CMPUT 428: Tracking

Roderick Lan

Contents

1 Lecture - Jan 30

lec06GeomIntro.pdf
Image prediction is nonlinear. (Pinhole) Camera collects light rays, projected onto plane (upside down).

$$(x, y, z) \to \left(f\frac{x}{z}, f\frac{y}{z}\right)$$

or Projectively: $x = PX$
3D $\to$ 2D plane.
Choice (and positioning) of coordinate system very important, can simplify problem a lot. (we project img onto new coordinate system?)

1.1 Matrix Rep. + Homogenous Coords

Chain transformations together to build total transformation. Good if all transformations can be represented as matrix mult. (comb. of transformation only involves mult. of matrices); Affine involves addition and multiplication. Translations don't have 2×2, use Homogenous Coords to get 3×3 matrix rep for transformation.
Append extra coords to 2D:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \Rightarrow \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$\mathbb{R}^4$ space sufficient for all Euclidean operaitons.

1.1.1 General Case

Homog coord corresponding to (x, y, z) is the triple (x_h, y_h, z_h, w) s.t.

$$x_h = wx$$
$$y_h = wy$$
$$z_h = wz$$

Initially set $w = 1$. May perform differently w/ transformations (?)

1.1.2 Transformations

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: Translation:

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

Using general idea of Homog. Transformation:

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

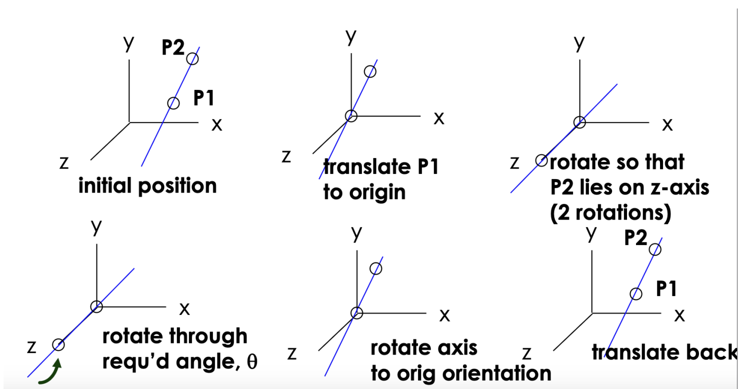
Inverse:

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

(gen. rotation in slides (42))

1.2 Rotation abt Specified Axis

Useful to rotate about any axis in 3D space → compose 7 elementary transformations.
Issue w/ just rotating: rotates about origin (object 'translates' and rotates). Need to center object and align w/ coordinate system. (often times rotate on z axis).



1.3 Basic Transformations

Translation, scaling, etc.
Scaling be rewritten for fixed point.

1.4 Camera Matrix

$$(U, V, W) \rightarrow \left(\frac{U}{W}, \frac{V}{W}\right) = (u, v)$$

Expln 1.4.1

W is often called the **scale factor**

Homog Coords for 3D:

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

coord sys not aligned:

- projective (proj coords $\times$ arbitrary (3×4) matrix P)
  $x = PX$
- euclidean
  $x = [R|T]X$

Expln 1.4.2

Projective geometry preserves collinearity

1.5 Intersection (preview)

can extend when we know img in plane and where camera is geometrically.
need tracking for complex imgs (very important for correspondence)

1.6 SFM (preview)

Only have (unorganized) imgs, dont know where they were taken from. Multiple images of same 3d object is redundant (?)

1.7 N-view geometry; Affine Factorization

stack points in measurement matrix W, perform SVD (heath ch3), get mathematically equivalent camera position ($\hat{M}$) and points ($\hat{X}$)
$W = UDV^\top$
$\hat{W} = U_{2m\times 3} D_{3\times 3} V^\top_{n\times 3} = \hat{M}\hat{X}$

$$W = \begin{bmatrix} \tilde{x}_1^1 & \cdots & \tilde{x}_n^1 \\ \vdots & \ddots & \vdots \\ \tilde{x}_1^m & \cdots & \tilde{x}_n^m \end{bmatrix} = \begin{bmatrix} M^1 \\ \vdots \\ M^m \end{bmatrix} \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

1.8 Limitations in images

distant objects smaller; img prediction can rearrange position and scale arbitrarily (can't trust size relationships)
Visual ambiguity.

Parallel lines meet at a point in projective geometry (vanishing point)

1.9 Invariants

Things that stay the same after every projective transform

- points map to points
- intersections are preserved (not necessarily at the same real/metric distances)
- lines map to lines
- collinearity preserved
- ratio of ratios (cross ratio); relative distances (no abs. distance)
- horizon

1.10 Vanishing points

Each set of parallel ines (direction) meets at a diff. point - **vanishing point** for this direction.
Sets of parallel lines on same plane lead to **collinear** vanishing point (**horizon** for that plane)

1.11 Geometric Properties of Projection

- points go to points
- lines go to lines
- (mathematical) planes go to whole img
- polygons go to polygons

Degenerate cases:

- line through focal point to point
- plane through focal point to line