Rectification

Elad Osherov
Taxonomy of SfM problems

Given image correspondences:

- **Camera extrinsic and intrinsic calibration (Resectioning):**
  Calculate camera poses and calibration from known 3D points

- **Triangulation, 3D reconstruction:**
  Calculate 3D point(s) from known camera poses

- **Motion estimation:**
  Compute camera motion (up to scale). Involves multiple view geometry

- **Bundle Adjustment:**
  Compute camera poses and 3D points (up to a similarity transformation)
Epipolar geometry

- Epipolar Plane
- Epipoles
- Baseline
- Epipolar Lines

Adapted from M. Pollefeys, UNC
Epipolar geometry

- **Baseline**: line joining the camera centers
- **Epipole**: point of intersection of baseline with the image plane
- **Epipolar plane**: plane containing baseline and world point
- **Epipolar line**: intersection of epipolar plane with the image plane

Adapted from M. Pollefeys, UNC
Epipolar geometry: terms

- **Epipole**: point of intersection of baseline with the image plane
- **Epipolar plane**: plane containing baseline and world point
- **Epipolar line**: intersection of epipolar plane with the image plane

- All epipolar lines intersect at the epipole
- An epipolar plane intersects the left and right image planes in epipolar lines
• Potential matches for $p$ have to lie on the corresponding epipolar line $l'$.

• Potential matches for $p'$ have to lie on the corresponding epipolar line $l$.

http://www.ai.sri.com/~luong/research/Meta3DViewer/EpipolarGeo.html

Source: M. Pollefeys
Example
From geometry to algebra

\[ X' = RX + T \]

\[ T \times X' = T \times RX + T \times T = T \times RX \]

Normal to the plane

\[ X' \cdot (T \times X') = X' \cdot (T \times RX) = 0 \]
Matrix form of cross product

\[
\bar{a} \times \bar{b} = \begin{bmatrix}
0 & -a_z & a_y \\
 a_z & 0 & -a_x \\
-\ a_y & a_x & 0
\end{bmatrix}
\begin{bmatrix}
b_x \\
b_y \\
b_z
\end{bmatrix} = \bar{c}
\]

\[
\bar{a} \cdot \bar{c} = 0
\]

\[
\bar{b} \cdot \bar{c} = 0
\]

Can be expressed as a matrix multiplication.

\[
[a_x] = \begin{bmatrix}
0 & -a_z & a_y \\
 a_z & 0 & -a_x \\
-\ a_y & a_x & 0
\end{bmatrix}
\]

\[
\bar{a} \times \bar{b} = [a_x] \bar{b}
\]
Essential matrix

\[
X' \cdot (T \times RX) = 0
\]

\[
X' \cdot (T_x \ RX) = 0
\]

Let \( E = T_x R \)

\[
X'^T EX = 0
\]

This holds for the rays \( p \) and \( p' \) that are parallel to the camera-centered position vectors \( X \) and \( X' \), so we have:

\[
p'^T E p = 0
\]

\( E \) is called the essential matrix, which relates corresponding image points [Longuet-Higgins 1981].
Essential matrix and epipolar lines

\[
p' E^T p = 0
\]

Epipolar constraint: if we observe point \( p \) in one image, then its position \( p' \) in second image must satisfy this equation.

\( E \) is the coordinate vector representing the epipolar line associated with point \( p \)

\( E^T p' \) is the coordinate vector representing the epipolar line associated with point \( p' \)
Essential matrix: properties

- Relates image of corresponding points in both cameras, given rotation and translation
- Assuming intrinsic parameters are known

\[ E \triangleq T_x R \]
Stereo image rectification

In practice, it is convenient if image scanlines are the epipolar lines.

- reproject image planes onto a common plane parallel to the line between optical centers
- pixel motion is horizontal after this transformation two homographies (3x3 transforms), one for each input image reprojection

Stereo image rectification: example

Source: Alyosha Efros
Stereo image rectification: why?

- With multiple cameras it can be difficult to find a corresponding point viewed by one camera in the image of the other camera (known as the correspondence problem).
- In most camera configurations, finding correspondences requires a search in two-dimensions. However, if the two cameras are aligned correctly to be coplanar, the search is simplified to one dimension - a horizontal line parallel to the line between the cameras!

