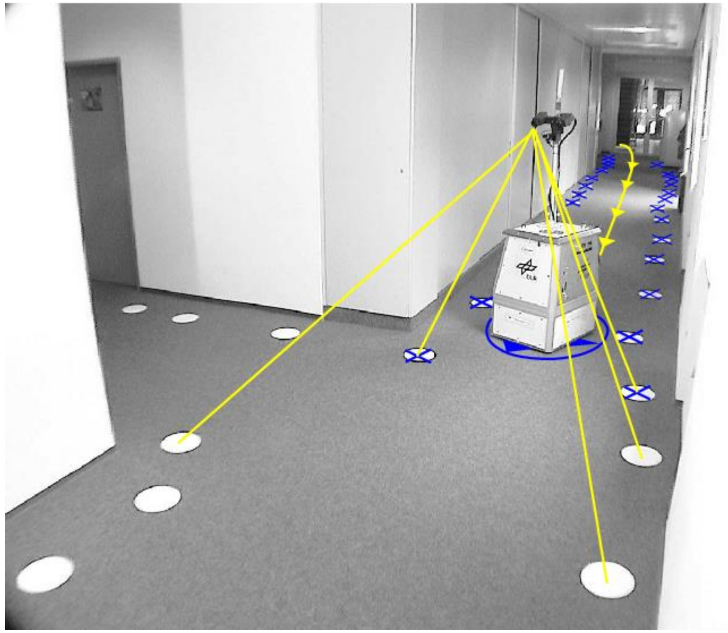
Robust Mechatronics

Localization and Mapping for Autonomous Mobile Systems



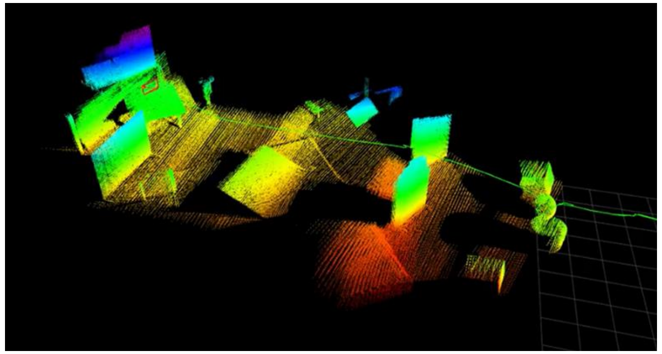
Dr Loukas Bampis, Assistant Professor

Mechatronics & Systems Automation Lab

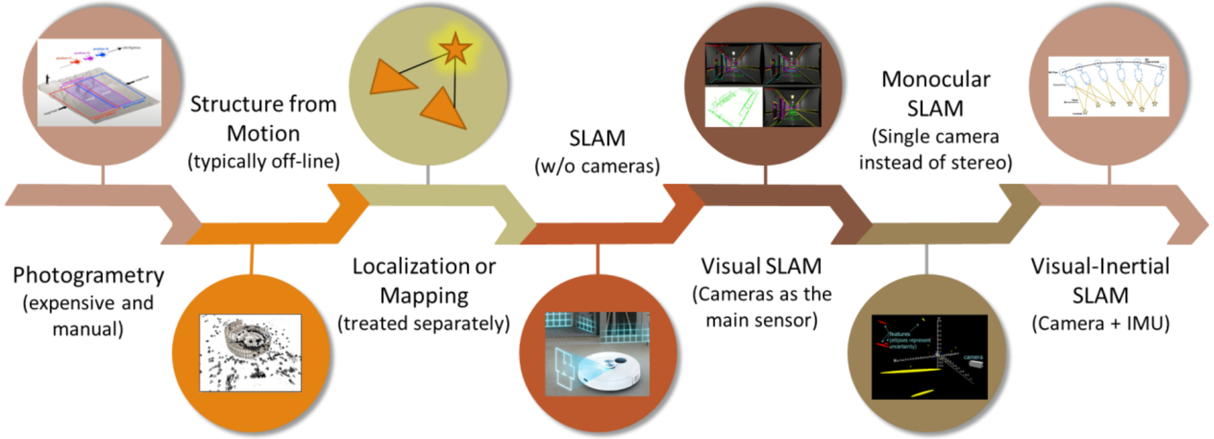
Localization and Mapping for Autonomous Mobile Systems

What we previously discussed:

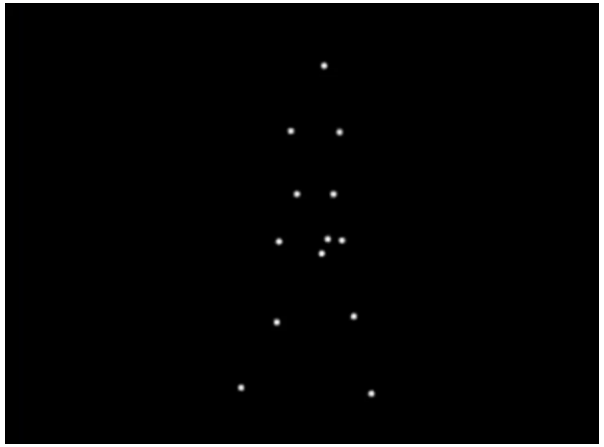
What is SLAM?



