Lecture 7

Multi-view geometry

• The SFM problem
• Affine SFM
• Perspective SFM
• Self-calibration
• Applications

Reading:

[HZ] Chapter 10 “3D reconstruction of cameras and structure”
Chapter 18 “N-view computational methods”
Chapter 19 “Auto-calibration”

[FP] Chapter 13 “projective structure from motion”

[Szelisky] Chapter 7 “Structure from motion”
Affine structure from motion
(simpler problem)

From the $m \times n$ observations $x_{ij}$, estimate:
• $m$ projection matrices $M_i$ (affine cameras)
• $n$ 3D points $X_j$
Affine structure from motion
(simpler problem)

Image 1

World point $X_j$

Image $i$

For the affine case (in Euclidean space)

$$x_{ij} = A_i X_j + b_i$$

[Eq. 4]
The Affine Structure-from-Motion Problem

Two approaches:

- Algebraic approach (affine epipolar geometry; estimate F; cameras; points)

- Factorization method
A factorization method – Tomasi & Kanade algorithm

C. Tomasi and T. Kanade

• Data centering
• Factorization
A factorization method - Centering the data

Centering: subtract the centroid of the image points

\[ \hat{x}_{ij} = x_{ij} - \frac{1}{n} \sum_{k=1}^{n} x_{ik} \]

\[ \bar{x}_i = \frac{1}{n} \sum_{k=1}^{n} x_{ik} \]
A factorization method - Centering the data

Centering: subtract the centroid of the image points

[Eq. 6] \[ \hat{x}_{ij} = x_{ij} - \frac{1}{n} \sum_{k=1}^{n} x_{ik} = A_i X_j + b_i - \frac{1}{n} \sum_{k=1}^{n} A_i X_k - \frac{1}{n} \sum_{k=1}^{n} b_i \]

\[ x_{ik} = A_i X_k + b_i \]  

[Eq. 4]

[Eq. 5] \[ \bar{x}_i = \frac{1}{n} \sum_{k=1}^{n} x_{ik} \]
A factorization method  - Centering the data

Centering: subtract the centroid of the image points

\[ \hat{x}_{ij} = x_{ij} - \frac{1}{n} \sum_{k=1}^{n} x_{ik} = A_i X_j + b_i - \frac{1}{n} \sum_{k=1}^{n} A_i X_k - \frac{1}{n} \sum_{k=1}^{n} b_i \]

\[ x_{ik} = A_i X_k + b_i \]

\[ \bar{X}_i = \frac{1}{n} \sum_{k=1}^{n} x_{ik} \]

Centroid of 3D points

\[ \bar{X} = \frac{1}{n} \sum_{k=1}^{n} X_k \]  

[Eq. 4]

[Eq. 6]

[Eq. 7]

[Eq. 8]
A factorization method - Centering the data

Thus, after centering, each \textbf{normalized} observed point is related to the 3D point by

\[
\hat{X}_{ij} = A_i \hat{X}_j \quad [\text{Eq. 8}]
\]

\[
\bar{X}_i = \frac{1}{n} \sum_{k=1}^{n} X_{ik} \quad [\text{Eq. 7}]
\]

Centroid of 3D points
A factorization method - Centering the data

If the centroid of points in 3D = center of the world reference system

\[ \hat{X}_{ij} = A_i \hat{X}_j = A_i X_j \]  
\[ \text{Eq. 9} \]

\[ \bar{X}_i = \frac{1}{n} \sum_{k=1}^{n} x_{ik} \]
\[ \text{Eq. 7} \]

Centroid of 3D points
A factorization method - factorization

Let’s create a $2m \times n$ data (measurement) matrix:

$$
D = \begin{bmatrix}
\hat{x}_{11} & \hat{x}_{12} & \cdots & \hat{x}_{1n} \\
\hat{x}_{21} & \hat{x}_{22} & \cdots & \hat{x}_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\hat{x}_{m1} & \hat{x}_{m2} & \cdots & \hat{x}_{mn}
\end{bmatrix}
$$