- Rectified images satisfy the following:
  - All epipolar lines are parallel to the horizontal axis.
  - Corresponding points have identical vertical coordinates.
Calibrated Rectification - How?

- “A compact Algorithm For Rectification Of Stereo Pairs” [Fusiello et al (2000)]
- 22 Matlab lines
Recap: Camera parameters

- **Extrinsic**: location and orientation of camera frame with respect to reference frame
- **Intrinsic**: how to map pixel coordinates to image plane coordinates
**Recap: Camera parameters**

- **Extrinsic**: location and orientation of camera frame with respect to reference frame
- **Intrinsic**: how to map pixel coordinates to image plane coordinates
Perspective Projection Matrix

\[
\begin{bmatrix}
\alpha_u & \gamma & u_0 \\
0 & \alpha_v & v_0 \\
0 & 0 & 1
\end{bmatrix}
\]

\[
M_{ext} = (R | t)
\]

\[
\tilde{m} \cong \tilde{P}\tilde{w}
\]

\[
\tilde{P} = M_{int} M_{ext}
\]

\[
\alpha_u = -fk_u \quad \alpha_v = -fk_v
\]

- \(\alpha_u\) and \(\alpha_v\) are the focal length in horizontal and vertical axes in pixels
- \(f\) is the focal length in mm
- \(k_u\) and \(k_v\) are the effective number of pixels per mm along \(u\) and \(v\) axes
- \(u_0\) and \(v_0\) are the coordinates of the principal point
- Assume the skew is 0
“A compact Algorithm For Rectification Of Stereo Pairs”

- Projection matrix can be written as
  \[ \tilde{P} = [Q | \tilde{q}] = \begin{bmatrix} q_1^T & q_{14} \\ q_2^T & q_{24} \\ q_3^T & q_{34} \end{bmatrix} \]

- For optical center:
  \[ 0 \succeq \tilde{P}c \succeq QC_{3x1} + \begin{bmatrix} q_{14} \\ q_{24} \\ q_{34} \end{bmatrix} \rightarrow c = -Q^{-1}\tilde{q} \]

- \(c\) are the optical center coordinates

- Projection matrix can be rewritten as:
  \[ \tilde{P} = [Q | -Qc] \]
“A compact Algorithm For Rectification Of Stereo Pairs”

- So if $\tilde{P} = [Q \mid -Qc]$.
- The **optical ray** (The line between image point $m$ and $C$) is:
  \[ w = c + \lambda Q^{-1}\tilde{m} \quad \lambda \in \mathbb{R} \]
- The **focal plane** is the plane that contains $C$ which is parallel to the image plane.
- Rectification – Both epipoles are at infinity, that happens when the line $C_1C_2$ (baseline) is contained in both focal planes.
“A compact Algorithm For Rectification Of Stereo Pairs”

- **Idea**: define two new PPMs, $\tilde{P}_n$ & $\tilde{P}_n'$, by rotating the old ones around their optical centers until focal planes become coplanar.

- **Results:**
  - Epipoles in infinity
  - Epipolar lines are parallel.
  - For horizontal lines, the baseline must be parallel to the new axis of both cameras.
  - Conjugate points must have the same vertical coordinates.
“A compact Algorithm For Rectification Of Stereo Pairs”
“A compact Algorithm For Rectification Of Stereo Pairs”

- Assume that the stereo rig is calibrated – we know \( \tilde{P}_{o1} \) & \( \tilde{P}_{o2} \)
- Positions of old optical centers stay the same, orientation of focal plane changes
- So the new PPM’s are \( \tilde{P}_{n1} \) & \( \tilde{P}_{n2} \)

\[
\tilde{P}_{n1} = A\begin{bmatrix} R & -Rc_1 \end{bmatrix} \quad \tilde{P}_{n2} = A\begin{bmatrix} R & -Rc_2 \end{bmatrix}
\]