Evolution of SLAM

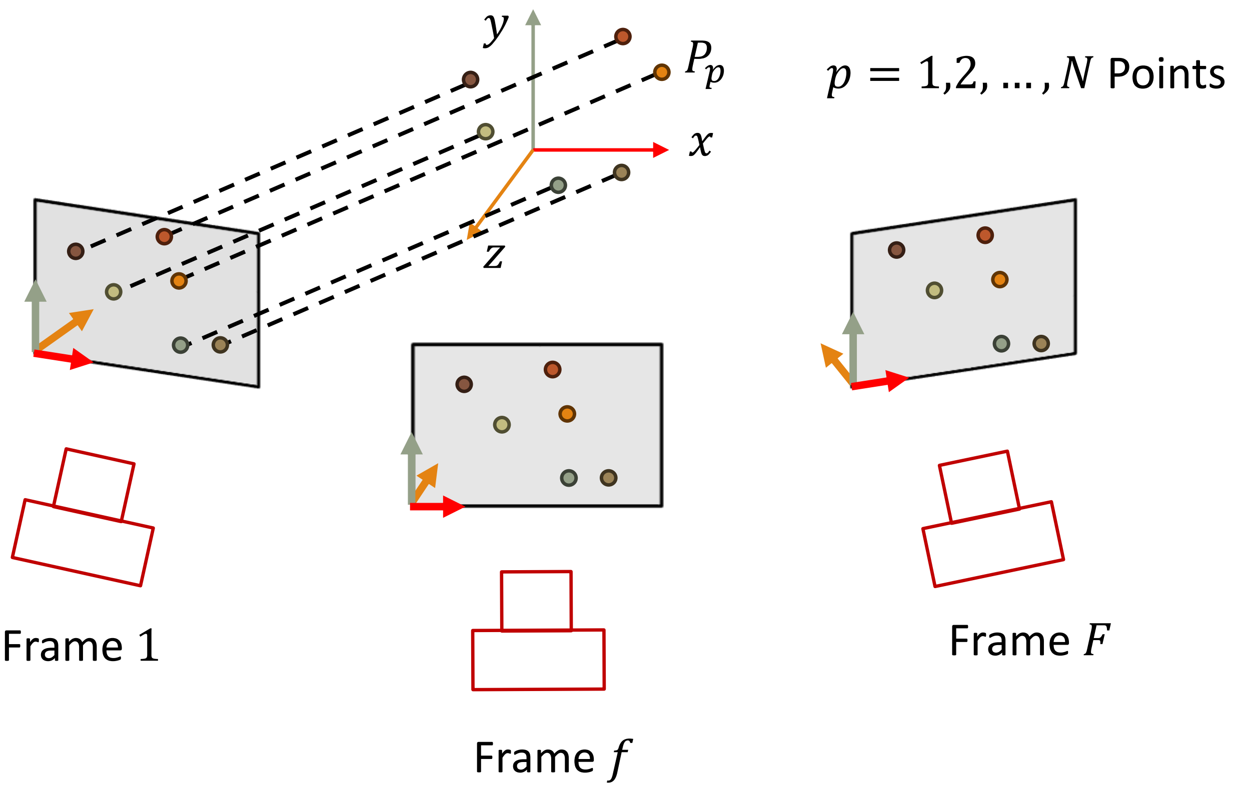


Structure from Motion



Localization and Mapping for Autonomous Mobile Systems

Structure from Motion – SfM: The early-days approach



Given sets of corresponding image points (2D): $(u_{f,p}, v_{f,p})$

Find scene points (3D):

$$P_p = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}$$

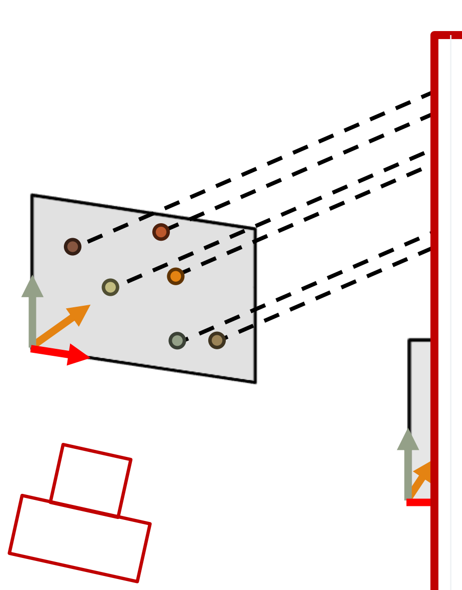
Early-days' approach assumption:

