CAP 5415 Computer Vision
Fall 2005

Dr. Alper Yilmaz

Univ. of Central Florida

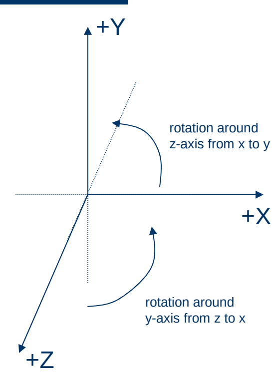
www.cs.ucf.edu/courses/cap5415/fall2005

Office: CSB 250

Confusion from Last Lecture

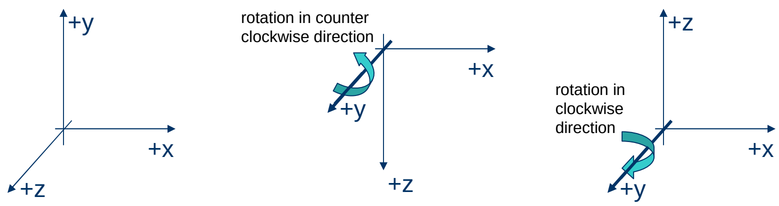
- Around Z-axis $R^Z = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$
- Around Y-axis $R^Y = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$

minus sign is here



Confusion from Last Lecture

- Given a coordinate system, rotations are always in counter clockwise direction!!



original coordinate system

rotated coordinate system

different coordinate system

Correction

- Euler's angles: If angle θ is small, then $\cos\theta = 1$ and $\sin\theta = \theta$

$$R = R_Z^{\alpha} R_Y^{\beta} R_X^{\gamma}$$

$$R = R_Y^{\beta} R_Z^{\alpha} R_X^{\gamma}$$

$$\boxed{R = \begin{bmatrix} 1 & -\alpha - \gamma\beta & \beta - \gamma\alpha \\ \alpha & 1 - \alpha\beta\gamma & -\beta\alpha - \gamma \\ -\beta & \gamma & 1 \end{bmatrix}}$$

$$R = R_X^{\gamma} R_Y^{\beta} R_Z^{\alpha}$$

$$R = \begin{bmatrix} 1 & -\alpha & \beta \\ \alpha - \gamma\beta & 1 + \alpha\beta\gamma & -\gamma \\ -\beta - \gamma\alpha & \gamma - \beta\alpha & 1 \end{bmatrix}$$

$$\vdots$$

Summary

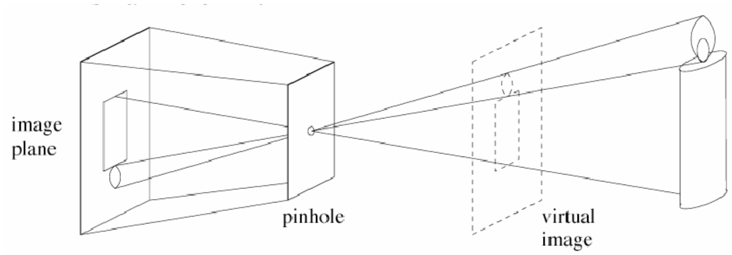


image coordinate system

camera coordinate system

real world coordinate system

5

Alper Yilmaz, CAP5415, Fall 2005

Summary

- Perspective projection

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

- Intrinsic camera parameters

$$x_{image} = k_x x_{camera} + o_x \qquad y_{image} = k_y y_{camera} + o_y$$

- Extrinsic camera parameters

o_x and o_y has various names:
• Principal point
• Center of projection

$$\begin{bmatrix} X_{camera} \\ Y_{camera} \\ Z_{camera} \\ 1 \end{bmatrix} = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_x \\ r_{2,1} & r_{2,2} & r_{2,3} & t_y \\ r_{3,1} & r_{3,2} & r_{3,3} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_{world} \\ Y_{world} \\ Z_{world} \\ 1 \end{bmatrix}$$