- The intrinsic parameters are the same (can choose arbitrarily)
- \( R \) is the same for both PPM’s. Define it by the rows:

\[
R = \begin{bmatrix}
r_1^T \\
r_2^T \\
r_3^T
\end{bmatrix}
\]

- These are the X,Y,Z axes of the camera reference frame in world coordinates
According to the previous definitions we take:

1. The new X axis parallel to the baseline:
   \[ r_1 = \frac{(c_1 - c_2)}{||c_1 - c_2||_2} \]

2. The new Y axis orthogonal to X and to k:
   \[ r_2 = k \land r_1 \]

3. The new Z axis orthogonal to XY:
   \[ r_3 = r_1 \land r_2 \]

4. k is arbitrary unit vector that fixes the position of new Y in the plane orthogonal to X. We take it equal to the Z unit vector of the old left matrix, thereby constraining the new Y to be orthogonal to both X and Z (in the old left coordinates).
In order to rectify (Let’s say the left image) we need to compute the transformation mapping the old image plane \( P_{o1} \approx [Q_{o1} \mid \tilde{q}_{o1}] \) to the new image plane \( P_{n1} \approx [Q_{n1} \mid \tilde{q}_{n1}] \).

We will see that’s the transformation is the collinearity given by the 3x3 matrix \( T_{1} = Q_{n1}Q^{-1}_{o1} \).

Why? For any 3D point \( w \) we can write:

\[
\begin{align*}
\tilde{m}_{o1} & \approx \tilde{P}_{o1}\tilde{w} \\
\tilde{m}_{n1} & \approx \tilde{P}_{n1}\tilde{w}
\end{align*}
\]

The optical rays are:

\[
\begin{align*}
w & = c_{1} + \lambda_{1}Q_{o1}^{-1}\tilde{m}_{o1} & \lambda_{1} & \in \mathbb{R} \\
w & = c_{2} + \lambda_{2}Q_{n1}^{-1}\tilde{m}_{n1} & \lambda_{2} & \in \mathbb{R}
\end{align*}
\]

So \( \tilde{m}_{n1} = \lambda Q_{n1}Q_{o1}^{-1}\tilde{m}_{o1} \quad \lambda \in \mathbb{R} \).
% factorize old PPM
[A1,R1,t1] = art(Po1);
[A2,R2,t2] = art(Po2);

% optical centers (unchanged)
c1 = -R1'*inv(A1)*Po1(:,4);
c2 = -R2'*inv(A2)*Po2(:,4);

% new x axis (baseline, from c1 to c2)
v1 = (c2-c1);
% new y axes (orthogonal to old z and new x)
v2 = cross(R1(3,:)',v1);
% new z axes (no choice, orthogonal to baseline and y)
v3 = cross(v1,v2);

% new extrinsic (translation unchanged)
R = [v1'/norm(v1)
v2'/norm(v2)
v3'/norm(v3)];

% new intrinsic (arbitrary)
An1 = A2;
An1(1,2)=0; %no skew
An2 = A2;
An2(1,2)=0;

% new projection matrices
Pn1 = An1 * [R, -R*c1 ];
Pn2 = An2 * [R, -R*c2 ];

% rectifying image transformation
T1 = Pn1(1:3,1:3)* inv(Po1(1:3,1:3));
T2 = Pn2(1:3,1:3)* inv(Po2(1:3,1:3));
function [A,R,t] = art(P)
% ART: factorize a PPM as
P=A*[R;t]
Q = inv(P(1:3, 1:3));
[U,B] = qr(Q);
R = inv(U);
t = B*P(1:3,4);
A = inv(B);
A = A ./A(3,3)
Uncalibrated case

- What if we don’t know the camera parameters?
Uncalibrated case

For a given camera:

\[ \bar{p} = M_{\text{int}} p \]