- Orthographic camera
- In subsequent years, this assumption was revoked

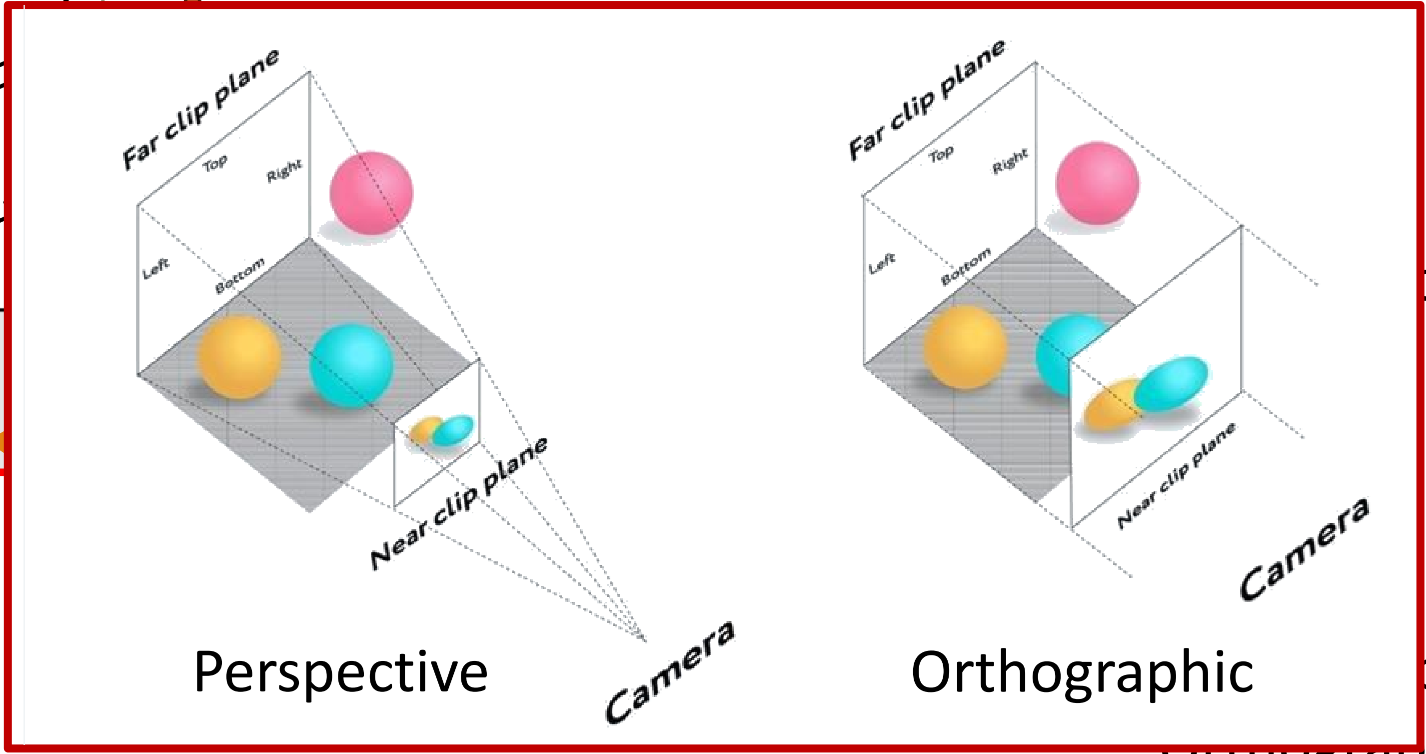
Tomasi and Kanade, 1992

Localization and Mapping for Autonomous Mobile Systems

Structure from Motion – SfM: The early-days approach



Frame 1



Perspective

Orthographic

corresponding
(2D): $(u_{f,p}, v_{f,p})$

oints (3D):

