Multiple View Geometry II

CS 534: Introduction to Computer Vision
Rutgers University
TA: Ali Elqursh
Spring 2010

Sources

• Medioni, G. and Kang, S. B. 2004 Emerging Topics in Computer Vision. Prentice Hall PTR.
• Slides for Trifocal Tensor by Marc Pollefeys
Agenda

- Applications
- Projective Geometry
- Structure and Motion problem
  - Two Views
  - Three Views
  - N-Views

Applications

- 3D graphical models
  - Computing the structure
- Video Augmentation
  - Must compute camera motion, no need to compute structure
    - [http://www.youtube.com/watch?v=ZKwMp5YkaE](http://www.youtube.com/watch?v=ZKwMp5YkaE)
  - You have already done it!
Applications

• Generating novel views from existing photographs

Fig. 1:2. Single view reconstruction, or Hogarth's study. The original painting is on the right, with the reconstructed views on the left. The second column shows the reconstructed views from different angles. The third column shows the original painting with additional details added.

Applications

• Panoramic mosaicing

Fig. 3:4. Panoramic mosaicing. Each image is a view of the subject, arranged in a circular pattern around the center. The images are stitched together using special software to create a seamless panoramic view.
Applications

• Measuring heights, angles

Projective Geometry

Central Perspective Transformation

• Consider mappings of points using central perspective transformation
• Points on $\Pi_r, \Pi_0$ can be represented as $\mathbb{R}^2$
• Problems
  – Points on $l$ do not map to any points on $\Pi$.
  – Points on $h$ do not correspond to any points on $\Pi_0$
• Solution
  – Append an extra line to $\Pi_r$ that are images of points on $l$ ($\mathbb{R}^2 +$ extraline)
  – Append an extra line to $\Pi_0$ that are points that are imaged at $h$ ($\mathbb{R}^2 +$ extraline)
Projective Geometry
Projective Space

• Projective Plane \( \mathbb{P}^2 \)
  
  \[ P^2 = \mathbb{R}^2 \cup \{ \text{ideal points} \} \]

  • Ideal line \( l_\infty \) or line at infinity
  
  \[ l_\infty = \{ \text{ideal points} \} \]

  • Two different points define a line
  
  • Two different line intersect in a point

• Projective Line (\( \mathbb{P}^1 \))

• Projective Space (\( \mathbb{P}^3 \))
  
  • Ideal points in \( \mathbb{P}^3 \) builds a plane
  
  • We call it \textit{ideal plane} Or \textit{plane at infinity}

---

Projective Geometry
Homogeneous coordinates

• Algebraic tool to represent entities in projective geometry

• Consider intersections of vectors through the origin and the line \( y=1 \)

• We say \((p_1, p_2)\) equivalent to \((q_1, q_2)\) if \((p_1, p_2) = \lambda (q_1, q_2)\)

• One-One mapping between equivalent vectors and intersection points

• Use vectors as representation of points on the line

• Every vectors intersects except \((1, 0)\)

• It must intersect at \((1, 0)\), ideal point

• Call these vectors “homogeneous coordinates” of points on \( l \)

\[ P^1 = \{(p_1, 1) \in P^1 \} \cup \{(p_1, 0) \in P^1 \} \]
Projective Geometry

- Generalize to $\mathbb{P}^2$
- Plane $\Pi$
- Represent each point $p$ as intersections of $(p_1, p_2, p_3)$ with $\Pi$
- $(p_1, p_2, p_3)$ equivalent to $(q_1, q_2, q_3)$ if $(p_1, p_2, p_3) = \lambda (q_1, q_2, q_3)$
- $\mathbb{P}^2 = \{(p_1, p_2, 1) \in \mathbb{P}^2\} \cup \{(p_1, p_2, 0) \in \mathbb{P}^2\}$
- Try to imagine the line at infinity

---

Projective Geometry

- Generalize to $\mathbb{P}^3$
- We get a plane at infinity
Projective Geometry

Projective Transformations

- A linear transformation in homogeneous coordinates
  \[ x' \sim Hx \]

- \( H \) does not have to be square
- If it is square mapping from \( \mathbb{P}^n \) to \( \mathbb{P}^n \), we call it a homography
- A projective transformation can move the plane at infinity!