6

Alper Yilmaz, CAP5415, Fall 2005

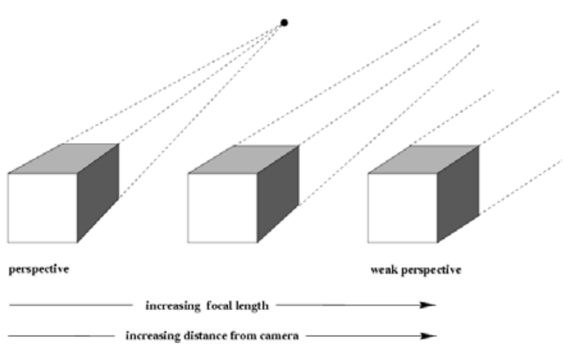
Summary

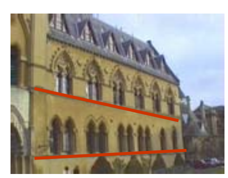
$$\overbrace{\begin{bmatrix} U_{image} \\ V_{image} \\ S \end{bmatrix}}^{\substack{\text{homogenous} \\ \text{image} \\ \text{coordinates}}} = \overbrace{\begin{bmatrix} -fk_x & 0 & 0 & o_x \\ 0 & -fk_y & 0 & o_y \\ 0 & 0 & 1 & 0 \end{bmatrix}}^{\text{intrinsic camera parameters}} \overbrace{\begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_x \\ r_{2,1} & r_{2,2} & r_{2,3} & t_y \\ r_{3,1} & r_{3,2} & r_{3,3} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}}^{\text{extrinsic camera parameters}} \overbrace{\begin{bmatrix} X_{world} \\ Y_{world} \\ Z_{world} \\ 1 \end{bmatrix}}^{\substack{\text{homogenous} \\ \text{world} \\ \text{coordinates}}}$$

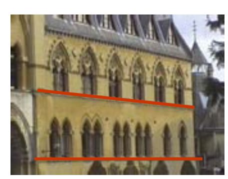
$$x_{image} - o_x = -f_x \frac{r_{1,1}X_{world} + r_{1,2}Y_{world} + r_{1,3}Z_{world} + t_x}{r_{3,1}X_{world} + r_{3,2}Y_{world} + r_{3,3}Z_{world} + t_z}$$

$$y_{image} - o_y = -f_y \frac{r_{2,1}X_{world} + r_{2,2}Y_{world} + r_{2,3}Z_{world} + T_y}{r_{3,1}X_{world} + r_{3,2}Y_{world} + r_{3,3}Z_{world} + T_z}$$

Affine cameras







Hierarchy of Affine Cameras

Distant camera
We do not care about intrinsic parameters
We do not care about extrinsic parameters

$$P_\infty = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Orthographic projection
We do not care about intrinsic parameters

$$P_\infty = \begin{bmatrix} r^{1T} & t_1 \\ r^{1T} & t_2 \\ 0 & 1 \end{bmatrix}$$

Scaled orthographic projection
We do not care about intrinsic parameters

$$P_\infty = \begin{bmatrix} r^{1T} & t_1 \\ r^{1T} & t_2 \\ 0 & 1/k \end{bmatrix}$$

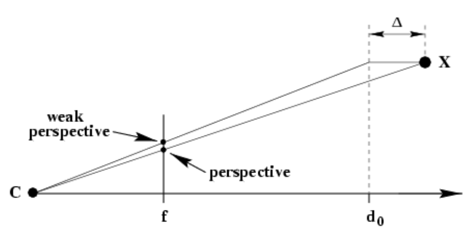
9

Alper Yilmaz, CAP5415, Fall 2005

Hierarchy of Affine Cameras

Weak perspective projection
We care about a set of intrinsic parameters

$$P_\infty = \begin{bmatrix} f_x & & \\ & f_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} r^{1T} & t_1 \\ r^{1T} & t_2 \\ 0 & 1/k \end{bmatrix}$$



10

Alper Yilmaz, CAP5415, Fall 2005

Hierarchy of Affine Cameras

Affine camera

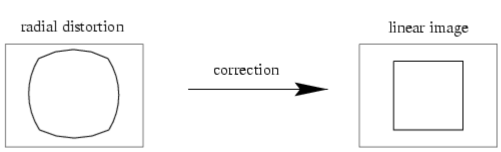
s is shearing affect

$$P_A = \begin{bmatrix} f_x & s & \\ & f_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} r^{1T} & t_1 \\ r^{1T} & t_2 \\ 0 & 1/k \end{bmatrix}$$