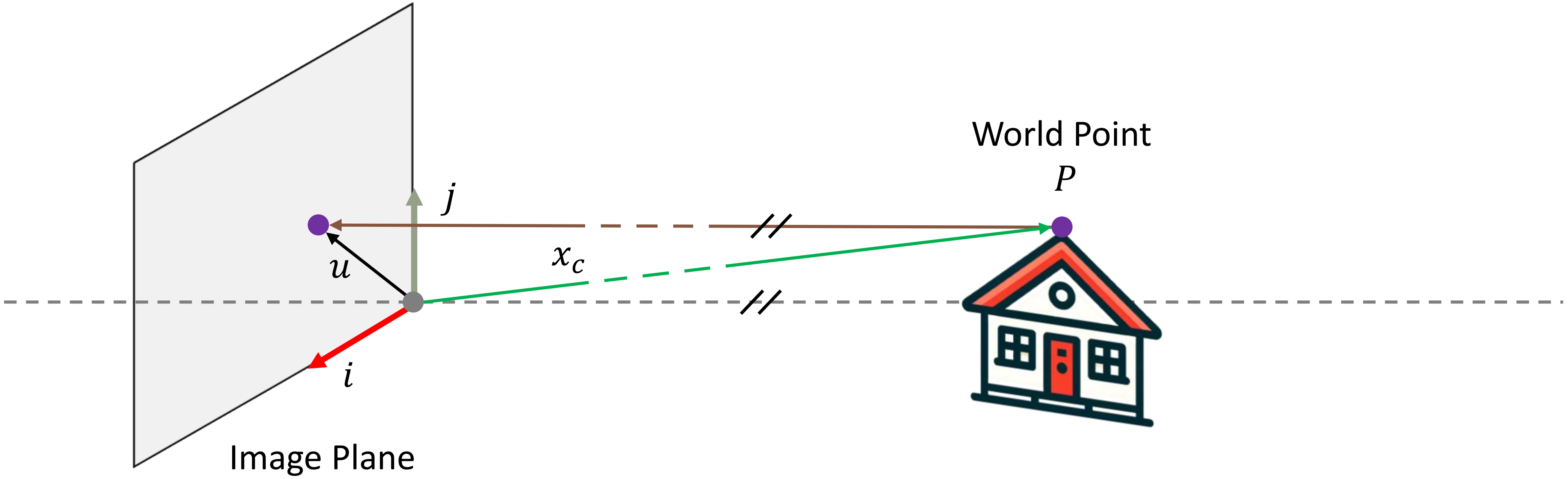
pproach assumption:
Orthographic camera

- In subsequent years, this assumption was revoked

Tomasi and Kanade, 1992

Localization and Mapping for Autonomous Mobile Systems

Structure from Motion – SfM: The early-days approach

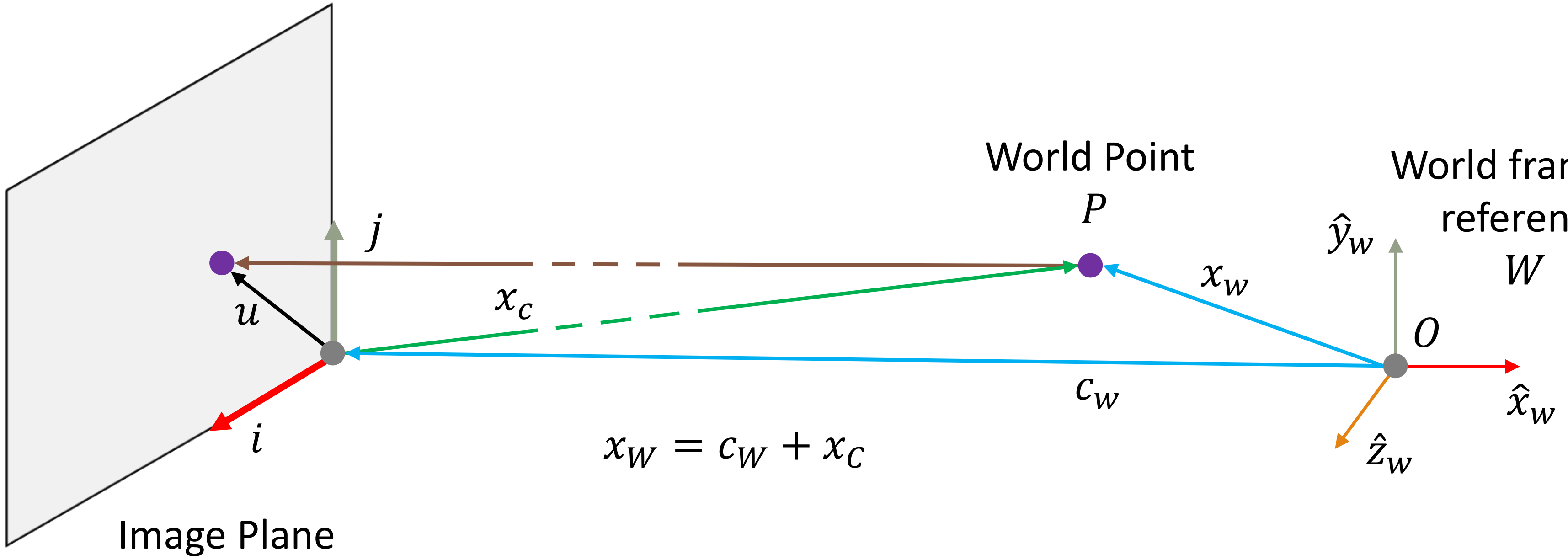


$$u = \mathbf{i} \cdot x_C = \mathbf{i}^T x_C$$

$$v = \mathbf{j} \cdot x_C = \mathbf{j}^T x_C$$

The orthographic camera assumption better-resembles the true perspective camera projection when the distance of the scene from the camera is large compared to the depth variation within the scene.

