Computer Vision: Lecture 10

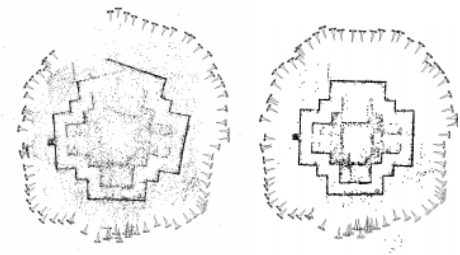
2023-11-30

Today's Lecture

Non-Sequential Methods for SfM

- Rotation averaging and translation registration
- Project presentation

Why non-sequential SfM: drift



Minimizing the Reprojection Error

Main goal

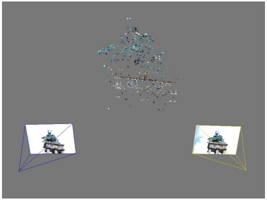
For given $\{\mathbf{x}_{ij}\}$ and $\{m_{ij}\}$ find a minimizer

$$\sum_{i,j} m_{ij} \left\| \mathbf{x}_{ij} - \pi(\mathtt{P}_i \mathbf{X}_j) \right\|^2 \to \min_{\{\mathtt{P}_i\},\{\mathbf{X}_j\}} .$$

Local optimization needs good starting point

$N = 2$: Recall lecture 5+6

- Compute $\mathtt{F}/\mathtt{E}$
- Compute a pair of camera matrices: $\mathtt{P}_1 = (\mathtt{I} \mid \mathbf{0})$, $\mathtt{P}_2 = ([\mathbf{e}_2]_\times \mathtt{F} \mid \mathbf{e}_2)$
- Triangulate $\mathbf{X}_j$ using DLT



Minimizing the Reprojection Error

Main goal

For given $\{\mathbf{x}_{ij}\}$ and $\{m_{ij}\}$ find a minimizer

$$\sum_{i,j} m_{ij} \left\| \mathbf{x}_{ij} - \pi(\mathtt{P}_i \mathbf{X}_j) \right\|^2 \to \min_{\{\mathtt{P}_i\},\{\mathbf{X}_j\}} .$$

Local optimization needs good starting point

$N > 2$: Sequential / incremental SfM

How do we compute the entire reconstruction?

1. For a pair of images compute initial reconstruction ($N = 2$)
2. For a new image viewing some of the previously reconstructed scene points, find the camera matrix (via DLT)
3. Compute new scene points using triangulation (DLT)
4. If there are more cameras/images goto step 2

Minimizing the Reprojection Error

Main goal

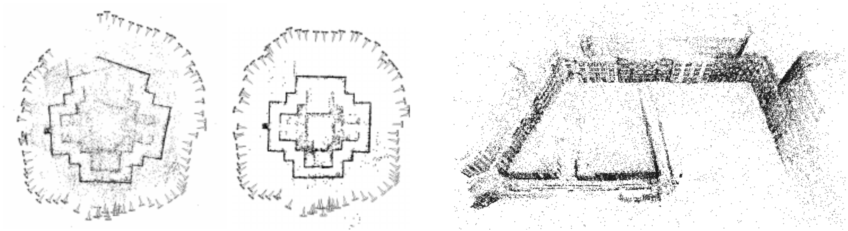
For given $\{\mathbf{x}_{ij}\}$ and $\{m_{ij}\}$ find a minimizer

$$\sum_{i,j} m_{ij} \left\| \mathbf{x}_{ij} - \pi(\mathtt{P}_i \mathbf{X}_j) \right\|^2 \to \min_{\{\mathtt{P}_i\},\{\mathbf{X}_j\}} .$$