Each $\hat{x}_{ij}$ entry is a 2x1 vector!
A factorization method - factorization

Let’s create a $2m \times n$ data (measurement) matrix:

$$D = \begin{bmatrix}
\hat{X}_{11} & \hat{X}_{12} & \cdots & \hat{X}_{1n} \\
\hat{X}_{21} & \hat{X}_{22} & \cdots & \hat{X}_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
\hat{X}_{m1} & \hat{X}_{m2} & \cdots & \hat{X}_{mn}
\end{bmatrix} = \begin{bmatrix}
A_1 \\
A_2 \\
\vdots \\
A_m
\end{bmatrix}\begin{bmatrix}
X_1 \\
X_2 \\
\vdots \\
X_n
\end{bmatrix}$$

(points $3 \times n$)

cameras $(2m \times 3)$

$$D = MS$$

[Eq. 10]

Each $\hat{X}_{ij}$ entry is a 2x1 vector!
$A_i$ is 2x3 and $X_j$ is 3x1

The measurement matrix $D = MS$ has rank 3
(it’s a product of a 2mx3 matrix and 3xn matrix)
Factorizing the Measurement Matrix

\[ D = MS \]
Factorizing the Measurement Matrix

- How to factorize D? By computing the Singular value decomposition of D!

\[ D = U \times W \times V^T \]
Since rank (D)=3, there are only 3 non-zero singular values $\sigma_1$, $\sigma_2$ and $\sigma_3$

Factorizing the Measurement Matrix

Where $W_3 = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix}$ [Eq. 11]
Factorizing the Measurement Matrix

\[ D = U_3 \times W_3 \times V_3^T \]
Factorizing the Measurement Matrix

\[ D = U_3 W_3 V_3^T = U_3 (W_3 V_3^T) = M S \quad [\text{Eq. 12}] \]
Factorizing the Measurement Matrix

\[ D = U_3 W_3 V_3^T = U_3 (W_3 V_3^T) = M S \]  \hspace{1cm} [Eq. 12]

What is the issue here? \( D \) has rank > 3 because of:

- measurement noise
- affine approximation

**Theorem:** When \( D \) has a rank greater than 3, \( U_3 W_3 V_3^T \) is the best possible rank-3 approximation of \( D \) in the sense of the Frobenius norm.

\[ D = U_3 W_3 V_3^T \begin{cases} M \approx U_3 \\ S \approx W_3 V_3^T \end{cases} \]

\[ \|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2} \]
Reconstruction results

Affine Ambiguity

\[ D = M S \]
Affine Ambiguity

The decomposition is not unique. We get the same \( \mathbf{D} \) by applying the transformations:

\[
\mathbf{M}^* = \mathbf{M} \mathbf{H}
\]

\[
\mathbf{S}^* = \mathbf{H}^{-1} \mathbf{S}
\]

where \( \mathbf{H} \) is an arbitrary 3x3 matrix describing an affine transformation.

Additional constraints must be enforced to resolve this ambiguity.
Affine Ambiguity

$$A^* = AH$$

$$S^* = H^{-1}S$$

$$A'^* = A'H$$
Similarity Ambiguity

- The scene is determined by the images only up a similarity transformation (rotation, translation and scaling)

- This is called **metric reconstruction**

- The ambiguity exists even for (intrinsically) calibrated cameras
- For calibrated cameras, the similarity ambiguity is the **only** ambiguity

[Longuet-Higgins '81]
• It is impossible, based on the images alone, to estimate the absolute scale of the scene
While calibrating a camera, we make assumptions about the geometry of the world.
Lecture 7

Multi-view geometry

• The SFM problem
• Affine SFM
• Perspective SFM
• Self-calibration
• Applications
Structure from motion problem

From the m×n observations \( x_{ij} \), estimate:
- \( m \) projection matrices \( M_i \) = motion
- \( n \) 3D points \( X_j \) = structure
Structure from motion problem