Structure from Motion – SfM: The early-days approach



$$u = \mathbf{i}^T x_c = \mathbf{i}^T (x_w - c_w) = \mathbf{i}^T (P - C)$$

$$v = \mathbf{j}^T x_c = \mathbf{j}^T (x_w - c_w) = \mathbf{j}^T (P - C)$$

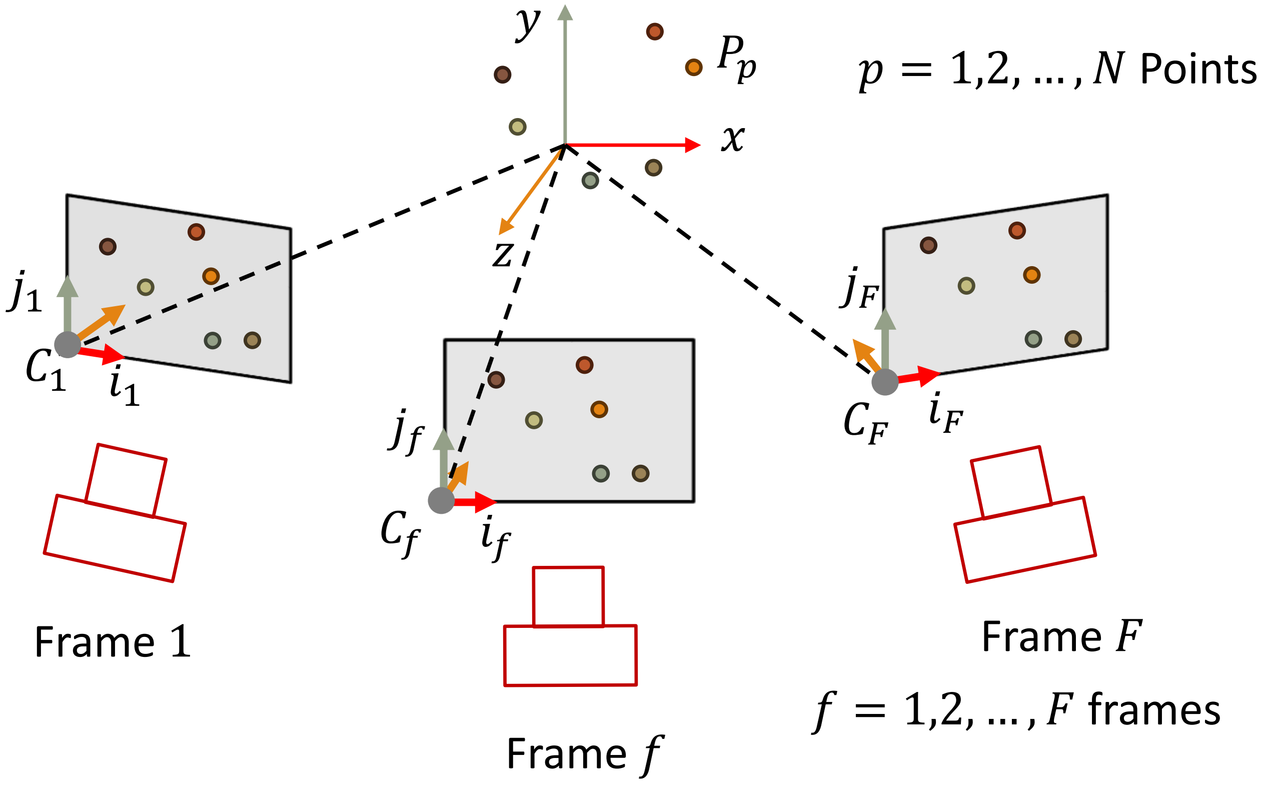
C: camera w.r.t. W FoR

P: Point w.r.t. W FoR

$$\boxed{u = \mathbf{i}^T (P - C)}$$

$$\boxed{v = \mathbf{j}^T (P - C)}$$

Structure from Motion – SfM: The early-days approach



Given sets of corresponding image points (2D) in F frames: $(u_{f,p}, v_{f,p})$