Local optimization needs good starting point

$N > 2$: Sequential / incremental SfM

Main drawback: drift due to error accumulation



Minimizing the Reprojection Error

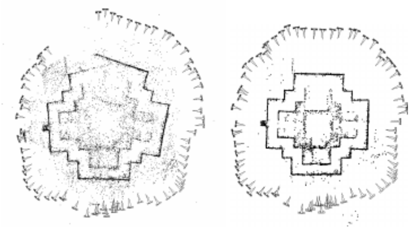
Main goal

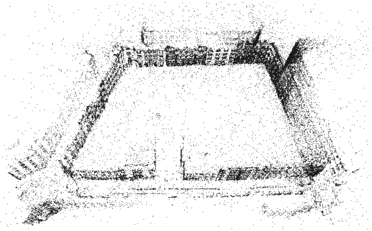
For given $\{\mathbf{x}_{ij}\}$ and $\{m_{ij}\}$ find a minimizer

$$\sum\nolimits_{i,j} m_{ij} \left\| \mathbf{x}_{ij} - \pi(\mathtt{P}_i \mathbf{X}_j) \right\|^2 \rightarrow \min_{\{\mathtt{P}_i\},\{\mathbf{X}_j\}} .$$

Local optimization needs good starting point

$N > 2$: Sequential / incremental SfM
Main drawback: drift due to error accumulation





Minimizing the Reprojection Error

Main goal

For given $\{\mathbf{x}_{ij}\}$ and $\{m_{ij}\}$ find a minimizer

$$\sum\nolimits_{i,j} m_{ij} \left\| \mathbf{x}_{ij} - \pi(\mathtt{P}_i \mathbf{X}_j) \right\|^2 \to \min_{\{\mathtt{P}_i\},\{\mathbf{X}_j\}}$$

Local optimization needs good starting point

Why is this objective hard to minimize?

- Bilinear, non-convex terms $\mathtt{P}_i \mathbf{X}_j$
- Perspective division $\pi(\mathbf{X}) = \frac{1}{X_3}\binom{X_1}{X_2}$
- Calibrated SfM: constraints $\mathtt{R}_i \in SO(3)$

Minimizing the Reprojection Error

Main goal

For given $\{\mathbf{x}_{ij}\}$ and $\{m_{ij}\}$ find a minimizer

$$\sum_{i,j} m_{ij} \left\| \mathbf{x}_{ij} - \pi(\mathtt{P}_i \mathbf{X}_j) \right\|^2 \to \min_{\{\mathtt{P}_i\},\{\mathbf{X}_j\}}$$

Local optimization needs good starting point

Why is this objective hard to minimize?

- Bilinear, non-convex terms $\mathtt{P}_i \mathbf{X}_j$
- Perspective division $\pi(\mathbf{X}) = \frac{1}{X_3}\binom{X_1}{X_2}$
- Calibrated SfM: constraints $\mathtt{R}_i \in SO(3)$

Minimizing the Reprojection Error

Question

We have a good idea how to compute $\mathtt{F}/\mathtt{E}$ robustly for image pairs:
RANSAC + 5/7/8-point method
Can we use this pairwise information?

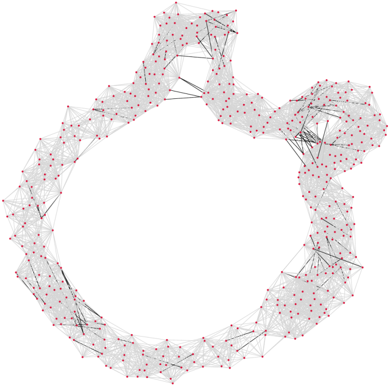
Issues:

- Ambiguity: projective ambiguity ($\mathtt{F}$), similarity transformation ($\mathtt{E}$)
- Perceptual aliasing

Calibrated Structure and Motion

- Estimate $\mathtt{E}_{ij}$ between images i and j if possible

$\approx$ 500 images



Camera graph

Calibrated Structure and Motion

- Each $\mathtt{E}_{ij}$ yields relative rotations $\mathtt{R}_{ij}$ and translations $\mathbf{T}_{ij}$
- Pair of camera matrices in local coordinate system

$$\tilde{\mathtt{P}}_{ij,0} = (\mathtt{I} \mid \mathbf{0}) \qquad \tilde{\mathtt{P}}_{ij,1} = (\mathtt{R}_{ij} \mid \mathbf{T}_{ij})$$