---

Projective Geometry

Duality

- For each statement in Projective geometry there is a dual statement

  \[ l = \{ x = (x_1, x_2, x_3) \in \mathbb{P}^3 \mid x = t_1 d_1 + t_2 d_2, \ (t_1, t_2) \in \mathbb{P}^2 \} \]

  \[ l = \{ x = (x_1, x_2, x_3) \in \mathbb{P}^3 \mid m_1 x_1 + m_2 x_2 + m_3 x_3 = 0 \} \]

![Figure 3.5. Duality of points and lines in \( \mathbb{P}^3 \).](image)

- Given \( n = (n_1, n_2, n_3) \), the points \( x = (x_1, x_2, x_3) \) that fulfills (3.1) constitutes the line defined by \( n \).
- Given \( x = (x_1, x_2, x_3) \), the line \( n = (n_1, n_2, n_3) \) that fulfills (3.1) constitutes the line coincident by \( x \).
Projective Geometry
Conics

Definition 14. A conic, $c$, in $\mathbb{P}^2$ is defined as
$$c = \{x = (x_1, x_2, x_3) \in \mathbb{P}^2 \mid x^T C x = 0\},$$

Theorem 4. The dual $c^*$, to a conic $c$: $x^T C x$ is the set of lines
$$\{(l = (l_1, l_2, l_3) \in \mathbb{P}^2 \mid l^T C l = 0\},$$
where $C^* = C^{-1}$.

Fig. 22. (a) Plane $x$ satisfying $x^T C x = 0$ lie on a point conic, the lines $l$ satisfying $l^T C l = 0$ are tangent to the point conic $c$. The conic $C$ is the envelope of the lines $l$.

---

Projective Geometry
Conics

Degenerate conics. If the matrix $C$ is not of full rank, then the conic is termed degenerate. Degenerate point conics include two lines (rank 2), and a repeated line (rank 1).

Definition 18. The (singular, complex) conic, $\Pi$, in $\mathbb{P}^n$ defined by
$$x_1^2 + x_2^2 + \ldots + x_n^2 = 0 \quad \text{and} \quad x_{n+1} = 0$$
is called the absolute conic.

Observe that the absolute conic is located at the plane at infinity, it contains only complex points and it is singular.

Lemma 2. The dual to the absolute conic, denoted $\Pi'$, is given by the set of planes
$$\Omega = \{\Pi = (\Pi_1, \Pi_2, \ldots, \Pi_{n+1}) \mid \Pi_1^2 + \ldots + \Pi_n^2 = 0\}.$$

In matrix form $\Omega'$ can be written as $\Pi^T C' \Pi = 0$ with
$$C' = \begin{bmatrix} I_{n \times n} & 0_{n \times 1} \\ 0_{1 \times n} & 0 \end{bmatrix}.$$
Projective Geometry

Conics

Lemma 2. The dual to the absolute conic, denoted $\Omega'$, is given by the set of planes

$$\Omega' = \{ \Pi = (\Pi_1, \Pi_2, \ldots, \Pi_{n+1}) \mid \Pi_1^T + \ldots + \Pi_{n+1}^T = 0 \}.$$ 

In matrix form $\Omega'$ can be written as $\Pi^T C \Pi = 0$ with

$$C' = \begin{bmatrix} I_{n \times n} & 0_{n \times 1} \\ 0_{1 \times n} & 0 \end{bmatrix}.$$ 

Proposition 3. The subgroup $K$, of projective transformations, $G_P$, that preserves the absolute conic consists exactly of the projective transformations of the form (3.4), with

$$H = \begin{bmatrix} cR_{n \times n} & t_{n \times 1} \\ 0_{1 \times n} & 1 \end{bmatrix},$$

where $0 \neq c \in \mathbb{R}$ and $R$ denote an orthogonal matrix, i.e. $RR^T = R^T R = I$.