$m$ cameras $M_1... M_m$

$$M_i = \begin{bmatrix}
a_{11} & a_{12} & a_{13} & b_1 \\
a_{21} & a_{22} & a_{23} & b_2 \\
a_{31} & a_{32} & a_{33} & 1
\end{bmatrix}$$
In the general case (nothing is known) the ambiguity is expressed by an arbitrary 4x4 projective transformation.

\[
\begin{align*}
x_j &= M_i X_j \\
M_i &= K_i [R_i \quad T_i] \\
M_j H^{-1} \\
x_j &= M_i X_j = (M_i H^{-1})(H X_j)
\end{align*}
\]
The Structure-from-Motion Problem

Given \( m \) images of \( n \) fixed points \( X_j \) we can write

\[
x_{ij} = M_i X_j \quad \text{for } i = 1, \ldots, m \text{ and } j = 1, \ldots, n
\]

Problem: estimate \( m \) 3×4 matrices \( M_i \) and \( n \) positions \( X_j \) from \( m \times n \) observations \( x_{ij} \).

- If the cameras are not calibrated, cameras and points can only be recovered up to a 4x4 projective (where the 4x4 projective is defined up to scale)
- Given two cameras, how many points are needed?
- How many equations and how many unknown?

\( 2m \times n \) equations in \( 11m+3n - 15 \) unknowns
Projective Ambiguity

\[ S = \]

The problem of recovering the metric reconstruction from the perspective one is called **self-calibration**.
Structure-from-Motion methods

1. Recovering structure and motion up to perspective ambiguity
   - Algebraic approach (by fundamental matrix)
   - Factorization method (by SVD)
   - Bundle adjustment

2. Resolving the perspective ambiguity
Algebraic approach (2-view case)

1. Compute the fundamental matrix F from two views
2. Use F to estimate projective cameras
3. Use these cameras to triangulate and estimate points in 3D
Algebraic approach (2-view case)

From at least 8 point correspondences, compute $F$ associated to camera 1 and 2

$x_{1,j} = M_1 X_j$
$x_{2,j} = M_2 X_j$

For $j = 1, \ldots, n$

N. of points
1. Compute the fundamental matrix $F$ from two views (eg. 8 point algorithm)
2. Use $F$ to estimate projective cameras
3. Use these cameras to triangulate and estimate points in 3D
Algebraic approach (2-view case)

Because of the projective ambiguity, we can always apply a projective transformation $H$ such that:

$$ M_1 H^{-1} = \begin{bmatrix} I & 0 \end{bmatrix} \quad \text{Canonical perspective camera} $$

$$ M_2 H^{-1} = \begin{bmatrix} A & b \end{bmatrix} \quad \text{[Eq. 4]} $$

For $j = 1, \ldots, n$

$$ x_{1j} = M_1 X_j $$

$$ x_{2j} = M_2 X_j $$

N. of points
Algebraic approach (2-view case)

- Call $\mathbf{X}$ a generic 3D point $\mathbf{X}_{ij}$
- Call $\mathbf{x}$ and $\mathbf{x}'$ the corresponding observations to camera 1 and respectively

\[\begin{align*}
\tilde{M}_1 &= M_1 H^{-1} = \begin{bmatrix} I & 0 \end{bmatrix} \\
\tilde{M}_2 &= M_2 H^{-1} = \begin{bmatrix} A & b \end{bmatrix} \\
\tilde{X} &= H X
\end{align*}\]

\[\begin{align*}
\mathbf{x}' &= [A|b] \tilde{X} = [A|b] \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \\ \tilde{X}_3 \\ 1 \end{bmatrix} = A[I|0] \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \\ \tilde{X}_3 \\ 1 \end{bmatrix} + b = A[I|0] \tilde{X} + b = Ax + b \\
\mathbf{x}' \times b &= (Ax + b) \times b = Ax \times b \\
\mathbf{x}'^T \cdot (\mathbf{x}' \times b) &= \mathbf{x}'^T \cdot (Ax \times b) = 0 \\
\mathbf{x}'^T (b \times Ax) &= 0
\end{align*}\]
Cross product as matrix multiplication