- Camera matrices in common coordinate system $\mathtt{P}_i = (\mathtt{R}_i \mid \mathbf{T}_i)$
- Transformation from $\mathtt{P}_i/\mathtt{P}_j$ to a canonical pair $\tilde{\mathtt{P}}_{ij,0}/\tilde{\mathtt{P}}_{ij,1}$

$$\begin{aligned}
(\mathtt{R}_i \mid \mathbf{T}_i) \begin{pmatrix} \mathtt{R}_i^\top & -\mathtt{R}_i^\top \mathbf{T}_i \\ \mathbf{0}^\top & 1 \end{pmatrix} &= (\mathtt{I} \mid \mathbf{0}) \\
(\mathtt{R}_j \mid \mathbf{T}_j) \begin{pmatrix} \mathtt{R}_i^\top & -\mathtt{R}_i^\top \mathbf{T}_i \\ \mathbf{0}^\top & 1 \end{pmatrix} &= (\mathtt{R}_j \mathtt{R}_i^\top \mid \mathbf{T}_j - \mathtt{R}_j \mathtt{R}_i^\top \mathbf{T}_i) \\
&\stackrel{!}{=} (\mathtt{R}_{ij} \mid \mathbf{T}_{ij})
\end{aligned}$$

- Therefore $\mathtt{R}_{ij} = \mathtt{R}_j \mathtt{R}_i^\top$ and $\mathbf{T}_{ij} = \mathbf{T}_j - \mathtt{R}_j \mathtt{R}_i^\top \mathbf{T}_i$

Calibrated Structure and Motion

- Each $\mathtt{E}_{ij}$ yields relative rotations $\mathtt{R}_{ij}$ and translations $\mathbf{T}_{ij}$
- Pair of camera matrices in local coordinate system

$$\tilde{\mathtt{P}}_{ij,0} = (\mathtt{I} \mid \mathbf{0}) \qquad \tilde{\mathtt{P}}_{ij,1} = (\mathtt{R}_{ij} \mid \mathbf{T}_{ij})$$

- Camera matrices in common coordinate system $\mathtt{P}_i = (\mathtt{R}_i \mid \mathbf{T}_i)$
- Transformation from $\mathtt{P}_i/\mathtt{P}_j$ to a canonical pair $\tilde{\mathtt{P}}_{ij,0}/\tilde{\mathtt{P}}_{ij,1}$

$$\begin{aligned}
(\mathtt{R}_i \mid \mathbf{T}_i) \begin{pmatrix} \mathtt{R}_i^\top & -\mathtt{R}_i^\top \mathbf{T}_i \\ \mathbf{0}^\top & 1 \end{pmatrix} &= (\mathtt{I} \mid \mathbf{0}) \\
(\mathtt{R}_j \mid \mathbf{T}_j) \begin{pmatrix} \mathtt{R}_i^\top & -\mathtt{R}_i^\top \mathbf{T}_i \\ \mathbf{0}^\top & 1 \end{pmatrix} &= (\mathtt{R}_j \mathtt{R}_i^\top \mid \mathbf{T}_j - \mathtt{R}_j \mathtt{R}_i^\top \mathbf{T}_i) \\
&\stackrel{!}{=} (\mathtt{R}_{ij} \mid \mathbf{T}_{ij})
\end{aligned}$$

- Therefore $\mathtt{R}_{ij} = \mathtt{R}_j \mathtt{R}_i^\top$ and $\mathbf{T}_{ij} = \mathbf{T}_j - \mathtt{R}_j \mathtt{R}_i^\top \mathbf{T}_i$

Calibrated Structure and Motion

Rotation averaging

For given relative rotation matrices $\{\mathtt{R}_{ij}\}$ determine $\{\mathtt{R}_i\}$ such that

$$\mathtt{R}_{ij} \approx \mathtt{R}_j \mathtt{R}_i^\top \qquad \text{or} \qquad \mathtt{R}_{ij}\mathtt{R}_i - \mathtt{R}_j \approx 0$$

- Method 0: set $\mathtt{R}_1 = \mathtt{I}$ and chain rotations

$$\mathtt{R}_j \leftarrow \mathtt{R}_{ij}\mathtt{R}_i \qquad \text{if } \mathtt{R}_j \text{ is not assigned yet}$$