Projective Geometry

Quadrics

- In $\mathbb{P}^3$ a conic is called a quadric

Lemma 3. The projection of a quadric, $X^T CX = 0$ (dually $\Pi^T C \Pi = 0$, $C' = C^{-1}$), is an image conic, $x^T ex = 0$ (dually $\Pi e1 = 0$, $c' = c^{-1}$), with $c' = PC^T P^T$.

Proposition 6. The image of the absolute conic is given by the conic $x^T \omega x = 0$ (dually $\Pi^T \omega 1 = 0$), where $\omega' = K K^T$.

- Proof

$$\omega' \sim PT \bar{P}^T \sim KK^T [I - t] \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I \\ -t^T \end{bmatrix} = KK^T$$
Structure and Motion Problem

- Given:
  - sequence of images
  - corresponding feature points
  \[ \lambda_{ij} x_{ij} = P_i X_j \]
- Determine
  - Camera Matrices \( P_i \) (Motion)
  - 3D Points \( X_j \) (Structure)
- Important Results
  - Given un-calibrated image sequence, and without any assumptions, it is only possible to reconstruct the object up to an unknown projective transformation
  - Given calibrated camera then it is possible to reconstruct the scene up to an unknown similarity transformation
  - For some applications, projective reconstruction may be sufficient!

---

**Theorem 5.** Given an un-calibrated image sequence with corresponding points, it is only possible to reconstruct the object up to an unknown projective transformation.

**Proof:** Assume that \( X_j \) is a reconstruction of \( n \) points in \( m \) images, with camera matrices \( P_i \) according to
\[
    x_{ij} \sim P_i X_j, \quad i = 1, \ldots m, \quad j = 1, \ldots n.
\]
Then \( H X_j \) is also a reconstruction, with camera matrices \( P_i H^{-1} \), for every non-singular \( 4 \times 4 \) matrix \( H \), since
\[
    x_{ij} \sim P_i X_j \sim P_i H^{-1} H X_j \sim (P_i H^{-1})(H X_j).
\]
The transformation
\[
    X \mapsto H X
\]
corresponds to all projective transformations of the object.
In the same way it can be shown that if the cameras are calibrated, then it is possible to reconstruct the scene up to an unknown similarity transformation.
Two Views
Epipolar Geometry

C, C’, x, x’ and X are coplanar

What if only C, C’, x are known?
Two Views
Epipolar Geometry

All points on $\pi$ project on $I$ and $I'$

Two Views
Epipolar Geometry

Family of planes $\pi$ and lines $I$ and $I'$ Intersection in $e$ and $e'$
Two Views
Epipolar Geometry

epipoles $e, e'$
= intersection of baseline with image plane
= projection of projection center in other image

an epipolar plane = plane containing baseline (1-D family)
an epipolar line = intersection of epipolar plane with image
(always come in corresponding pairs)

Two Views
Example: converging cameras
Two Views
Example: motion parallel with image plane

Two Views
Example: Forward Motion
Two Views

• Epipole $e_{i,j}$ is the projections of the camera center of camera $i$ in image $j$

$$\begin{align*}
P_1 &= [A_1 \ | \ b_1] \quad \text{and} \quad P_2 = [A_2 \ | \ b_2] \\
e_{1,2} &= -A_2A_1^{-1}b_1 + b_2
\end{align*}$$

Two Views
Canonical Cameras

• Two cameras can be represented as

$$\begin{align*}
P_1 &= [A_1 \ | \ b_1] \\
P_2 &= [A_2 \ | \ b_2]
\end{align*}$$

• It will be useful to use canonical cameras. We can always convert non-canonical cameras to canonical by using the following transformation

$$H = \begin{bmatrix} A_1^{-1} & -A_1^{-1}b_1 \\ 0 & 1 \end{bmatrix}$$

$$\begin{align*}
P_1 &= P_1H = [I \ | \ 0] \\
P_2 &= P_2H = [A_2A_1^{-1} \ | \ b_2 - A_2A_1^{-1}b_1]
\end{align*}$$