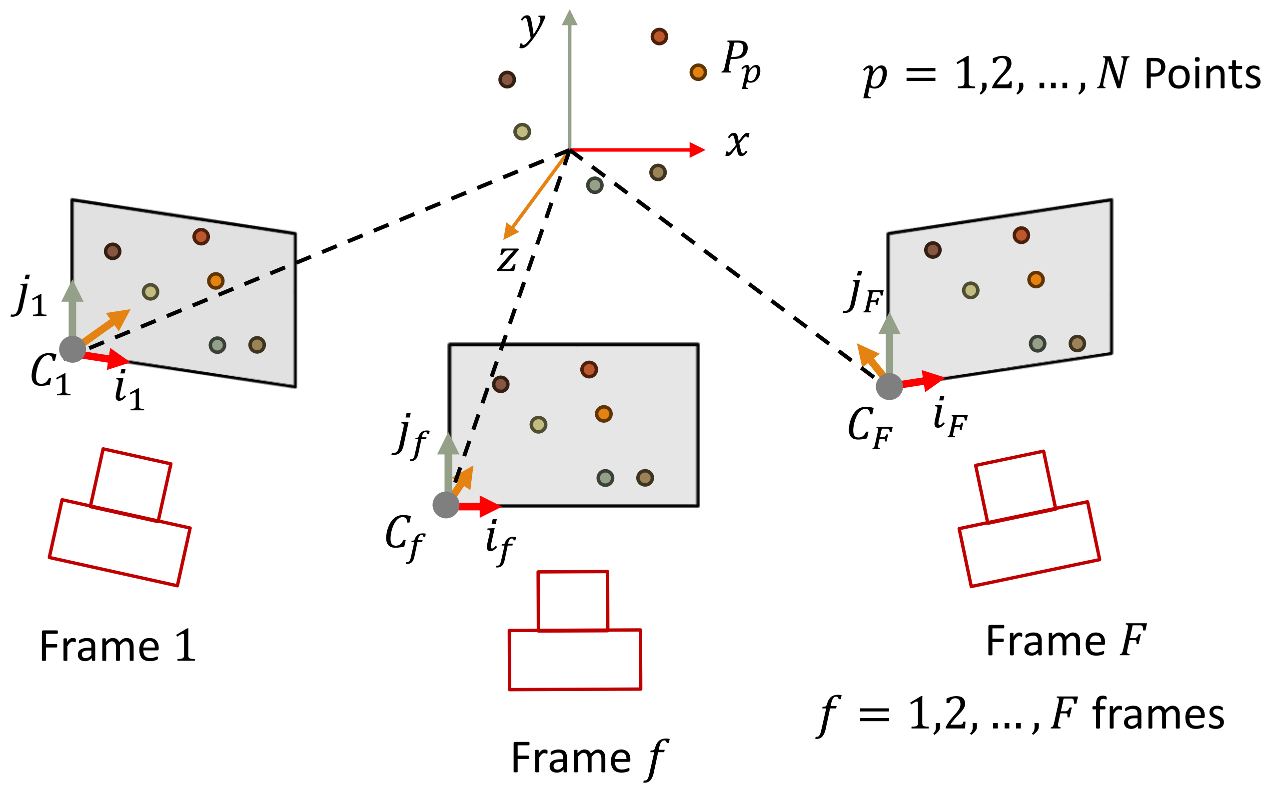
Find scene points (3D):

$$P_p = \begin{bmatrix} x_p \\ y_p \\ z_p \end{bmatrix}$$

Unknowns:

- Camera Positions: C_f
- Camera Orientations: $(\mathbf{i}_f, \mathbf{j}_f)$

Structure from Motion – SfM: The early-days approach



$p = 1,2, \dots, N$ Points

$f = 1,2, \dots, F$ frames

Point P_p representation in frame f:

$$u_{f,p} = \mathbf{i}_f^T (P_p - C_f)$$
$$v_{f,p} = \underline{\mathbf{j}_f^T (P_p - C_f)}$$