- *Affine camera:* Camera with principal plane is at infinity
- Affine camera maps parallel lines to parallel lines
- No center of projection

Radial Distortion

Will not be on the exams



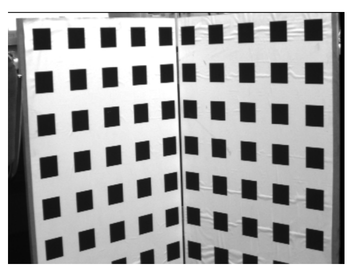
Radial Distortion

13

Alper Yilmaz, CAP5415, Fall 2005

Estimating Camera Parameters

(x_1, y_1)	(X_1, Y_1, Z_1)
(x_2, y_2)	(X_2, Y_2, Z_2)
(x_3, y_3)	(X_3, Y_3, Z_3)
(x_N, y_N)	(X_N, Y_N, Z_N)



14

Alper Yilmaz, CAP5415, Fall 2005

Estimating Camera Parameters

- For each corresponding pair (x_i, y_i) and (X_i, Y_i , Z_i), we have

$$x_i - o_x = -f_x \frac{r_{1,1}X_i + r_{1,2}Y_i + r_{1,3}Z_i + t_x}{r_{3,1}X_i + r_{3,2}Y_i + r_{3,3}Z_i + t_z} \qquad \textbf{(A)}$$

$$y_i - o_y = -f_y \frac{r_{2,1}X_i + r_{2,2}Y_i + r_{2,3}Z_i + t_x}{r_{3,1}X_i + r_{3,2}Y_i + r_{3,3}Z_i + t_z} \qquad \textbf{(B)}$$

15

Estimating Camera Parameters

- Divide equation (A) and (B)

$$\frac{x_i - o_x}{y_i - o_y} = \frac{-f_x}{-f_y} \frac{\dfrac{r_{1,1}X_i + r_{1,2}Y_i + r_{1,3}Z_i + t_x}{\cancel{r_{3,1}X_i + r_{3,2}Y_i + r_{3,3}Z_i + t_z}}}{\dfrac{r_{2,1}X_i + r_{2,2}Y_i + r_{2,3}Z_i + t_y}{\cancel{r_{3,1}X_i + r_{3,2}Y_i + r_{3,3}Z_i + t_z}}}$$

- Let $o_x = o_y = 0$ and $\alpha = f_x / f_y$

16

Estimating Camera Parameters

$$\frac{x_i}{y_i} = \alpha \frac{r_{1,1}X_i + r_{1,2}Y_i + r_{1,3}Z_i + t_x}{r_{2,1}X_i + r_{2,2}Y_i + r_{2,3}Z_i + t_z}$$

After division of (A) and (B)

$$x_i\left(r_{2,1}X_i + r_{2,2}Y_i + r_{2,3}Z_i + t_z\right) = \alpha y_i\left(r_{1,1}X_i + r_{1,2}Y_i + r_{1,3}Z_i + t_x\right)$$

$$x_iX_iv_1 + x_iY_iv_2 + x_iZ_iv_3 + x_iv_4 - y_iX_iv_5 - y_iY_iv_6 - y_iZ_iv_7 - y_iv_8 = 0 \quad \textbf{(C)}$$

where

$$\begin{matrix} v_1 = r_{21} & v_5 = \alpha r_{11} \\ v_2 = r_{22} & v_6 = \alpha r_{12} \\ v_3 = r_{23} & v_7 = \alpha r_{13} \\ v_4 = t_y & v_8 = \alpha t_x \end{matrix}$$

Estimating Camera Parameters

- Let's rearrange (C) into Av=0 form

$$x_iX_iv_1 + x_iY_iv_2 + x_iZ_iv_3 + x_iv_4 - y_iX_iv_5 - y_iY_iv_6 - y_iZ_iv_7 - y_iv_8 = 0$$