Notice the epipole
Two Views
Canonical Cameras

We can multiply again by without changing $P_1$

$$\begin{bmatrix} I & 0 \\ v^T & 1 \end{bmatrix}$$

$$\begin{bmatrix} I \end{bmatrix} \begin{bmatrix} A_{12} + ev^T \end{bmatrix} = [ A_{12} + ev^T ] \begin{bmatrix} e \end{bmatrix}$$

Definition 26. A pair of camera matrices is said to be in canonical form if

$$P_1 = [ I \mid 0 ] \quad \text{and} \quad P_2 = [ A_{12} + ev^T \mid e ] \quad (3.13)$$

where $v$ denote a three-parameter ambiguity.

Two Views
Fundamental Matrix

- The fundamental matrix $F$ is an algebraic representation of the epipolar geometry

$$\lambda_1 x_1 = P_1 X = [ A_1 \mid b_1 ] X, \quad \lambda_2 x_2 = P_2 X = [ A_2 \mid b_2 ] X.$$  

$$\lambda_1 x_1 = P_1 X = [ A_1 \mid b_1 ] X = A_1 \begin{bmatrix} X^T \\ Y \\ Z \end{bmatrix} + b_1 = A_1^{-1} (\lambda_1 x_1 - b_1)$$

$$\lambda_2 x_2 = A_2 A_1^{-1} (\lambda_1 x_1 - b_1) + b_2 = \lambda_1 A_{12} x_1 + (-A_{12} b_1 - b_2)$$

$$x_2^T A_{12}^T T x_2 = x_2^T F x_2 = 0$$

- Epipolar constraint

$$x_2^T F x_2 = 0$$
Two Views
Fundamental Matrix

• Epipolar line

**Definition 28.** The line \( l = F^T x \) is called the epipolar line corresponding to \( x \).

\[
\begin{align*}
\text{Image 1} & \quad \text{Image 2} \\
C_1 & \quad C_2 \\
\mathbf{a}_1 & \quad \mathbf{b}_1 \\
\mathbf{e}_{12} & \quad \mathbf{e}_{21} \\
L & 
\end{align*}
\]

Two Views
Fundamental Matrix

• \( F e' = 0 \) (Since \( xFe' = 0 \) for all \( x \))
• \( e^T F = 0 \) (Since \( e^T Fx' = 0 \) for all \( x' \))
• \( F \) has Rank 2
• \( \det F = 0 \)
• Given \( F \) we can extract the projection matrices

\[
F = A_{12} T_e \quad \leftrightarrow \quad P_1 = [I | 0], \quad P_2 = [A_{12} | e].
\]
Two Views
Fundamental Matrix

Observe that

\[ F = A_{12}^T T_e = (A_{12} + e v^T)^T T_e \]

for every vector \( v \), since

\[ (A_{12} + e v)^T T_e(x) = A_{12}^T(e \times x) + v e^T(e \times x) = A_{12}^T T_e x , \]

since \( e^T(e \times x) = e \cdot (e \times x) = 0 \). This ambiguity corresponds to the transformation

\[ H^P_2 = [ A_{12} + e v^T ] [ e ] . \]

We conclude that there are three free parameters in the choice of the second camera matrix when the first is fixed to \( P_1 = [ I | 0 ] \).