- Method 1: null-space method
  Place $\mathtt{R}_{ij}$ and $-\mathtt{I}$ in the columns corresponding to $\mathtt{R}_i$ and $\mathtt{R}_j$

$$\underbrace{\begin{pmatrix} & \vdots & \\ \mathtt{R}_{ij} & & -\mathtt{I} \\ & \vdots & \end{pmatrix}}_{=\mathtt{A}} \begin{pmatrix} \mathtt{R}_1 \\ \vdots \\ \mathtt{R}_2 \end{pmatrix} = 0_{3N\times 3}$$

  Extract 3 rightmost columns of $\mathtt{V}$, $\mathtt{U}\Sigma\mathtt{V}^\top = \mathrm{SVD}(\mathtt{A})$

- Method 2: solve a semi-definite program (SDP)

Calibrated Structure and Motion

Rotation averaging

For given relative rotation matrices $\{\mathtt{R}_{ij}\}$ determine $\{\mathtt{R}_i\}$ such that

$$\mathtt{R}_{ij} \approx \mathtt{R}_j \mathtt{R}_i^\top \qquad \text{or} \qquad \mathtt{R}_{ij}\mathtt{R}_i - \mathtt{R}_j \approx 0$$

- Method 0: set $\mathtt{R}_1 = \mathtt{I}$ and chain rotations

$$\mathtt{R}_j \leftarrow \mathtt{R}_{ij}\mathtt{R}_i \qquad \text{if } \mathtt{R}_j \text{ is not assigned yet}$$

- Method 1: null-space method
  Place $\mathtt{R}_{ij}$ and $-\mathtt{I}$ in the columns corresponding to $\mathtt{R}_i$ and $\mathtt{R}_j$

$$\underbrace{\begin{pmatrix} & \vdots & \\ \mathtt{R}_{ij} & & -\mathtt{I} \\ & \vdots & \end{pmatrix}}_{=\mathtt{A}} \begin{pmatrix} \mathtt{R}_1 \\ \vdots \\ \mathtt{R}_2 \end{pmatrix} = 0_{3N\times 3}$$

  Extract 3 rightmost columns of $\mathtt{V}$, $\mathtt{U}\Sigma\mathtt{V}^\top = \mathrm{SVD}(\mathtt{A})$

- Method 2: solve a semi-definite program (SDP)

Calibrated Structure and Motion

Rotation averaging

For given relative rotation matrices $\{\mathtt{R}_{ij}\}$ determine $\{\mathtt{R}_i\}$ such that

$$\mathtt{R}_{ij} \approx \mathtt{R}_j \mathtt{R}_i^\top \qquad \text{or} \qquad \mathtt{R}_{ij}\mathtt{R}_i - \mathtt{R}_j \approx 0$$

- Method 0: set $\mathtt{R}_1 = \mathtt{I}$ and chain rotations

$$\mathtt{R}_j \leftarrow \mathtt{R}_{ij}\mathtt{R}_i \qquad \text{if } \mathtt{R}_j \text{ is not assigned yet}$$

- Method 1: null-space method
  Place $\mathtt{R}_{ij}$ and $-\mathtt{I}$ in the columns corresponding to $\mathtt{R}_i$ and $\mathtt{R}_j$

$$\underbrace{\begin{pmatrix} & \vdots & \\ \mathtt{R}_{ij} & & -\mathtt{I} \\ & \vdots & \end{pmatrix}}_{=\mathtt{A}} \begin{pmatrix} \mathtt{R}_1 \\ \vdots \\ \mathtt{R}_2 \end{pmatrix} = 0_{3N\times 3}$$

  Extract 3 rightmost columns of $\mathtt{V}$, $\mathtt{U}\Sigma\mathtt{V}^\top = \mathrm{SVD}(\mathtt{A})$

- Method 2: solve a semi-definite program (SDP)

Calibrated Structure and Motion

Translation registration I

For fixed rotation matrices $\{\mathtt{R}_i\}$ and given relative translations $\{\mathbf{T}_{ij}\}$ determine translations $\{\mathbf{T}_i\}$ such that

$$\mathbf{T}_{ij} \sim \mathbf{T}_j - \mathtt{R}_j \mathtt{R}_i^\top \mathbf{T}_i$$

Problem 1: $\mathbf{T}_{ij}$ are only known up to scale.