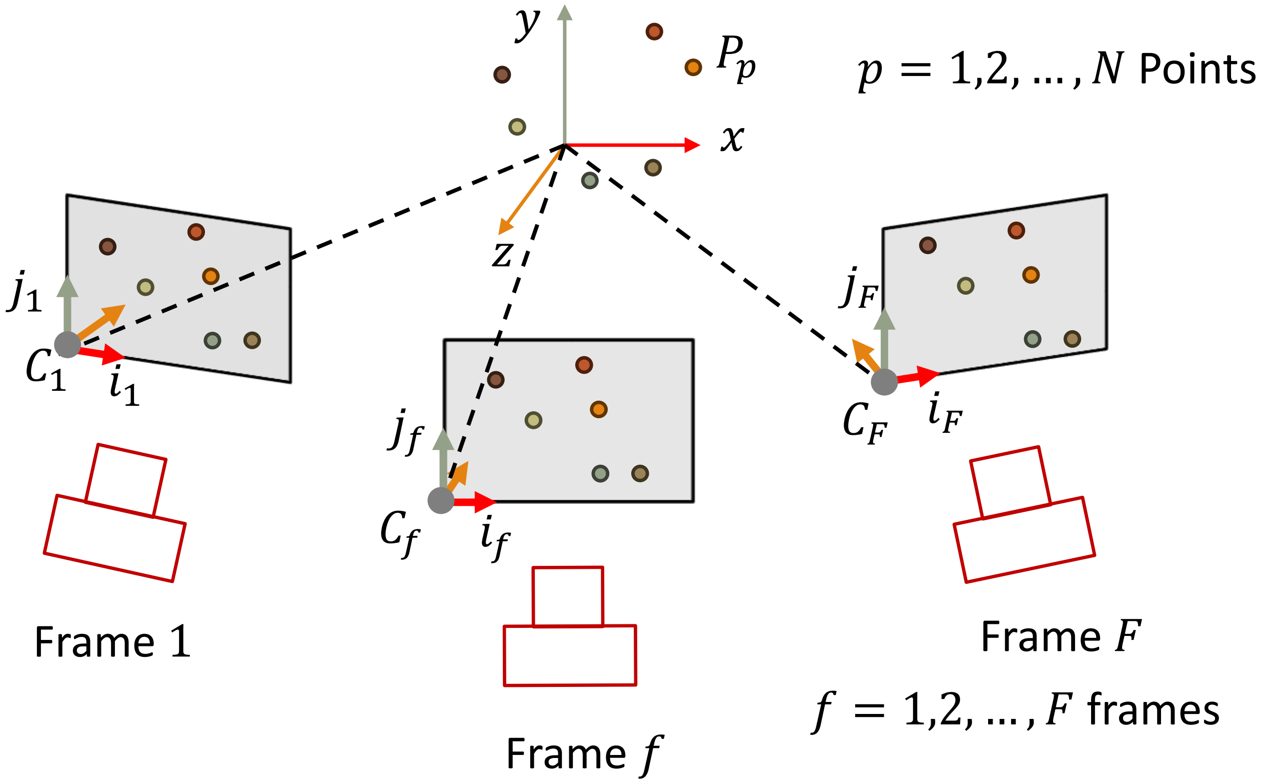
We know those

We don't know those

How do we solve for P_p?

While also not knowing $\{\mathbf{i}_f^T, \mathbf{j}_f^T\}$ and C_f.

Structure from Motion – SfM: The early-days approach



Point P_p representation in frame f:

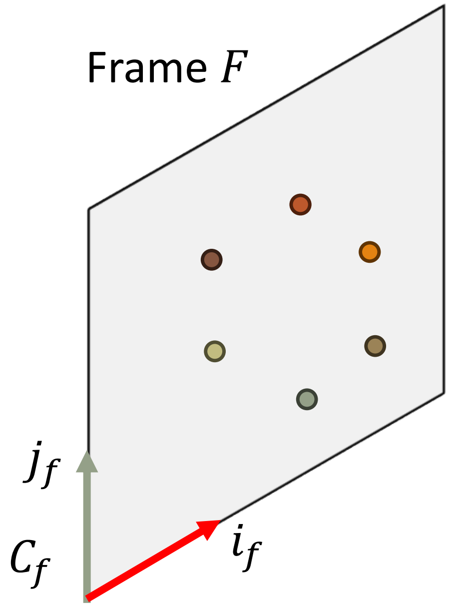
$$u_{f,p} = \mathbf{i}_f^T(P_p - C_f)$$
$$v_{f,p} = \underline{\mathbf{j}_f^T(P_p - C_f)}$$

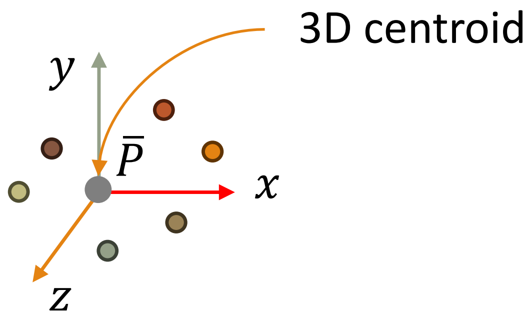
We know those

We don't know those

We can implement a trick to remove C_f from the equations.