Two Views
Fundamental Matrix

- The fundamental matrix has 9 entries
- However it is a homogeneous quantity
- Has 7 dof (9 elements - det F=0 - homogeneous)
- Can be estimated from 8 point correspondences

\[ x_i^T F x_i = 0 \]

\[ x' x f_{11} + x' y f_{12} + x' f_{13} + y' x f_{21} + y' y f_{22} + y' f_{23} + x f_{31} + y f_{32} + f_{33} = 0. \] (11.2)

\[
\begin{align*}
\mathbf{A} \mathbf{f} &= \begin{bmatrix}
x'_1 x_1 & x'_1 y_1 & 1 \\
x'_2 x_2 & x'_2 y_2 & 1 \\
\vdots & \vdots & \ddots \\
x'_n x_n & x'_n y_n & 1
\end{bmatrix} \\
\begin{bmatrix}
x'_1 y_1 \\
x'_2 y_2 \\
\vdots \\
x'_n y_n
\end{bmatrix} \\
\begin{bmatrix}
x'_1 x_1 \\
x'_2 x_2 \\
\vdots \\
x'_n x_n
\end{bmatrix} \\
\begin{bmatrix}
1 \\
1 \\
\vdots \\
1
\end{bmatrix} \mathbf{f} = 0.
\end{align*}
\]
Two Views
Fundamental Matrix

Here the convention is $x^T F x = 0$ instead of $xFx' = 0$

- $F$ is a rank 2 homogeneous matrix with 7 degrees of freedom.
- Point correspondence: If $x$ and $x'$ are corresponding image points, then $x^T F x = 0$.
- Epipolar line:
  - $F = F_0$ is the epipolar line corresponding to $x$.
  - $x - x' F^T 0$ is the epipolar line corresponding to $x'$.
- Epipole:
  - $F x = 0$.
  - $F^T x = 0$.
- Computation from camera matrices $P, P'$:
  - General camera, $F = [e]_{	imes}$, $P' = [e]_{	imes} P$, where $e^T$ is the pseudo-inverse of $e$, and $e' = P'C$, with $PC = 0$.
  - Canonical camera, $F = [I | 0]$, $P' = [I | 0] P$.
    - $F = [I | 0]$, $P = [0 | 0]$, where $e' = 0$ and $e = 0$.
  - Cameras not at infinity $F = [R | t]$, $P' = [R | t] P$.
    - $F = [R | t]$, $P = [0 | 0]$.

---

Two Views
Essential matrix

- Calibrated camera ($K$ is known)
- $P_1 = K[I | 0]$ , $P_2 = K'[R | t]$
- $F = K^{-1} R[t] x K^{-1}$
- If we use normalized points then
- $E = R[t] x$
- $x' E x = x' R[t] x x = 0$
Two Views
Retrieving Projection Matrices

E is essential matrix if and only if
two singular values are equal (and third=0)
\[ E = U \text{diag}(1, 1, 0)V^T \]
\[ W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \]

**Result 9.19.** For a given essential matrix \( E = U \text{diag}(1, 1, 0)V^T \), and first camera matrix
\( P = [1 \mid 0] \), there are four possible choices for the second camera matrix \( P' \), namely
\[ P' = [\hat{X}W^T \mid \pm \hat{u}_3] \text{ or } [\hat{X}W^T \mid -\hat{u}_3] \text{ or } [\hat{X}W^T \mid +\hat{u}_3] \text{ or } [\hat{X}W^T \mid -\hat{u}_3]. \]

These four solutions can be disambiguated. Only one will
yield reconstructed points in front of the two cameras.

---

Two Views
Structure Computation

- We know how to compute \( P' \)'s for Two views
  (Compute \( F \), then compute \( P' \)'s)
- We can define two new sub-problems
  - Resection: Assume that \( X_j' \)'s (Structure) are given,
    we have \( x_{ij} \), calculate \( P'_j \)'s (Motion)
  - Intersection: Assume that \( P'_j \)'s (Motion) are given,
    we have \( x_{ij} \), calculate \( X_j' \)'s (Structure)
Two Views
Intersection

• We have these constraints
  \[
  \begin{align*}
  \lambda_1 x_1 &= P_1 x, \\
  \lambda_2 x_2 &= P_2 x,
  \end{align*}
  \]