\[
a \times b = \begin{bmatrix}
0 & -a_z & a_y \\
a_z & 0 & -a_x \\
-a_y & a_x & 0
\end{bmatrix}
\begin{bmatrix}
b_x \\
b_y \\
b_z
\end{bmatrix} = \begin{bmatrix} a_x \end{bmatrix} b
\]
Algebraic approach (2-view case)

\[
\begin{align*}
\tilde{M}_1 &= M_1 H^{-1} = \begin{bmatrix} I & 0 \end{bmatrix} \\
\tilde{M}_2 &= M_2 H^{-1} = \begin{bmatrix} A & b \end{bmatrix} \\
\tilde{X} &= H X
\end{align*}
\]

\[
\begin{align*}
x &= M_1 H^{-1} H X = [I \mid 0] \tilde{X} \\
x' &= M_2 H^{-1} H X = [A \mid b] \tilde{X}
\end{align*}
\]

\[
x'^T (b \times A x) = 0 \quad \text{[Eq. 10]}
\]

\[
x'^T [b_x] A x = 0 \quad \text{is this familiar?}
\]

\[
F = [b_x] A
\]

\[
x'^T F x = 0
\]

fundamental matrix!
Compute cameras

\[ x'^T F x = 0 \quad F = [b_x]A = b \times A \quad \text{[Eq. 11]} \]

Compute \( b \):

- Let’s consider the product \( F b \)

\[ F \cdot b = [b_x]A \cdot b = b \times A \cdot b = 0 \quad \text{[Eq. 12]} \]

- Since \( F \) is singular, we can compute \( b \) as least sq. solution of \( F b = 0 \), with \( |b| = 1 \) using SVD

- Using a similar derivation, we have that \( b^T F = 0 \) \quad \text{[Eq. 12-bis]}
Compute cameras

\[ x'^T F x = 0 \quad \text{F} = [b_x] A \]

[Eq. 11]

Compute A:

- Define: \( A' = -[b_x] F \)

- Let’s verify that \([b_x] A' \) is equal to \( F \):

Indeed:

\[
[b_x] A' = -[b_x] [b_x] F = -(b b^T - |b|^2 I) F = -b b^T F + |b|^2 F = 0 + 1 \cdot F = F
\]

[Eq. 13]

- Thus, \( A = A' = -[b_x] F \)

\[
\tilde{M}_1 = \begin{bmatrix} I & 0 \end{bmatrix} \quad \tilde{M}_2 = \begin{bmatrix} -[b_x] F & b \end{bmatrix}
\]

[Eqs. 14]
Interpretation of $b$

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

$$F = [b_x] A$$

[Eq. 11]

$$F b = 0 \quad [Eq. 12]$$

$$b^T F = 0 \quad [Eq. 12\text{-bis}]$$

What’s $b$??
Epipolar Constraint [lecture 5]

F \times_2 \text{ is the epipolar line associated with } x_2 (l_1 = F \times_2)

F^T \times_1 \text{ is the epipolar line associated with } x_1 (l_2 = F^T \times_1)

F \text{ is singular (rank two)}

\boxed{F e_2 = 0 \text{ and } F^T e_1 = 0}

F \text{ is 3x3 matrix; 7 DOF}
Interpretation of $\mathbf{b}$

\[ \mathbf{x}^T \mathbf{F} \mathbf{x} = 0 \quad \mathbf{F} = [\mathbf{b}_x] \mathbf{A} \]

\[
\begin{align*}
\mathbf{F} \mathbf{b} &= 0 \\
\mathbf{b}^T \mathbf{F} &= 0
\end{align*}
\]

[Eq. 11]

\[ \mathbf{b} \text{ is an epipole!} \]

\[ \begin{align*}
\tilde{\mathbf{M}}_1 &= \begin{bmatrix}
I & 0 \\
\downarrow & \downarrow
\end{bmatrix} \\
\tilde{\mathbf{M}}_2 &= \begin{bmatrix}
- & [\mathbf{b}_x] \mathbf{F} & \mathbf{b} \\
\downarrow & \downarrow
\end{bmatrix}
\end{align*} \]

