = \mathbf{0}$$

- 11 DoF $\rightarrow$ need at least 6 points to solve for P
  - Typically use more than 6

23

Solving for $\mathbf{p}$

- Homogeneous Equations

$$A\mathbf{p} = \mathbf{0}$$

- Multiply by both sides by A^T

$$\underbrace{(A^T A)}_{12\text{x}12}\underbrace{\mathbf{p}}_{12\text{x}1} = B\mathbf{p} = \underbrace{\mathbf{0}}_{12\text{x}1}$$

- Recall Eigen analysis

$$B\mathbf{x} = \lambda\mathbf{x}$$

Least squares would give $\mathbf{p} = \mathbf{0}$
Thus, minimize $\|B\mathbf{p}\|^2$ subject to $\|\mathbf{p}\|^2 = 1$

- Solution of $\mathbf{p}$ given by Eigen decomposition of $(A^T A)$
  - Set $\mathbf{p}$ equal to Eigenvector of $(A^T A)$ corresponding to **smallest** Eigenvalue
  - Normalize $\mathbf{p}$ to unit vector to enforce 11 DoF (most Eigensolvers already do)
- Unrasterize $\mathbf{p}$ to give camera matrix P

Another good technique to remember!

24

Extracting Parameters

- Assuming P is scaled appropriately, it is possible to extract the actual intrinsic and extrinsic camera parameters
- Otherwise, parameters will be correct "up to scale"

25

Overview

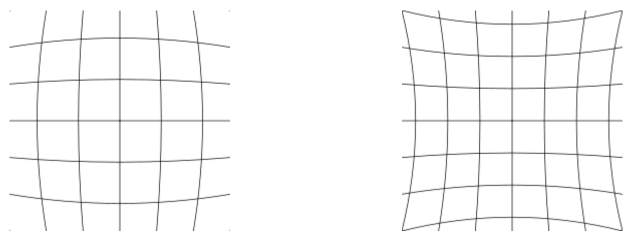
- Thin Lens Model
- Projective Camera Model
  - Intrinsic and Extrinsic Camera Parameters
  - Incorporating Radial Distortion
- Additional Camera Models and Effects

26

Radial Distortion

- Thus far, we have assumed a linear model for the imaging process
  - Assumption does not hold for real lenses
- Lenses generally cause some form of radial distortion
  - Straight lines no longer straight

Barrel Distortion Pincushion Distortion



Zooming Out Zooming In

27

Radial Distortion

- Remove distortion by correcting image measurements to those that would have been obtained under perfect linear camera action

- Actual projection point $\mathbf{x}_d = (x_d, y_d)$ is related to the ideal point $\mathbf{x} = (x, y)$ by a **radial factor** $L(r)$ as $\mathbf{x}_d = L(r)\mathbf{x}$
  - Distortion factor $L(r)$ is a function of the radial distance r from the center of radial distortion
    - $r = (x^2 + y^2)^{1/2}$, where x and y are the horizontal and vertical displacements of the ideal point $\mathbf{x}$ from the center of distortion
  - Approximate function $L(r)$ by Taylor series expansion

    $L(r) = 1 + \kappa_1 r + \kappa_2 r^2 + \kappa_3 r^3 + \ldots$
  - Estimate radial distortion coefficients

28

Radial Distortion Example

29

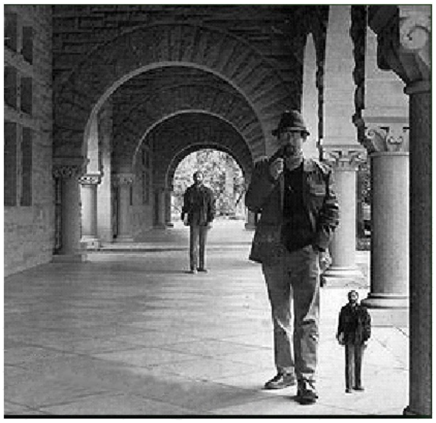
Overview

- Thin Lens Model
- Projective Camera Model
  - Intrinsic and Extrinsic Camera Parameters
  - Incorporating Radial Distortion
- Additional Camera Models and Effects

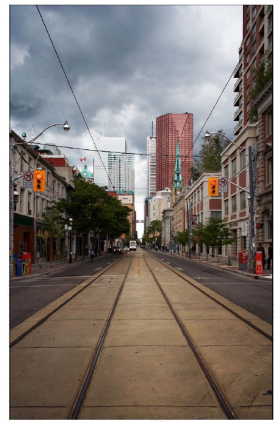
30

Perspective Effects

Far away objects appear smaller



Parallel lines converge at horizon



31

"<u>Forced</u> Perspective"

32

Affine Camera Models

- Special case of projective camera where

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

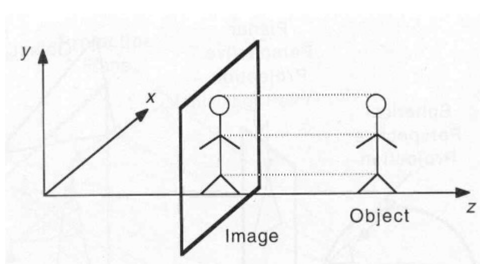
Scaling with Z removed for inhomogeneous points

- Preserve parallelism
- Common affine camera models:
  - Orthographic projection
  - Weak perspective projection

33

Orthographic Projection

- Assume object points are all at the same constant distance from the camera
  - World points projected along rays parallel to optical axis
  - Appropriate when object points are very far away from camera***



$$x' = x$$
$$y' = y$$

34

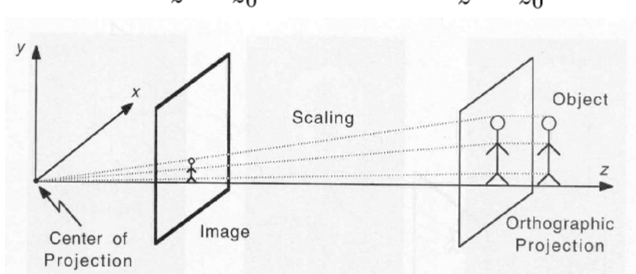
Orthographic Projection

35

Weak Perspective Projection

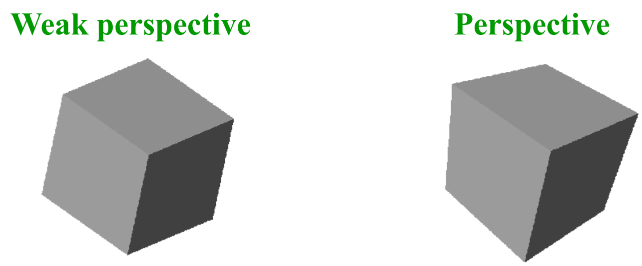
- Assume relative depth of object points << average distance to camera
  - Assume every object point is at constant depth z_0 away from camera and then perform perspective projection (**perform orthographic then perspective**)
  - Closer object bigger than farther object (perspective)

$$x' = f\frac{x}{z} \approx \frac{f}{z_0}x \qquad y' = f\frac{y}{z} \approx \frac{f}{z_0}y$$



36

Pictorial Comparison



37

Summary

- Thin lens model used to model cameras with real lenses
- Pinhole camera model results from assuming aperture diameter is infinitesimally small
- General projective camera model incorporates internal and external parameters
- Several approaches exist to calibrate cameras and extract camera parameters
- Orthographic and weak perspective projections are also common camera models

38