So, for two cameras (left and right):

\[ p_{(\text{left})} = M_{\text{left}, \text{int}}^{-1} \bar{p}_{(\text{left})} \]

\[ p_{(\text{right})} = M_{\text{right}, \text{int}}^{-1} \bar{p}_{(\text{right})} \]
Uncalibrated case: fundamental matrix

\[
\begin{align*}
\mathbf{p}_{\text{left}} &= \mathbf{M}_{\text{left, int}}^{-1} \bar{\mathbf{p}}_{\text{left}} \\
\mathbf{p}_{\text{right}} &= \mathbf{M}_{\text{right, int}}^{-1} \bar{\mathbf{p}}_{\text{right}} \\
\end{align*}
\]

\[
\begin{align*}
\mathbf{p}_{\text{right}}^T \mathbf{E} \mathbf{p}_{\text{left}} = 0
\end{align*}
\]

From before, the essential matrix \( \mathbf{E} \).

\[
\tilde{\mathbf{P}} = [\mathbf{Q} \mid -\mathbf{Qc}]
\]

\[
\begin{align*}
\left( \mathbf{M}_{\text{right, int}}^{-1} \bar{\mathbf{p}}_{\text{right}} \right)^T \mathbf{E} \left( \mathbf{M}_{\text{left, int}}^{-1} \bar{\mathbf{p}}_{\text{left}} \right) = 0
\end{align*}
\]
Uncalibrated case: fundamental matrix

\[ \tilde{P} = [Q \mid -Qc] \]

\[ \left( M^{-1}_{\text{right, int}} \bar{p}_{\text{right}} \right)^T E \left( M^{-1}_{\text{left, int}} \bar{p}_{\text{left}} \right) = 0 \]

\[ \bar{p}_{\text{right}}^T \left( M^{-T}_{\text{right, int}} E M^{-1}_{\text{left, int}} \right) \bar{p}_{\text{left}} = 0 \]

\[ \bar{p}_{\text{right}}^T F \bar{p}_{\text{left}} = 0 \]

Fundamental matrix
Fundamental matrix

- Relates pixel coordinates in the two views
- More general form than essential matrix: we remove the need to know intrinsic parameters
- If we estimate fundamental matrix from correspondences in pixel coordinates, can reconstruct epipolar geometry without intrinsic or extrinsic parameters
Uncalibrated Rectification - How?

- “Quasi-Euclidean Uncalibrated Epipolar Rectification” [Fusiello et al (2008)]
**Idea:** Find two rectifying homographies.

**Homographies:**

\[
H_r = P_{nr1:3} P_{or1:3}^{-1} = K_{nr} R_r K_{or}^{-1}
\]

A fundamental matrix of a rectified pair has the form of:

\[
F_r = \begin{bmatrix}
    0 & 0 & 0 \\
    0 & 0 & -1 \\
    0 & 1 & 0
\end{bmatrix} = [u_1]_x
\]

\[u_1 = (1, 0, 0)\]
“Quasi-Euclidean Uncalibrated Epipolar Rectification”

- Therefore:
- Instead of algebraic error – Approximation to geometric error (Sampson error):

\[
\left( H_r m_r^j \right)^T [u_1] \times \left( H_l m_l^j \right) = 0
\]

\[
F = H_r^T [u_1] \times H_l
\]

\[
E^j_s = \frac{(m_r^j)^T F m_r^j)^2}{||[u_3] \times F m_r^j||^2 + ||m_r^j)^T F [u_3] \times||^2}
\]

- Solve \( \{E_s\}^j = 0 \) with least squares (SVD…)

\[
u_3 = (0, 0, 1)
\]
“Quasi-Euclidean Uncalibrated Epipolar Rectification” - Assumptions

- \( K_{nl}, K_{nr} \) Set arbitrarily: \( K_{nl} = K_{nr} = K_{ol} \)
- Original calibration matrices:
  \[
  K_{ol} = K_{or} = \begin{bmatrix}
  \alpha & 0 & w/2 \\
  0 & \alpha & h/2 \\
  0 & 0 & 1 
  \end{bmatrix}
  \]
- \( X \) rotation of one image assumed 0.
- Total 6 unknowns: 5 angles and focal length
“Quasi-Euclidean Uncalibrated Epipolar Rectification” - Results
“Quasi-Euclidean Uncalibrated Epipolar Rectification” - Results

Correct correspondences according to RANSAC
“Quasi-Euclidean Uncalibrated Epipolar Rectification” - Results