[Eq. 15]

\[ \begin{align*}
\tilde{\mathbf{M}}_1 &= \begin{bmatrix}
I & 0 \\
\downarrow & \downarrow
\end{bmatrix} \\
\tilde{\mathbf{M}}_2 &= \begin{bmatrix}
- & [\mathbf{e}_x] \mathbf{F} & \mathbf{e} \\
\downarrow & \downarrow
\end{bmatrix}
\end{align*} \]

[Eq. 16]
Algebraic approach (2-view case)

1. Compute the fundamental matrix $F$ from two views (eg. 8 point algorithm)
2. Use $F$ to estimate projective cameras
3. Use these cameras to triangulate and estimate points in 3D
Triangulation

\[ x_{1j} = \tilde{M}_2 \tilde{X}_j \]

\[ x_{2j} = \tilde{M}_1 \tilde{X}_j \]

\[ \tilde{M}_1 = \begin{bmatrix} I & 0 \end{bmatrix} \]

\[ \tilde{M}_2 = \begin{bmatrix} -[e_x]F & e \end{bmatrix} \]

\[ \tilde{X}_j \quad \text{For } j = 1, \ldots, n \]

3D points can be computed from camera matrices via SVD (see page 312 of HZ for details)
Algebraic approach: the N-views case

- From $I_k$ and $I_h$ $\Rightarrow \tilde{M}_k, \tilde{M}_h, \tilde{X}_{[k,h]}$

- Pairwise solutions may be combined together using bundle adjustment
Structure-from-Motion Algorithms

- Algebraic approach (by fundamental matrix)
- Factorization method (by SVD)
- Bundle adjustment
Limitations of the approaches so far

• Factorization methods assume all points are visible. This is not true if:
  • occlusions occur
  • failure in establishing correspondences

• Algebraic methods work with 2 views
Bundle adjustment

- Non-linear method for refining structure and motion
- Minimizes re-projection error

\[ E(M, X) = \sum_{i=1}^{m} \sum_{j=1}^{n} D(x_{ij}, M_i X_j)^2 \]
General Calibration Problem

\[ E(M, X) = \sum_{i=1}^{m} \sum_{j=1}^{n} D(x_{ij}, M_i X_j)^2 \]

D is the nonlinear mapping

- Newton Method
- Levenberg-Marquardt Algorithm
  - Iterative, starts from initial solution
  - May be slow if initial solution far from real solution
  - Estimated solution may be function of the initial solution
  - Newton requires the computation of J, H
  - Levenberg-Marquardt doesn’t require the computation of H
Bundle adjustment

**Advantages**
- Handle large number of views
- Handle missing data

**Limitations**
- Large minimization problem (parameters grow with number of views)
- Requires good initial condition

- Used as the final step of SFM (i.e., after the factorization or algebraic approach)
- Factorization or algebraic approaches provide a initial solution for optimization problem
Lecture 7
Multi-view geometry

• The SFM problem
• Affine SFM
• Perspective SFM
• Self-calibration
• Applications
Self-calibration

- **Self-calibration** is the problem of recovering the metric reconstruction from the perspective (or affine) reconstruction
- We can self-calibrate the camera by making some assumptions about the cameras
Self-calibration

Several approaches:
- Use single-view metrology constraints (lecture 4)
- Direct approach (Kruppa Eqs) for 2 views
- Algebraic approach
- Stratified approach

[HZ] Chapters 19 “Auto-calibration”
Inject information about the camera during the bundle adjustment optimization.

