Agenda

• Project 3
  • Grades & soln in EEE dropbox
• Project 4: Due in 1 week!
  • Reduced load & head start with skeleton code
• Homography estimation
• Least squares
• Warping

Slides from Andrew Zisserman, Serge Belongie, Steve Seitz, Rick Szeliski
Some of your results...  

Jason Newton
Some of your results...

Sam Hallman
2-view geometry

Corresponding points are images of the same scene point

Triangulation

The back-projected points generate rays which intersect at the 3D scene point
Human brain triangulates as well

cross-eye viewing random dot stereogram

Eyes = 2 cameras with small baseline
Epipolar geometry continued

Epipolar geometry is a consequence of the coplanarity of the camera centres and scene point.

The camera centres, corresponding points and scene point lie in a single plane, known as the epipolar plane.
Epipolar lines from pure translation
Parallax

E is determined by relative location of cameras \((R,T)\)

- Let \(X_1\) be 3d non-homogenous point “p” viewed from cam1 (likewise for cam2)
- Define position of cam1 to be at origin \((R_1 = I, T_1 = 0)\)
- Assume instrinsics \(K\) are identical for both cameras

\[
X_2 = RX_1 + T
\]
Map 3D to 2D

Write 3D points as 2D homogenous points with unknown depth

\[ X_1 = \lambda_1 x_1, \quad X_2 = \lambda_2 x_2 \]

\[ \lambda_2 x_2 = R \lambda_1 x_1 + T \]
Matrix representation of the vector cross product

The vector product $\mathbf{v} \times \mathbf{x}$ can be represented as a matrix multiplication

$$
\mathbf{v} \times \mathbf{x} = \begin{pmatrix}
  v_2x_3 - v_3x_2 \\
  v_3x_1 - v_1x_3 \\
  v_1x_2 - v_2x_1
\end{pmatrix} = [\mathbf{v}]_\times \mathbf{x}
$$

where

$$
[\mathbf{v}]_\times = \begin{bmatrix}
  0 & -v_3 & v_2 \\
  v_3 & 0 & -v_1 \\
  -v_2 & v_1 & 0
\end{bmatrix}
$$

- $[\mathbf{v}]_\times$ is a $3 \times 3$ skew-symmetric matrix of rank 2.
- $\mathbf{v}$ is the null-vector of $[\mathbf{v}]_\times$, i.e. $[\mathbf{v}]_\times \mathbf{v} = 0$,
  since $\mathbf{v} \times \mathbf{v} = [\mathbf{v}]_\times \mathbf{v} = 0$
Essential matrix

\[
\lambda_2 x_2 = R \lambda_1 x_1 + T
\]

Take the cross product of both sides with \(T\),

\[
\lambda_2 \hat{T} x_2 = \hat{T} R \lambda_1 x_1 + \hat{T}T \underline{=} 0
\]

take the inner product with \(x_2\),

\[
\lambda_2 x_2^\top \hat{T} x_2 \underline{=} x_2^\top \hat{T} R \lambda_1 x_1 \underline{=} 0
\]

\[
x_2^\top \hat{T} R x_1 = 0
\]

\[
x_2^\top E x_1 = 0
\]
Essential matrix

\[ x_2^\top E x_1 = 0 \]

\[ E = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \in \mathbb{R}^{3\times3} \]

Thus to map a point in one image to a line in the other using the essential matrix, we apply the following equations:

\[ l_2 \sim E x_1 \]

(2.13)

\[ x_2^\top l_2 = 0 \]

Alternatively, you can go the other way:

\[ l_1 \sim E^\top x_2 \]

(2.14)

\[ x_1^\top l_1 = 0 \]

where \( l_1, l_2 \) are epipolar lines (specified in homogeneous coordinates).
Can we find a direct linear map
\[ x_2 = Hx_1 \]?

- Recall such a transformation is a homography
- Let’s make some simplifying assumptions
Scene is planar

Let $X_1$ be any point on plane $P$
Let $N$ be a unit-norm vector pointing from cam1 to plane
Let $d = \text{distance from plane to cam1}$

$$N^\top X_1 = n_1 X + n_2 Y + n_3 Z = d$$

$$\frac{1}{d} N^\top X_1 = 1 \quad \forall X_1 \in P$$
Homography derivation

\[ X_2 = RX_1 + T \]

\[ \frac{1}{d} N^T X_1 = 1 \quad \forall X_1 \in P \]

Multiply both together....

\[ X_2 = RX_1 + T \frac{1}{d} N^T X_1 = HX_1 \]

\[ H = R + \frac{1}{d} TN^T, \quad H \in \mathbb{R}^{3 \times 3} \]

Almost there...
We’ve defined a linear mapping for 3D points on a plane
\[ X_2 = HX_1 \]
3D to 2D

\[ X_2 = H X_1 \quad X_1, X_2 \in \mathbb{R}^3 \]

\[ \lambda_1 x_1 = X_1, \quad \lambda_2 x_2 = X_2, \quad \text{therefore} \quad \lambda_2 x_2 = H \lambda_1 x_1 \]

We can write

\[ x_2 \sim H x_1 \]

where \( \sim \) means equal in homogenous coordinates

Done! - given 3x3 matrix H, we know how to map every image coordinate in image1 to image2
Purely rotating camera

\[ X_2 = RX_1 \]

\[ X_2 = HX_1 \quad X_1, X_2 \in \mathbb{R}^3 \]

\[ H = R + \frac{1}{d}TN^\top \]

H approaches R as d approaches infinity

How do pixels change as camera rotates?