Left tie points (trapezoid)  Right tie points (square)
“Quasi-Euclidean Uncalibrated Epipolar Rectification” - Results
“Quasi-Euclidean Uncalibrated Epipolar Rectification” - Results
Image based motion estimation

- **Input:** 2 images, n image correspondences, camera calibration
- **Required:** Camera motion $R, t$ (up to scale)

Process:
- Estimate essential matrix $E$ from image correspondences
- Recover camera motion from estimated $E$

Solutions:
- (normalized) 8-point algorithm
- 6-point algorithm
- 5-point algorithm
Image based motion estimation via 8-Point Algorithm

- **Input**: 2 images, camera calibration
- **Required**: Camera motion $R, t$ (up to scale)

- **Steps**:
  - Calculate image correspondences (previous lecture)
  - Compute essential matrix
  - Extract camera motion
Fundamental Matrix Song
Fundamental Matrix Song

After this song, there will be a quiz. Listen carefully!
Exercise

What is the difference between Fundamental Matrix and Homography?  
(Both of them are to explain 2D point to 2D point correspondences.)
Fundamental matrix - Properties

- It can be shown that $\forall x : \ l' = Fx$

- Since $e'$ lies on $l'$: $\ e'^Tl' = 0$

- Hence: $\forall x : \ e'^TFx = 0$
  - $e'$ is the left null-space of $F$
  - Similarly, $e$ is the right null-space of $F$

- $\det(F) = 0$ - Singularity constraint

- $F$ has 7 degrees of freedom (DOFs):
  - a $3\times3$ homogenous matrix has 8 DOFs (why?)
  - The singularity constraint removes an additional DOF
Fundamental matrix - Estimation

- Each image correspondence $x \leftrightarrow x'$ contributes a single equation, linear in the unknown entries of $F$:

  $$x'^T F x = 0$$

- Recalling

  $$x = (x, y, 1)^T$$
  $$x' = (x', y', 1)^T$$

  we get:

  $$x' x f_{11} + x' y f_{12} + x' f_{13} + y' x f_{21} + y' y f_{22} + y' f_{23} + x f_{31} + y f_{32} + f_{33} = 0.$$
Fundamental matrix - Estimation

Given sufficiently many image correspondences \( x_i \leftrightarrow x_i' \) we can compute \( F \)

Single correspondence:
\[
(x'x, x'y, x', y'x, y'y, y', x, y, 1) f = 0
\]

For \( n \) image correspondences:
\[
Af = \begin{bmatrix}
x'_1x_1 & x'_1y_1 & x'_1 & y'_1x_1 & y'_1y_1 & y'_1 & x_1 & y_1 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x'_nx_n & x'_ny_n & x'_n & y'_nx_n & y'_ny_n & y'_n & x_n & y_n & 1
\end{bmatrix} f = 0
\]

What is minimal \( n \)?
Fundamental matrix - Estimation

\[ \mathbf{F} = \begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix} \mathbf{f} = 0 \]

- Need to enforce singularity constraint \( \det(F') = 0 \)

Recall
All epipolar lines should intersect at the epipole.
Fundamental matrix - Estimation

- Need to enforce singularity constraint $\det(F) = 0$
  - Most convenient: Correct F that was estimated via SVD over A
  - How:
    - Replace F by $F'$ that minimizes
      $$||F - F'|| \quad \text{subject to} \quad \det(F') = 0$$
    - via SVD on F

- To reduce sensitivity to noise:
  - Normalize all image points
  - Results in the “Normalized 8-point algorithm”

- Further reading
  - H&Z MVG book, Chapter 11.1
  - Nister’s 5-point algorithm
Fundamental matrix - Estimation

- Given camera calibration, E can be calculated from F
- Recall $E = [t]_\times \; R \in \mathbb{R}^{3 \times 3}$
- Therefore: $t^T E = 0$
  - i.e. $t^T$ is the nullspace of E
- Extract $t^T$ via svd (only up to scale, why?)
  - Left singular value corresponding to zero singular value
- Extract R
Slide Credits

- Trevor Darrell
- Kristen Grauman for most,
- Rick Szeliski and others as noted…
- Vadim Indelman