• In matrix form
  \[
  \begin{bmatrix}
  P_1 & 0 \\
  P_2 & 0
  \end{bmatrix}
  \begin{bmatrix}
  X \\
  -\lambda_1 \\
  -\lambda_2
  \end{bmatrix} = 0
  \]

• Can be computed linearly (6 constraints, 6 unknowns)

N-Views
Affine Factorization

• Assume cameras are affine
  – Write here how an affine camera looks like
• Compute both structure and motion at the same time
  \[
  \psi = \begin{bmatrix}
  x_1^1 & x_2^1 & \ldots & x_n^1 \\
  x_1^2 & x_2^2 & \ldots & x_n^2 \\
  \vdots & \vdots & \ddots & \vdots \\
  x_1^n & x_2^n & \ldots & x_n^n
  \end{bmatrix},
  \mu = \begin{bmatrix}
  \mu_1^1 \\
  \mu_2^1 \\
  \vdots \\
  \mu_m^1
  \end{bmatrix}
  \begin{bmatrix}
  x_1 & x_2 & \ldots & x_n
  \end{bmatrix}
  \]

• With noise \( \rightarrow \) equations not satisfied exactly
• Seek \( W_{\text{hat}} \) that is closest to \( W \) in Frobenius norm
  \[
  \| \mathbf{w} - \hat{\mathbf{w}} \|^2 = \sum_{ij} (w_{ij} - \hat{w}_{ij})^2 = \sum_{ij} \| x_j - \hat{x}_j \|^2 = \sum_{ij} \| x_j - \hat{w}_i x_i \|^2
  \]
N-Views
Affine Factorization

This can be achieved by SVD
Note $W$ must be rank 3
$W = USVT, \ W = U_{2mx3}D_{3x3}V_{3xn}^T$
Remember, we can only compute Structure and Motion up to projective transformation
One solution $M = U_{2mx3}D_{3x3}, X = V_{3xn}^T$
Another solution $M = U_{2mx3}, X = D_{3x3}V_{3xn}^T$

N-Views
Projective Factorization

- In general we need to do projective factorization
  $\begin{bmatrix}
  \lambda_1 x_1 \\
  \lambda_2 x_2 \\
  \cdots \\
  \lambda_n x_n \\
  \lambda_1^p x_1^p \\
  \lambda_2^p x_2^p \\
  \cdots \\
  \lambda_n^p x_n^p
  \end{bmatrix} = \begin{bmatrix}
  f_1^p \\
  f_2^p \\
  \cdots \\
  f_n^p \\
  [X_1, X_2, \ldots, X_n]
  \end{bmatrix}$

- Need to compute $\lambda$’s, $P$’s, and $X$’s
- Iterative process using Expectation-Maximization (EM) Algorithm
- Must choose initial values for $\lambda$’s
N-Views
Projective Factorization

Algorithm

(i) Normalize the image data using isotropic scaling as in section 4.4.4 of [107].
(ii) Start with an initial estimate of the projective depths \( \lambda_j \). This may be obtained by techniques such as an initial projective reconstruction, or else by setting all \( \lambda_j = 1 \).
(iii) Normalize the depths \( \lambda_j \) by multiplying rows and columns by constant factors. One method is to do a pass setting the norms of all rows to 1, then a similar pass on columns.
(iv) From the \( 3n \times n \) measurement matrix on the left of (18.9), find its nearest rank-4 approximation using the SVD and decompose to find the camera matrices and 3D points.
(v) Optional iteration. Reproject the points into each image to obtain new estimates of the depths and repeat from step (iii).

Algorithm 18.2. Projective reconstruction through factorization.

N-Views
Bundle Adjustment

- Nonlinear optimization
- Minimizes a geometric distance

\[
\min_{F^*, \hat{X}_j} \sum_{ij} d(p^j \hat{x}_j, x^j)^2
\]

- Can be performed very efficiently by using sparse optimization
- Requires good initialization