We can model as a homography transformation where the plane is infinitely far away
Rotation mosaics

Project 4

Reference material: online notes from UCSD course (will place link)
Project 4

- “Mosaic” = composite all images transformed to align with base image
- For every base+im pair, label a bunch of point correspondences between 2 images
- We discussed automatic ways of doing this (harris corner + sift matching)
- You’ll do this by hand
- Given point correspondences for a given base+im pair, we want to estimate H
- H tells us how to transform im to align with base
  \[ x_2 \sim Hx_1 \]
How many points do we need?

• In homogenous coordinates...

\[
\begin{bmatrix}
  x_2 \\
  y_2 \\
  z_2
\end{bmatrix} = \begin{bmatrix}
  H_{11} & H_{12} & H_{13} \\
  H_{21} & H_{22} & H_{23} \\
  H_{31} & H_{32} & H_{33}
\end{bmatrix} \begin{bmatrix}
  x_1 \\
  y_1 \\
  z_1
\end{bmatrix} \Leftrightarrow x_2 = Hx_1
\]

How many DOFs in H?
How many constraints does a single correspondence give?
Homography estimation

In inhomogenous coordinates \( (x'_2 = x_2/z_2 \text{ and } y'_2 = y_2/z_2) \),

\[
x'_2 = \frac{H_{11}x_1 + H_{12}y_1 + H_{13}z_1}{H_{31}x_1 + H_{32}y_1 + H_{33}z_1}
\]

\[
y'_2 = \frac{H_{21}x_1 + H_{22}y_1 + H_{23}z_1}{H_{31}x_1 + H_{32}y_1 + H_{33}z_1}
\]

Without loss of generality, set \( z_1 = 1 \) and rearrange:

\[
x'_2(H_{31}x_1 + H_{32}y_1 + H_{33}) = H_{11}x_1 + H_{12}y_1 + H_{13}
\]

\[
y'_2(H_{31}x_1 + H_{32}y_1 + H_{33}) = H_{21}x_1 + H_{22}y_1 + H_{23}
\]

We want to solve for \( H \). Even though these inhomogeneous equations involve the coordinates nonlinearly, the coefficients of \( H \) appear linearly.
Cont’d

To estimate $H$, we start from the equation $\mathbf{x}_2 \sim H\mathbf{x}_1$ and cross both sides with $\mathbf{x}_2$:

$$\hat{x}_2\mathbf{x}_2 \sim \hat{x}_2 H\mathbf{x}_1$$
$$\Rightarrow \hat{x}_2 H\mathbf{x}_1 = 0$$

To solve, stack $H$ into $H^s \in \mathbb{R}^9$ (i.e. $\mathbb{H}(:) \text{ in matlab}$).

$$a^\top H^s = 0.$$  

“$a$” is a 9X3 matrix of cross-terms

The above is a constraint from a single correspondence

(on board)
Homogenous linear system

Collect the $a$'s for each correspondence into a “design matrix” $\chi$, 

$$\chi = [a^1 \ a^2 \ \cdots \ a^n]^\top \in \mathbb{R}^{3n \times 9}, \text{ then } \chi H^s = 0$$

If we have $N$ correspondences, $X = 3N \times 9$ matrix

How to solve?
Homogenous linear system

Given a (possibly non-square) matrix $A$, we want to solve for $z$

$$Az = 0$$

Any matrix $A$ can be decomposed into a product of 3 matrices

$$A = UΣV^T$$

(On board description)

Solution $c$ is column from $V$ associated with smallest singular value
Homogenous linear system

Collect the $a$’s for each correspondence into a “design matrix” $\chi$,

$$\chi = [a^1 \ a^2 \ \cdots \ a^n]^\top \in \mathbb{R}^{3n \times 9}, \ \text{then} \quad \chi H^s = 0$$

Given $X$, we can find $H$ by

$$\gg [U,S,V] = \text{svd}(X);$$

If singular values are already sorted from largest to smallest,

$$\gg H = V(:,\text{end});$$

Reshape 9x1 vector $H$ into 3x3 matrix

$$\gg H = \text{reshape}(H, 3, 3);$$
Warping