Camera geometry

C306

Fall 2001

Martin Jagersand

Remember: Affine Geometric Transforms

In general, a point in n-D space transforms by

P' = rotate(point) + translate(point)

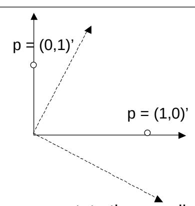
In 2-D space, this can be written as a matrix equation:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} Cos(q) & -Sin(q) \\ Sin(q) & Cos(q) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} tx \\ ty \end{pmatrix}$$

In 3-D space (or n-D), this can generalized as a matrix equation:

p' = R p + T or p = R^t (p' – T)

A Simple 2-D Example



Suppose we rotate the coordinate system through 45 degrees (note that this is measured relative to the *rotated* system!

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} Cos(p/4) & -Sin(p/4) \\ Sin(p/4) & Cos(p/4) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} Cos(p/4) \\ Sin(p/4) \end{pmatrix}$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} Cos(p/4) & -Sin(p/4) \\ Sin(p/4) & Cos(p/4) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} -Sin(p/4) \\ Cos(p/4) \end{pmatrix}$$

Remember: Homogeneous coordinates

- The Homogeneous coordinate corresponding to the point (x,y,z) is the triple (x_h, y_h, z_h, w) where:

  x_h = wx
  y_h = wy
  z_h = wz

  We can (initially) set w = 1.
- Suppose a point P = (x,y,z,1) in the homogeneous coordinate system is mapped to a point P' = (x',y',z',1) by a transformations, then the transformation can be expressed in matrix form.

Matrix representation and Homogeneous coordinates

- For the basic transformations we have:
  - Translation

$$P' = \begin{bmatrix} x' \\ y' \\ z' \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & T_x \\ 0 & 1 & 0 & T_y \\ 0 & 0 & 1 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$$

  - Scaling

$$P' = \begin{bmatrix} x' \\ y' \\ z' \\ w \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}$$

Geometric Transforms

Using the idea of homogeneous transforms, we can write:

$$p' = \begin{pmatrix} & R & & T \\ 0 & 0 & 0 & 1 \end{pmatrix} p$$

R and T both require 3 parameters.

$$R = \begin{bmatrix} \cos & -\sin & 0 \\ \sin & \cos & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos & 0 & \sin \\ 0 & 1 & 0 \\ -\sin & 0 & \cos \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos & -\sin \\ 0 & \sin & \cos \end{bmatrix}$$

Geometric Transforms

If we compute the matrix inverse, we find that

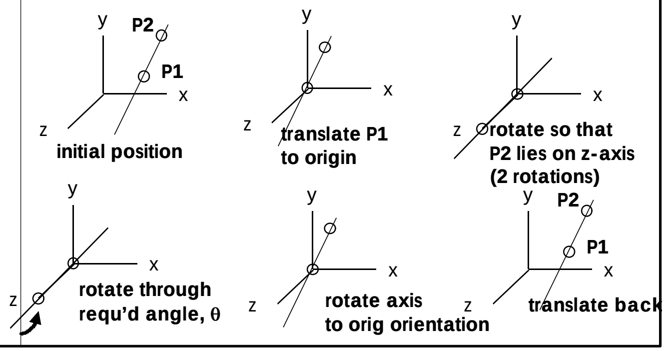
$$p = \begin{pmatrix} & R' & & -R'T \\ 0 & 0 & 0 & 1 \end{pmatrix} p'$$

R and T both require 3 parameters. These correspond to the 6 extrinsic parameters needed for camera calibration

Rotation about a Specified Axis

- It is useful to be able to rotate about any axis in 3D space
- This is achieved by composing 7 elementary transformations (next slide)

Rotation through θ about Specified Axis



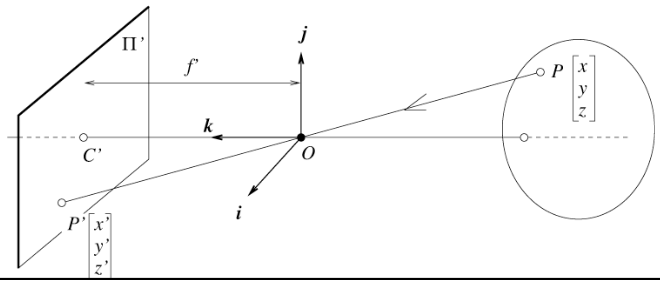
Comparison:

- Homogeneous coordinates
  - Rotations and translations are represented in a uniform way
  - Successive transforms are composed using matrix products: y = Pn*..*P2*P1*x
- Affine coordinates
  - Non-uniform representations: y = Ax + b
  - Difficult to keep track of separate elements

The equation of projection

- Cartesian coordinates:
  - We have, by similar triangles, that (x, y, z) -> (f x/z, f y/z, -f)
  - Ignore the third coordinate, and get

$$(x, y, z) \rightarrow (f\frac{x}{z}, f\frac{y}{z})$$

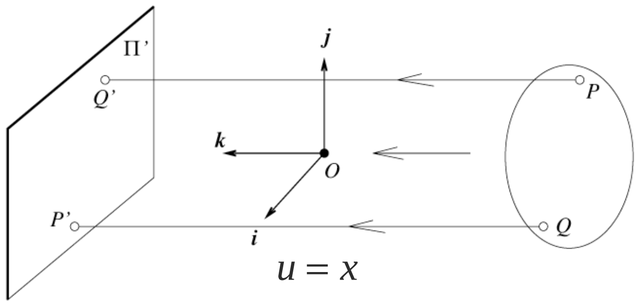


The camera matrix

- Homogenous coordinates for 3D
  - four coordinates for 3D point
  - equivalence relation (X,Y,Z,T) is the same as (k X, k Y, k Z,k T)
- Turn previous expression into HC's
  - HC's for 3D point are (X,Y,Z,T)
  - HC's for point in image are (U,V,W)

