Transformations Between Two Images

- Translation
- Rotation
- Rigid
- Similarity (scaled rotation)
- Affine
- Projective
- Pseudo Perspective
- Bi-linear
Fundamental Matrix

Lecture-13
Applications

• Stereo
• Structure from Motion
• View Invariant Action Recognition
• ..
Stereo Pairs and Depth Maps (from Szeliski’s book)
Image Rectification For Stereo
Photosynth: Structure From Motion
Fundamental Matrix

• Longuet Higgins (1981)
• Hartley (1992)
• Faugeras (1992)
• Zhang (1995)
Fundamental Matrix Song

http://www.youtube.com/watch?v=DgGV3l82NTk
Fundamental Matrix Song
Preliminaries

- Linear Independence
- Rank of a Matrix
- Matrix Norm
- Singular Value Decomposition
- Vector Cross product to Matrix Multiplication
- RANSAC
A finite subset of $n$ vectors, $v_1, v_2, ..., v_n$, from the vector space $V$, is linearly dependent if and only if there exists a set of $n$ scalars, $a_1, a_2, ..., a_n$, not all zero, such that

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n = 0.$$
Rank of a Matrix

- The **column rank** of a matrix $A$ is the maximum number of linearly independent column vectors of $A$.
- The **row rank** of a matrix $A$ is the maximum number of linearly independent row vectors of $A$.
- The column rank of $A$ is the dimension of the column space of $A$.
- The row rank of $A$ is the dimension of the row space of $A$. 

Copyright  Mubarak Shah 2003
Example (Row Echelon)

\[
A = \begin{bmatrix}
1 & 2 & 1 \\
-2 & -3 & 1 \\
3 & 5 & 0
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 2 & 1 \\
-2 & -3 & 1 \\
3 & 5 & 0
\end{bmatrix} \rightarrow 2r_1 + r_2
\]

Rank is 2
Singular Value Decomposition (SVD)

Theorem: Any $m$ by $n$ matrix $A$, for which $m \geq n$, can be written as

$$A = O_1 \Sigma O_2$$

$\Sigma$ is diagonal

$O_1, O_2$ are orthogonal

$$O_1^T O_1 = O_2^T O_2 = I$$
Singular Value Decomposition (SVD)

If $A$ is square, then $O_1, \Sigma, O_2$ are all square.

\[
O_1^{-1} = O_1^T
\]
\[
O_2^{-1} = O_2^T
\]
\[
\Sigma^{-1} = \text{diag}\left(\frac{1}{w_j}\right)
\]
\[
A = O_1 \Sigma O_2
\]
\[
A^{-1} = O_2 \text{diag}\left(\frac{1}{w_j}\right) O_1
\]
Matrix Norm

$L_1$ matrix norm is maximum of absolute column sum.

$$\|A\|_1 = \max_{1 \leq j \leq m} \sum_{i=1}^{n} |a_{ij}|,$$
Vector Cross Product to Matrix-vector multiplication

\[
A \times B = \begin{bmatrix}
  i & j & k \\
  A_x & A_y & A_z \\
  B_x & B_y & B_z
\end{bmatrix} = \begin{bmatrix}
  -A_z B_y + A_y B_z & A_z B_x - A_x B_z & -B_x A_y + B_y A_x \\
  \end{bmatrix}
\]

\[
A \times B = S.B = \begin{bmatrix}
  0 & -A_z & A_y \\
  A_z & 0 & -A_x \\
  -A_y & A_x & 0
\end{bmatrix} \begin{bmatrix}
  B_x \\
  B_y \\
  B_z
\end{bmatrix} = \begin{bmatrix}
  -A_z B_y + A_y B_z & A_z B_x - A_x B_z & -B_x A_y + B_y A_x \\
  \end{bmatrix}
\]
How to Fit A Line?

• Least squares Fit (over constraint)
• RANSAC (constraint)
• Hough Transform (under constraint)
RANSAC Song

http://www.youtube.com/watch?v=1YNjMxxXO-E&feature=relmfu
How to Fit A Line?

\[ y = mx + c \]
Least Squares Fit

- Standard linear solution to estimating unknowns.
  - If we know which points belong to which line
  - Or if there is only one line

\[
y = mx + c = f(x,m,c)
\]

Minimize \( E = \sum_i [y_i - f(x_i,m,c)]^2 \)

Take derivative wrt \( m \) and \( c \) set to 0
Line Fitting

\[ y = mx + c \]

\[ y_1 = mx_1 + c \]
\[ y_2 = mx_2 + c \]
\[ \vdots \]
\[ y_n = mx_n + c \]