- Given N points

$$\underbrace{\begin{bmatrix} x_1X_1 & x_1Y_1 & x_1Z_1 & x_1 & y_1X_1 & y_1Y_1 & y_1Z_1 & y_1 \\ x_2X_2 & x_2Y_2 & x_2Z_2 & x_2 & y_2X_2 & y_2Y_2 & y_2Z_2 & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_NX_N & x_NY_N & x_NZ_N & x_N & y_NX_N & y_NY_N & y_NZ_N & y_N \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \\ v_7 \\ v_8 \end{bmatrix}}_{v} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \textbf{(D)}$$

Estimating Camera Parameters

- Given $Av=0$ compute v by Singular Value Decomposition of A into $A=UDV^{T}$
- Solution of v "up to a scale factor γ" is the column of V corresponding to smallest singular value in D.

$$\tilde{v} = \gamma \begin{bmatrix} r_{21} & r_{22} & r_{23} & t_y & \alpha r_{11} & \alpha r_{12} & \alpha r_{13} & \alpha t_x \end{bmatrix}$$

19

Alper Yilmaz, CAP5415, Fall 2005

Estimating Camera Parameters

- Computing the scale γ

Since $r_{21}^2 + r_{22}^2 + r_{23}^2 = 1$, γ can be computed from:

$$\sqrt{v_1^2 + v_2^2 + v_3^2} = \sqrt{\gamma^2 \left(r_{21}^2 + r_{22}^2 + r_{23}^2\right)} = |\gamma|$$

- Computing 3rd row of rotation matrix

$$r_3 = \underbrace{r_1 \times r_2}_{\text{cross product of two vectors}}$$

because every rotation matrix is an orthonormal matrix

20

Alper Yilmaz, CAP5415, Fall 2005

Dot and Cross Products

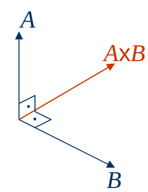
- *Dot product*

$$A^T.B = \begin{bmatrix} A_x & A_y & A_z \end{bmatrix} \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} = A_x B_x + A_y B_y + A_z B_z$$

- *Cross product:* creates an orthogonal vector

$$A \times B = (A_x, A_y, A_z) \times (B_x, B_y, B_z)$$

$$= \begin{vmatrix} i & j & k \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} = (A_y B_z - A_z B_y, -(A_x B_z - A_z B_x), A_x B_y - A_y B_x)$$



21

Alper Yilmaz, CAP5415, Fall 2005

$$x_i - o_x = -f_x \frac{r_{1,1} X_i + r_{1,2} Y_i + r_{1,3} Z_i + t_x}{r_{3,1} X_i + r_{3,2} Y_i + r_{3,3} Z_i + t_z}$$

Estimating Camera Parameters

- Computing f_x, f_y and t_z

For each corresponding pair (x_i, y_i) and (X_i, Y_i, Z_i)

$$x_i (r_{31} X_i + r_{32} Y_i + r_{33} Z_i + t_z) = -f_x (r_{11} X_i + r_{12} Y_i + r_{13} Z_i + t_x)$$

$$\underbrace{\begin{bmatrix} x_1 & r_{11} X_1 + r_{12} Y_i + r_{13} Z_i + t_x \\ x_2 & r_{11} X_2 + r_{12} Y_2 + r_{13} Z_2 + t_x \\ \vdots & \vdots \\ x_N & r_{11} X_N + r_{12} Y_N + r_{13} Z_N + t_x \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} t_z \\ f_x \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} -x_1 r_{31} X_1 - x_1 r_{32} Y_1 - x_1 r_{33} Z_1 \\ -x_2 r_{31} X_2 - x_2 r_{32} Y_2 - x_2 r_{33} Z_2 \\ \vdots \\ -x_N r_{31} X_N - x_N r_{32} Y_N - x_N r_{33} Z_N \end{bmatrix}}_{C}$$

22

Alper Yilmaz, CAP5415, Fall 2005

Solving System of Equations

$A_{N\times M} B_{M\times 1} = C_{N\times 1}$ N equations M unknowns

$$A^T AB = A^T C$$

$$\underbrace{\left(A^T A\right)^{-1}\left(A^T A\right)}_{\text{identity}} B = \left(A^T A\right)^{-1} A^T C$$

$$B = \left(A^T A\right)^{-1} A^T C \qquad \textbf{(E)}$$

- Difference between **(D)** and **(E)** is:
  - **(D)** is a homogenous system
  - **(D)** may have many solutions
  - **(D)** can be solved by SVD