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix} \qquad (U,V,W) \rightarrow (\frac{U}{W}, \frac{V}{W}) = (u,v)$$

Orthographic projection



$$u = x$$

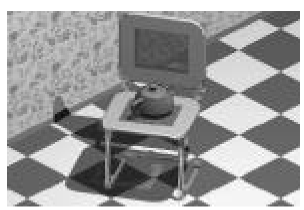
$$v = y$$

The fundamental model for orthographic projection

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Perspective and Orthographic Projection

perspective



Orthographic (parallel)

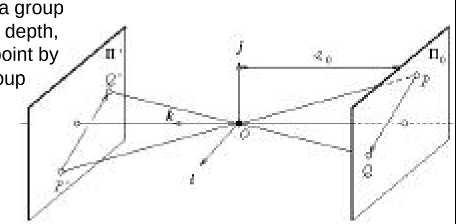
Weak perspective

$$u = Tx$$

$$v = Ty$$

$$T = f / Z$$

- Issue
  - perspective effects, but not over the scale of individual objects
  - collect points into a group at about the same depth, then divide each point by the depth of its group
  - Adv: easy
  - Disadv: wrong



The fundamental model for weak perspective projection

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & f/Z^* \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Camera parameters

- Issue
  - camera may not be at the origin, looking down the z-axis
    - extrinsic parameters
  - one unit in camera coordinates may not be the same as one unit in world coordinates
    - intrinsic parameters - focal length, principal point, aspect ratio, angle between axes, etc.

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{intrinsic parameters} \end{pmatrix} \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{projection model} \end{pmatrix} \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{extrinsic parameters} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Intrinsic Parameters

Intrinsic Parameters describe the conversion from metric to pixel coordinates (and the reverse)

$x_{mm} = -(x_{pix} - o_x)\, s_x$
$y_{mm} = -(y_{pix} - o_y)\, s_y$

or

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix}_{pix} = \begin{pmatrix} -f/s_x & 0 & o_x \\ 0 & -f/s_y & o_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix}_{mm} = M_{\text{int}}\, p$$

Camera Model Structure

Assume R and T express camera *in world coordinates, then*

$$^{c}p = \begin{pmatrix} & R' & & -R'T \\ 0 & 0 & 0 & 1 \end{pmatrix} {}^{w}p$$

Combining with a perspective model (and neglecting internal parameters) yields

$$^{c}u = M\, {}^{w}p = \begin{pmatrix} -R'_x & R'_x T \\ -R'_y & R'_y T \\ \frac{R_z}{f} & \frac{-R_z T}{f} \end{pmatrix} {}^{w}p$$

Note the M is defined only up to a scale factor at this point! If M is viewed as a 3x4 matrix defined up to scale, it is called the *projection matrix.*

Camera Model Structure

Assume R and T express camera *in world coordinates, then*

$$^{c}p = \begin{pmatrix} R' & & & -R'T \\ 0 & 0 & 0 & 1 \end{pmatrix} {}^{w}p$$

Combining with a weak perspective model (and neglecting internal parameters) yields

$$^{c}u = M\,{}^{w}p = \begin{pmatrix} -R'_x & R'_x T \\ -R'_y & R'_y T \\ 0 & \dfrac{R_z(\bar{P} - T)}{f} \end{pmatrix} {}^{w}p$$

Where $\bar{P}$ is the nominal distance to the viewed object

Other Models

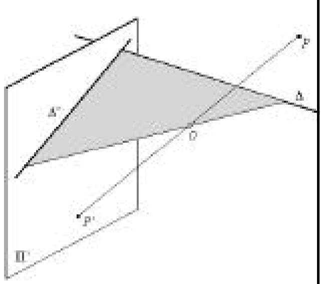
- The *affine camera* is a generalization of weak perspective.
- The *projective camera* is a generalization of the perspective camera.
- Both have the advantage of being linear models on real and projective spaces, respectively.

Camera calibration

- Issues:
  - what are intrinsic parameters of the camera?
  - what is the camera matrix? (intrinsic+extrinsic)
- General strategy:
  - view calibration object
  - identify image points
  - obtain camera matrix by minimizing error
  - obtain intrinsic parameters from camera matrix
- Error minimization:
  - Linear least squares
    - easy problem numerically
    - solution can be rather bad
  - Minimize image distance
    - more difficult numerical problem
    - solution usually rather good, but can be hard to find
      - start with linear least squares
  - Numerical scaling is an issue

Geometric properties of projection

- Points go to points
- Lines go to lines
- Planes go to whole image
- Polygons go to polygons
- Degenerate cases
  - line through focal point to point
  - plane through focal point to line



Polyhedra project to polygons

- (because lines project to lines)