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix}_B =
\begin{bmatrix}
  x_1 & 1 \\
  x_2 & 1 \\
  \vdots & \vdots \\
  x_n & 1
\end{bmatrix}_A
\begin{bmatrix}
  m \\
  c
\end{bmatrix}_D
\Rightarrow B = AD
\]

\[
A^T B = A^T AD
\]
\[
(A^T A)^{-1} A^T B = (A^T A)^{-1} (A^T A) D
\]
\[
D = (A^T A)^{-1} A^T B
\]

Alper Yilmaz, Mubarak Shah, Fall 2011 UCF
RANSAC: Random Sampling and Consensus
RANSAC Song

http://www.youtube.com/watch?v=1YNjMxxXOE&feature=relmfu
RANSAC: Random Sampling and Consensus

1. Randomly select two points to fit a line

2. Find the error between the estimated solution and all other points. If the error is less than tolerance, then quit, else go to step (1).
Derivation of Fundamental Matrix
Epipolar Geometry

- Defined for two static cameras
Epipolar Geometry

Epipolar line: set of world points that project to the same point in left image, when projected to right image forms a line.

Epipolar plane: plane defined by the camera centers and world point.

Epipole: intersection of image plane with line connecting camera centers. Image of a left camera center in the right, and vice versa.
Essential Matrix

Coplanarity constraint between vectors \((P_l - T), T, P_l\).

\[
\begin{align*}
(P_l - T)^T T \times P_l &= 0 \\
R^T P_r &= R(P_l - T) \\
R^T P_r &= (P_l - T)^T \\
P_r^T R &= (P_l - T)^T
\end{align*}
\]
Vector Cross Product to Matrix-vector multiplication

\[ A \times B = \begin{bmatrix} i & j & k \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{bmatrix} = \begin{bmatrix} -A_zB_y + A_yB_z \\ A_zB_x - A_xB_z \\ -B_xA_y + B_yA_x \end{bmatrix} \]

\[ A \times B = S.B = \begin{bmatrix} 0 & -A_z & A_y \\ -A_z & 0 & -A_x \\ -A_y & A_x & 0 \end{bmatrix} \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} = \begin{bmatrix} -A_zB_y + A_yB_z \\ A_zB_x - A_xB_z \\ -B_xA_y + B_yA_x \end{bmatrix} \]
Essential Matrix

\[ P_r^T R S P_l = 0 \]

\[ P_r^T E P_l = 0 \]

\( E = R \ S \)
Apply Camera model

\[ M_l^{-1} x_l = P_l \]
\[ M_r^{-1} x_r = P_r \]
\[ x_r^T M_r^{-T} = P_r^T \]

\[ x_l^T M_r^{-T} E M_l^{-1} x_l = 0 \]
\[ x_r^T M_r^{-T} E M_l^{-1} x_l = 0 \]
\[ x_r^T F x_l = 0 \]

fundamental matrix
**Fundamental Matrix**

\[
\begin{align*}
\mathbf{P}_r &= \mathbf{R}(\mathbf{P}_l - \mathbf{T}) \\
\mathbf{P}_l &= \mathbf{R}_l \mathbf{P} + \mathbf{T}_l \\
\mathbf{P}_r &= \mathbf{R}_r \mathbf{P} + \mathbf{T}_r
\end{align*}
\]  

\[
\begin{align*}
\mathbf{R} &= \mathbf{R}_r \mathbf{R}_l^\top \\
\mathbf{T} &= \mathbf{T}_l - \mathbf{R}_l^\top \mathbf{T}_r
\end{align*}
\]  

(A)

\[
\begin{align*}
\mathbf{x}_l &= \mathbf{M}_l \mathbf{P}_l \\
\mathbf{x}_r &= \mathbf{M}_r \mathbf{P}_r \\
\mathbf{P}_r^\top \mathbf{E} \mathbf{P}_l &= 0
\end{align*}
\]  

(B)

\[
\mathbf{x}_r^\top \left( \mathbf{M}_r^{-\top} \mathbf{E} \mathbf{M}_l^{-1} \right) \mathbf{x}_l = 0
\]

fundamental matrix

\[
\mathbf{x}_r^\top \mathbf{F} \mathbf{x}_l = 0
\]
**Fundamental Matrix**

Given a point $x$ in left camera, epipolar line in right camera is: $u_r = Fx$

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & m \end{bmatrix}$$
Fundamental Matrix

\[ x_l^T F x_r = x_l^T \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & m \end{pmatrix} x_r = 0 \]