For calibrated cameras, the similarity ambiguity is the only ambiguity. [Longuet-Higgins '81]
Lecture 7
Multi-view geometry

• The SFM problem
• Affine SFM
• Perspective SFM
• Self-calibration
• Applications
Structure from motion problem

Lucas & Kanade, 81
Chen & Medioni, 92
Debevec et al., 96
Levoy & Hanrahan, 96
Fitzgibbon & Zisserman, 98
Triggs et al., 99
Pollefeys et al., 99
Kutulakos & Seitz, 99
Levoy et al., 00
Hartley & Zisserman, 00
Dellaert et al., 00
Rusinkiewic et al., 02
Nistér, 04
Brown & Lowe, 04
Schindler et al., 04
Lourakis & Argyros, 04
Colombo et al. 05
Golparvar-Fard, et al. JAEI 10
Pandey et al. IFAC , 2010
Pandey et al. ICRA 2011
Microsoft's PhotoSynth
Snavely et al., 06-08
Schindler et al., 08
Agarwal et al., 09
Frahm et al., 10

Courtesy of Oxford Visual Geometry Group
Reconstruction and texture mapping

M. Pollefeys et al 98---
Incremental reconstruction of construction sites
Initial pair – 2168 & Complete Set 62,323 points, 160 images
Golparvar-Fard. Pena-Mora, Savarese 2008
The registration of images (08.27.08) within the reconstructed scene + Site photos for the Student Dining and Residence Hall project in Champaign, IL. Images courtesy of Turner Construction.
Reconstructed scene + Site photos
Results and applications

Next lecture

• Active Stereo & Volumetric Stereo
Direct approach

We use the following results:

1. A relationship that maps conics across views
2. Concept of absolute conic and its relationship to K
3. The Kruppa equations
Projections of conics across views

\[ X^T C_w X = 0 \quad \text{[Eq. 1]} \]

\[ X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ 1 \end{bmatrix} \]

\[ \left[ e' \right]_x C'^{-1} \left[ e' \right]_x = F C^{-1} F^T \quad \text{[Eq. 2]} \]
Projection of absolute conics across views

\[ [e']_x \omega'^{-1} [e']_x = F \omega^{-1} F^T \]

[Eq. 3]

\[ \omega = (K K^T)^{-1} \]

[Eq. 4]

\[ \omega' = (K' K'^T)^{-1} \]

[Eq. 5]

From lecture 4, [HZ] page 210, sec. 8.5.1
Kruppa equations

\[
\begin{pmatrix}
  u_2^T K' K'^T u_2 \\
  -u_1^T K' K'^T u_2 \\
  u_1^T K' K'^T u_1
\end{pmatrix}
\times
\begin{pmatrix}
  \sigma_1^2 \nu_1^T K K^T \nu_1 \\
  \sigma_1 \sigma_2 \nu_1^T K K^T \nu_2 \\
  \sigma_2^2 \nu_2^T K K^T \nu_2
\end{pmatrix} = 0
\]  

[Eq. 6]

where \( u_i \), \( \nu_i \) and \( \sigma_i \) are the columns and singular values of SVD of \( F \)

These give us two independent constraints in the elements of \( K \) and \( K' \)
Kruppa equations

[Eq. 7] \[
\begin{pmatrix}
    u_2^T K' K'^T u_2 \\
    -u_1^T K' K'^T u_2 \\
    u_1^T K' K'^T u_1
\end{pmatrix}
\times
\begin{pmatrix}
    \sigma_1^2 v_1^T K K^T v_1 \\
    \sigma_1 \sigma_2 v_1^T K K^T v_2 \\
    \sigma_2^2 v_2^T K K^T v_2
\end{pmatrix}
= 0
\]

[Eq. 8] \[
K' = K = \begin{pmatrix}
    f & 0 & 0 \\
    0 & f & 0 \\
    0 & 0 & 1
\end{pmatrix}
\]

[Eq. 9] \[
\alpha f^2 + \beta f + \gamma = 0 \quad \longrightarrow \quad f
\]
Kruppa equations

[Faugeras et al. 92]

• Powerful if we want to self-calibrate 2 cameras with unknown focal length

• Limitations:
  • Work on a camera pair
  • Don’t work if R=0

\[
\begin{align*}
[e']_x \omega^{-1} [e']_x &= F \omega^{-1} F^T \quad \text{becomes trivial} \\
\text{Since:} \quad F &= [e']_x
\end{align*}
\]
Self-calibration