$$\mathbf{T}_{ij} \times (\mathbf{T}_j - \mathtt{R}_j \mathtt{R}_i^\top \mathbf{T}_i) \approx \mathbf{0}$$

Problem 2: numerical stability when $\mathtt{R}_j \mathtt{R}_i^\top \approx \mathtt{I}$ (e.g. linear motion)

Translation registration II

For fixed rotation matrices $\{\mathtt{R}_i\}$ determine translations $\{\mathbf{T}_i\}$ and scene points $\{\mathbf{X}_j\}$ such that

$$\mathbf{x}_{ij} \approx \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)$$

Calibrated Structure and Motion

Translation registration II

For fixed rotation matrices $\{\mathtt{R}_i\}$ determine translations $\{\mathbf{T}_i\}$ and scene points $\{\mathbf{X}_j\}$ such that

$$\mathbf{x}_{ij} \approx \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)$$

- Method 1: linear DLT solution

$$\mathbf{x}_{ij} \times (\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i) = \mathbf{0} \iff [\mathbf{x}_{ij}]_\times \mathtt{R}_i \mathbf{X}_j + [\mathbf{x}_{ij}]_\times \mathbf{T}_i$$

- Homogeneous least-squares instance

$$\begin{pmatrix} & \vdots & \\ [\mathbf{x}_{ij}]_\times & & [\mathbf{x}_{ij}]_\times \mathtt{R}_i \\ & \vdots & \end{pmatrix} \begin{pmatrix} \vdots \\ \mathbf{T}_i \\ \vdots \\ \mathbf{X}_j \\ \vdots \end{pmatrix} = \mathbf{0}$$

Calibrated Structure and Motion

Translation registration II

For fixed rotation matrices $\{\mathtt{R}_i\}$ determine translations $\{\mathbf{T}_i\}$ and scene points $\{\mathbf{X}_j\}$ such that

$$\mathbf{x}_{ij} \approx \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)$$

- Method 1: linear DLT solution

$$\mathbf{x}_{ij} \times (\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i) = \mathbf{0} \iff [\mathbf{x}_{ij}]_\times \mathtt{R}_i \mathbf{X}_j + [\mathbf{x}_{ij}]_\times \mathbf{T}_i$$

- Homogeneous least-squares instance

$$\begin{pmatrix} & \vdots & \\ [\mathbf{x}_{ij}]_\times & & [\mathbf{x}_{ij}]_\times \mathtt{R}_i \\ & \vdots & \end{pmatrix} \begin{pmatrix} \vdots \\ \mathbf{T}_i \\ \vdots \\ \mathbf{X}_j \\ \vdots \end{pmatrix} = \mathbf{0}$$

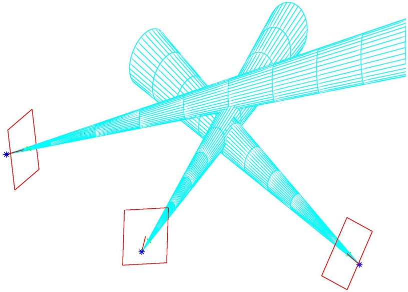
Calibrated Structure and Motion

Translation registration II

For fixed rotation matrices $\{\mathtt{R}_i\}$ determine translations $\{\mathbf{T}_i\}$ and scene points $\{\mathbf{X}_j\}$ such that

$$\mathbf{x}_{ij} \approx \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)$$

Method 2:



Translation registration II

For fixed rotation matrices $\{\mathtt{R}_i\}$ determine translations $\{\mathbf{T}_i\}$ and scene points $\{\mathbf{X}_j\}$ such that

$$\mathbf{x}_{ij} \approx \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)$$

Method 2:

- Constraint $\|\mathbf{x}_{ij} - \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)\| \le b$ is essentially convex in $\mathbf{X}_j$ and $\mathbf{T}_i$

$$\begin{aligned} & \|\mathbf{x}_i - \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)\| \le b \wedge (\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)_3 \ge 0 \\ \iff & \|(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)_3 \mathbf{x}_i - (\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)_{1,2}\| - (\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)_3 b \le 0 \wedge (\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)_3 \ge 0 \end{aligned}$$

  Convex cone

- Linear or SOCP constraint (recall lecture 9)
- Convex feasibility problem
- Bisection search (lecture 9) to determine $b \ge 0$