- 3x3 matrix with 9 components
- Rank 2 matrix (due to $S$)
- 7 degrees of freedom
- Given a point in left camera $x$, epipolar line in right camera is: $u_r = Fx$
Fundamental Matrix

• Longuet Higgins (1981)
• Hartley (1992)
• Faugeras (1992)
• Zhang (1995)
Fundamental Matrix

- Fundamental matrix captures the relationship between the corresponding points in two views.

\[
\begin{bmatrix}
  x_i \\
y_i \\
1
\end{bmatrix}^T \begin{bmatrix}
f_{11} & f_{12} & f_{13} \\
f_{21} & f_{22} & f_{23} \\
f_{31} & f_{32} & f_{33}
\end{bmatrix} \begin{bmatrix}
x'_i \\
y'_i \\
1
\end{bmatrix} = 0,
\]

\[
\begin{bmatrix}
x_i \\
y_i \\
1
\end{bmatrix}^T \begin{bmatrix}
f_{11}x' + f_{12}y' + f_{13} \\
f_{21}x' + f_{22}y' + f_{23} \\
f_{31}x' + f_{32}y' + f_{33}
\end{bmatrix} = 0,
\]
Fundamental Matrix

\[ x_i(f_{11}x' + f_{12}y' + f_{13}) + y_i(f_{21}x' + f_{22}y' + f_{23}) + (f_{31}x' + f_{32}y' + f_{33}) = 0 \]

\[ x_i x' f_{11} + x_i y' f_{12} + x_i f_{13} + y_i x' f_{21} + x' y' f_{22} + y_i f_{23} + x' f_{31} + y_i f_{32} + f_{33} = 0 \]

One equation for one point correspondence

\[
Mf = \begin{bmatrix}
    x_1 x_1 & x_2 x_1 & \cdots & x_n x_1 & y_1 x_1 & y_2 x_1 & \cdots & y_n x_1 \\
    x_1 x_2 & x_2 x_2 & \cdots & x_n x_2 & y_1 x_2 & y_2 x_2 & \cdots & y_n x_2 \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
    x_1 x_n & x_2 x_n & \cdots & x_n x_n & y_1 x_n & y_2 x_n & \cdots & y_n x_n \\
\end{bmatrix} \begin{bmatrix}
    f_{11} \\
    f_{12} \\
    f_{13} \\
    f_{21} \\
    f_{22} \\
    f_{23} \\
    f_{31} \\
    f_{32} \\
    f_{33} \\
\end{bmatrix} = 0
\]

\( M \) is a 9 by \( n \) matrix

\( f = [f_{11} \ f_{12} \ f_{13} \ f_{21} \ f_{22} \ f_{23} \ f_{31} \ f_{32} \ f_{33}] \)

To solve the equation, the rank(\( M \)) must be 8.
Computation of Fundamental Matrix
Normalized 8-point algorithm (Hartley)

Objective:
Compute fundamental matrix $F$ such that $x_i'Fx_i = 0$

Algorithm
Normalize the image $\hat{x}_i = T x_i$, $\hat{x}_i' = T' x_i'$

Find centroid of points in each image, determine the range, and normalize all points between 0 and 1

Linear solution
determining the eigen vector corresponding to the smallest eigen value of $A$,

$$Af = \begin{bmatrix}
\hat{x}_i' \hat{x}_i & \hat{x}_i' \hat{y}_i & \hat{x}_i' & \hat{y}_i' \hat{x}_i & \hat{y}_i' & \hat{x}_i & \hat{y}_i & 1
\end{bmatrix}f = 0$$

$$A = \begin{bmatrix}
\hat{x}_s' \hat{x}_s & \hat{x}_s' \hat{y}_s & \hat{x}_s' & \hat{y}_s' \hat{x}_s & \hat{y}_s' & \hat{x}_s & \hat{y}_s & 1
\end{bmatrix}$$
Normalized 8-point algorithm (Hartley)

Construct

\[ \hat{F} = \begin{bmatrix} f_1 & f_2 & f_3 \\ f_4 & f_5 & f_6 \\ f_7 & f_8 & f_9 \end{bmatrix} \]

Normalize

\[ \hat{F} = \hat{F} / \| \hat{F} \| \]

Constraint enforcement SVD decomposition

\[ \hat{F} = U \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} V' \]