Several approaches:
- Use single-view metrology constraints (lecture 4)
- Direct approach (Kruppa Eqs) for 2 views
- Algebraic approach
- Stratified approach
Algebraic approach  Multi-view approach

Suppose we have a projective reconstruction \( \{ \tilde{M}_i, \tilde{X}_j \} \)

Let \( H \) be a homography such that:

\[
\begin{align*}
\text{First perspective camera is canonical: } & \quad \tilde{M}_1 = [ I \quad 0 ] \quad \text{[Eq. 11]} \\
\text{i}^{th} \text{ perspective reconstruction of the camera (known): } & \quad \tilde{M}_i = [ A_i \quad b_i ] \quad \text{[Eq. 12]}
\end{align*}
\]

\[
\text{[Eq. 13]} \quad \left( A_i - b_i p^T \right) K_1 K_1^T \left( A_i - b_i p^T \right)^T = K_i K_i^T \quad \text{i=2...m}
\]

\[
\text{[Eq. 14]} \quad H = \begin{bmatrix}
K_1 & 0 \\
-p^T & 1
\end{bmatrix} \quad p \text{ is an unknown 3x1 vector} \quad K_1...K_m \text{ are unknown}
\]
Suppose we have a projective reconstruction.

Let $H$ be a homography such that:

\[
\begin{align*}
\text{First perspective camera is canonical: } & \quad \tilde{M}_1 = [I \quad 0] \quad [\text{Eq. 11}] \\
\text{i\textsuperscript{th} perspective reconstruction of the camera (known): } & \quad \tilde{M}_i = [A_i \quad b_i] \quad [\text{Eq. 12}]
\end{align*}
\]

\[
\text{[Eq. 13]} \quad \left( A_i - b_i p^T \right) K_1 \ K_1^T \left( A_i - b_i p^T \right)^T = K_i \ K_i^T \quad i=2\ldots m
\]

How many unknowns?  
- 3 from $p$
- 5 m from $K_1\ldots K_m$

How many equations? 5 independent equations [per view]
Suppose we have a projective reconstruction

Let H be a homography such that:

\[
\begin{align*}
\text{First perspective camera is canonical: } & \quad \tilde{M}_1 = \begin{bmatrix} I & 0 \end{bmatrix} \quad [\text{Eq. 11}] \\
\text{i}^{\text{th}} \text{ perspective reconstruction of the camera (known): } & \quad \tilde{M}_i = \begin{bmatrix} A_i & b_i \end{bmatrix} \\
\end{align*}
\]

Assume all camera matrices are identical: \( K_1 = K_2 \ldots = K_m \) 

\[
[\text{Eq. 15}] \quad \left( A_i - b_i p^T \right) K \ K^T \left( A_i - b_i p^T \right)^T = K \ K^T \quad i=2\ldots m
\]

How many unknowns? \begin{itemize}  
\item 3 from \( p \)  
\item 5 from \( K \)  
\end{itemize}

How many equations? 5 independent equations [per view]

We need at least 3 views to solve the self-calibration problem
# Algebraic approach

## Art of self-calibration:
Use assumptions on Ks to generate enough equations on the unknowns

<table>
<thead>
<tr>
<th>Condition</th>
<th>N. Views</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Constant internal parameters</td>
<td>3</td>
</tr>
<tr>
<td>• Aspect ratio and skew known</td>
<td>4</td>
</tr>
<tr>
<td>• Focal length and offset vary</td>
<td></td>
</tr>
<tr>
<td>• Skew =0, all other parameters vary</td>
<td>8</td>
</tr>
</tbody>
</table>

Issue: the larger is the number of view, the harder is the correspondence problem

Bundle adjustment helps!
SFM problem - summary

1. Estimate structure and motion up perspective transformation
   1. Algebraic
   2. factorization method
   3. bundle adjustment

2. Convert from perspective to metric (self-calibration)

3. Bundle adjustment

** or **

1. Bundle adjustment with self-calibration constraints