Calibrated Structure and Motion

Translation registration II

For fixed rotation matrices $\{\mathtt{R}_i\}$ determine translations $\{\mathbf{T}_i\}$ and scene points $\{\mathbf{X}_j\}$ such that

$$\mathbf{x}_{ij} \approx \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)$$

Method 3:

- Shorthand notation: $\mathbf{X}'_{ij} = \mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i$ linear in $\mathbf{X}_j$ and $\mathbf{T}_i$
- Take a closer look at the squared reprojection error

$$\left\| \mathbf{x}_{ij} - \pi(\mathbf{X}'_{ij}) \right\|^2 = \frac{\left\| Z'_{ij} \mathbf{x}_{ij} - \binom{X'_{ij}}{Y'_{ij}} \right\|^2}{(Z'_{ij})^2} \qquad \mathbf{X}'_{ij} = (X'_{ij}, Y'_{ij}, Z'_{ij})^\top$$

- Replace with convex approximation

$$\frac{\left\| Z'_{ij} \mathbf{x}_{ij} - \binom{X'_{ij}}{Y'_{ij}} \right\|^2}{Z'_{ij}} \qquad \text{s.t. } Z'_{ij} \geq 0$$

- Convex optimization problem

Calibrated Structure and Motion

Translation registration II

For fixed rotation matrices $\{\mathtt{R}_i\}$ determine translations $\{\mathbf{T}_i\}$ and scene points $\{\mathbf{X}_j\}$ such that

$$\mathbf{x}_{ij} \approx \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)$$

Method 3:

- Shorthand notation: $\mathbf{X}'_{ij} = \mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i$ linear in $\mathbf{X}_j$ and $\mathbf{T}_i$
- Take a closer look at the squared reprojection error

$$\left\| \mathbf{x}_{ij} - \pi(\mathbf{X}'_{ij}) \right\|^2 = \frac{\left\| Z'_{ij}\mathbf{x}_{ij} - \binom{X'_{ij}}{Y'_{ij}} \right\|^2}{(Z'_{ij})^2} \qquad \mathbf{X}'_{ij} = (X'_{ij}, Y'_{ij}, Z'_{ij})^\top$$

- Replace with convex approximation

$$\frac{\left\| Z'_{ij}\mathbf{x}_{ij} - \binom{X'_{ij}}{Y'_{ij}} \right\|^2}{Z'_{ij}} \qquad \text{s.t. } Z'_{ij} \geq 0$$

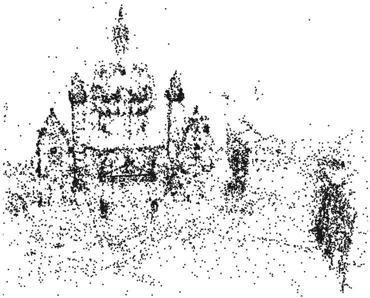
- Convex optimization problem

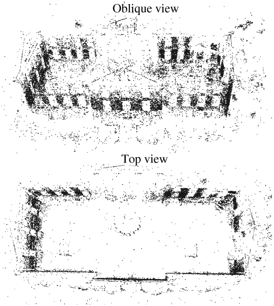
Calibrated Structure and Motion

Translation registration II

For fixed rotation matrices $\{\mathtt{R}_i\}$ determine translations $\{\mathbf{T}_i\}$ and scene points $\{\mathbf{X}_j\}$ such that

$$\mathbf{x}_{ij} \approx \pi(\mathtt{R}_i \mathbf{X}_j + \mathbf{T}_i)$$





Rotation averaging (method 1) + translation registration (method 3)

Calibrated Structure and Motion

Discussion

Non-sequential SfM (rotation averaging and translation registration) have shortcomings.

- Often not very scalable
  - Large non-linear convex programs to solve
  - Can only handle a very limited amount of outliers
- For city-scale 3D models
  - Combine with "bottom-up" approach: merge smaller sub-models
  - Exploit redundancy in images: apply non-sequential SfM on sparser image set
- Provides not very accurate initial models
  - Average reprojection error ≈ 5 pixels
  - Can be a problem for (robust) bundle adjustment

Upcoming

Project presentation