\( \sigma_1 \geq \sigma_2 \geq \sigma_3 \)

Rank enforcement

\[ \tilde{F} = U \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} V' \]

\( \sigma_3 = 0 \)

De-normalization:

\[ F = T'^T \tilde{F} T \]

L1 matrix norm is maximum of absolute column sum.

\[ \| A \|_1 = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}| \]
Robust Fundamental Matrix Estimation (by Zhang)

- Uniformly divide the image into 8×8 grid.
- Randomly select 8 grid cells and pick one pair of corresponding points from each grid cell, then use Hartley’s 8-point algorithm to compute Fundamental Matrix $F_i$.
- For each $F_i$, compute the median of the squared residuals $R_i$.
  - $R_i = \text{median}_k[d(p_{1k}, F_ip_{2k}) + d(p_{2k}, F'_ip_{1k})]$
- Select the best $F_i$ according to $R_i$.
- Determine outliers if $R_k > Th$.
- Using the remaining points compute the fundamental Matrix $F$ by weighted least square method.
Epi-polar Lines
Epi-polar lines
Figure 6.1: Stereo geometry.

\[
\frac{Z + f}{Z} = \frac{x_1 + x_2 + B}{B}, \quad Z = \frac{fB}{x_1 + x_2},
\]
Stereo Pairs and Depth Maps (from Szeliski’s book)
Image Rectification For Stereo
Correlation Based Stereo Methods

• Once disparity is available compute depth using

\[ Z = \frac{f_B}{d} \]
Correspondence using Search

Criterion function:
Correlation Based Stereo Methods

- Depth is computed only at tokens and interpolated/extrapolated to remaining pixel
- Disparity map is constructed based on a correlation measure

\[
\text{SSD} = \sum \sum (I_{\text{left}} - I_{\text{right}})^2 \\
\text{AD} = \sum \sum |(I_{\text{left}} - I_{\text{right}})| \\
\text{CC} = \sum \sum I_{t+1}I_t
\]

\[
\text{NC} = \frac{\sum \sum (I_{\text{left}}I_{\text{right}})}{\sqrt{\sum \sum I_{\text{left}}I_{\text{right}}}} \\
\text{MC} = \frac{1}{64\sigma_{\text{left}}\sigma_{\text{right}}} \sum \sum (I_{\text{left}} - \mu_{\text{right}})(I_{\text{left}} - \mu_{\text{right}})
\]
Barnard’s Stereo Method

- Similar intensity
  - Similar to brightness constraint
- Smoothness of disparity

\[
E = \sum_{i=-1}^{1} \sum_{j=-1}^{1} \left| I_{\text{left}}(x + i, y + j) - I_{\text{right}}(x + i + D_x(x, y), y + j) \right| + \lambda \| \nabla D(x, y) \|
\]

\[
\nabla D(x, y) = \sum_{i=-1}^{1} \sum_{j=-1}^{1} \left| D(x + i, y + j) - D(x, y) \right|
\]

Alper Yilmaz, Mubarak Shah, Fall 2011 UCF
Barnard’s Stereo Method

- Energy can be minimized using brute force search
  - Let max allowed disparity is 10 pixels
  - For 128x128 image for 10 possible levels of disparity
    - There $10^{16384}$ possible disparity values
  - We can select any minimization technique
    - Barnard choose simulated annealing
Simulated Annealing

- Select a random state $S$ (disparities)
- Select a high temperature
  - Select random $S'$
  - Compute $\Delta E = E(S') - E(S)$
  - If ($\Delta E < 0$) $S \leftarrow S'$
  - Else
    - $P \leftarrow \exp(-\Delta E/T)$
    - $X \leftarrow \text{random}(0,1)$
      - If $X < P$ then $S \leftarrow S'$
    - If no decrease in several iterations lower $T$
Examples

bread  toy  apple

Alper Yilmaz, Mubarak Shah, Fall 2011 UCF
Examples

Left Image
Right Image
Depth Map

Alper Yilmaz, Fall 2004 UCF
Stereo results

- Data from University of Tsukuba
- Similar results on other images without ground truth

Scene

Ground truth
Results with window correlation

Window-based matching (best window size)  Ground truth
Results with better method

State of the art method


Ground truth
Applications of Stereo (from Szeliski’s book)
Reading Material

• Fundamental of Computer Vision
  – 6.2.1, 6.2.4 and 6.2.5

• Computer Vision: Algorithms and Applications, Richard Szeliski
  – Chapter 11