Estimating Camera Parameters

- Computing o_x and o_y
  - We are given all other parameters

$$o_x = x_i + f_x \frac{r_{1,1}X_i + r_{1,2}Y_i + r_{1,3}Z_i + t_x}{r_{3,1}X_i + r_{3,2}Y_i + r_{3,3}Z_i + t_z}$$

$$o_y = y_i + f_y \frac{r_{2,1}X_i + r_{2,2}Y_i + r_{2,3}Z_i + t_x}{r_{3,1}X_i + r_{3,2}Y_i + r_{3,3}Z_i + t_z}$$

Application

$$M = \begin{bmatrix} .17237 & -.15879 & .01879 & 274.943 \\ .131132 & .112747 & .2914 & 258.686 \\ .000346 & .0003 & .00006 & 1 \end{bmatrix}$$

FIGURE 8 PHOTOGRAPH OF SAN FRANCISCO

FIGURE 9 MAP OF SAN FRANCISCO

Camera location: intersection of California and Mason streets, at an elevation of 435 feet above sea level. The camera was oriented at an angle of 8^0 above the horizon. $fs_x=495$, $fs_y=560$.

Application

$$M = \begin{bmatrix} -.175451 & -.10520 & .00435 & 297.83 \\ .02698 & -.09635 & .2303 & 249.574 \\ .00015 & -.00016 & .00001 & 1.0 \end{bmatrix}$$

FIGURE 10 ANOTHER PHOTOGRAPH OF SAN FRANCISCO

FIGURE 11 MAP OF SAN FRANCISCO

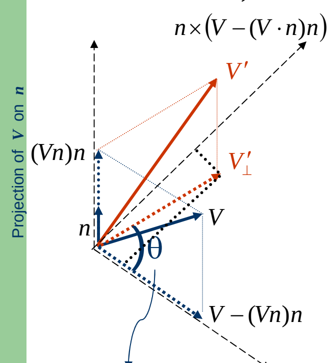
Camera location: at an elevation of 1200 feet above sea level. The camera was oriented at an angle of 4^0 above the horizon. $fs_x=876$, $fs_y=999$.

Back To Rotation

Rodriguez's Formula

Rotation Around Arbitrary Axis

Let V be a vector, n be the rotation axis (n is unit vector). What is V'?



θ Is amount of rotation

$$V = (V \cdot n)n + (V - (V \cdot n)n)$$

$$V'_{\perp} = \cos\theta(V - (V \cdot n)n) + \sin\theta(n \times (V - (V \cdot n)n))$$

$$V'_{\perp} = \cos\theta(V - (V \cdot n)n) + \sin\theta(n \times V)$$

$$V' = V'_{\perp} + (V \cdot n)n$$

$$V' = \cos\theta(V - (V \cdot n)n) + \sin\theta(n \times V) + (V \cdot n)n$$

Rotation Around Arbitrary Axis

$$n \times (n \times V) = (V.n)n - V \Rightarrow n \times (n \times V) + V = (V.n)n$$

$$V' = -\cos\theta\left(n \times (n \times V)\right) + \sin\theta\left(n \times V\right) + n \times (n \times V) + V$$

$$V' = V + \sin\theta\left(n \times V\right) + (1 - \cos\theta)\left(n \times (n \times V)\right)$$

$V' = R(n,\theta)V$ where $R(n,\theta) = I + \sin\theta\, X(n) + (1-\cos\theta) X^2(n)$

$$X(n) = \begin{bmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{bmatrix}$$

29

Alper Yilmaz, CAP5415, Fall 2005

- Homework: Derive $R(n,\theta)$
  - Due September 14,2005
- Suggested reading
  - Chapter 2.4 and 6.3, Emanuele Trucco, Alessandro Verri, "Introductory Techniques for 3-D Computer Vision", Prentice Hall, 1998
  - Chapter 1, Mubarak Shah, "Fundamentals of Computer Vision"

30

Alper Yilmaz, CAP5415, Fall 2005