Structure from Motion – SfM: The early-days approach





World Points

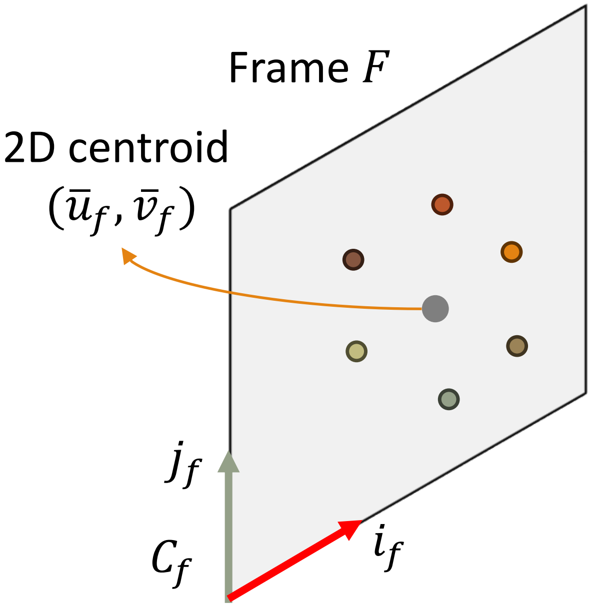
First, put the world's origin frame at the centroid of the scene's points:

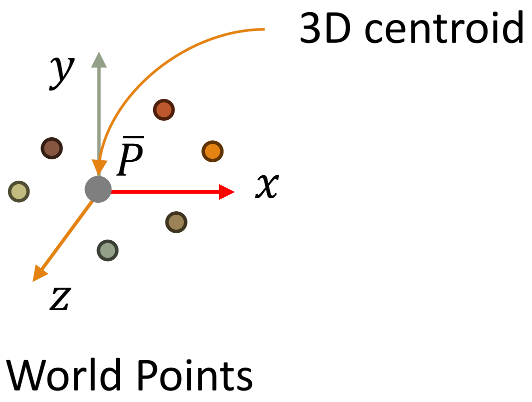
$$\frac{1}{N}\sum_{p=1}^{N} P_p = \bar{P} = 0$$

This means that the points to be found (P_p) will be expressed w.r.t. the centroid of the world.

Localization and Mapping for Autonomous Mobile Systems

Structure from Motion – SfM: The early-days approach

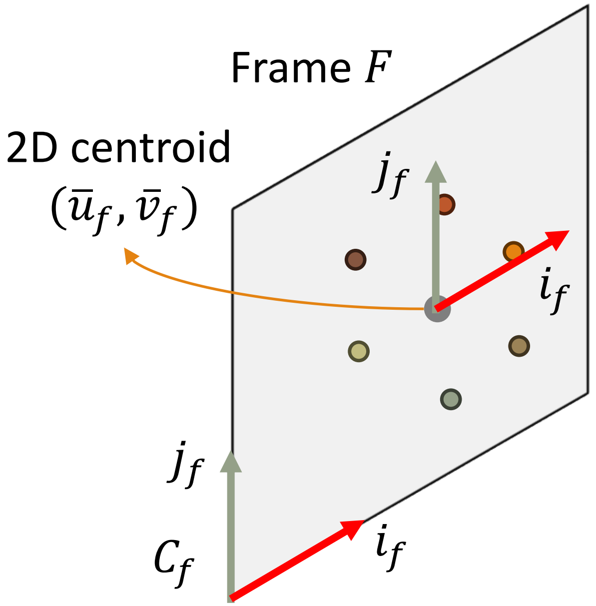


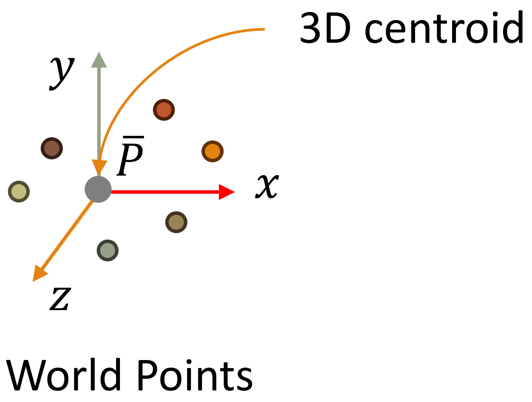


In accordance, the centroid $(\bar{u}_f, \bar{v}_f)$ of the image points in frame f can be written as:

$$\bar{u}_f = \frac{1}{N}\sum_{p=1}^{N} u_{f,p} = \frac{1}{N}\sum_{p=1}^{N} \mathbf{i}_f^T (P_p - C_f) = \mathbf{i}_f^T \cancelto{0}{\frac{1}{N}\sum_{p=1}^{N} P_p} - \frac{1}{N}\sum_{p=1}^{N} \mathbf{i}_f^T C_f = -\frac{1}{N}\sum_{p=1}^{N} \mathbf{i}_f^T C_f = -\mathbf{i}_f^T C_f$$

Structure from Motion – SfM: The early-days approach





The centroid $(\bar{u}_f, \bar{v}_f)$ of the image points in frame f can be written as:

$$\bar{u}_f = -\mathbf{i}_f^T C_f$$

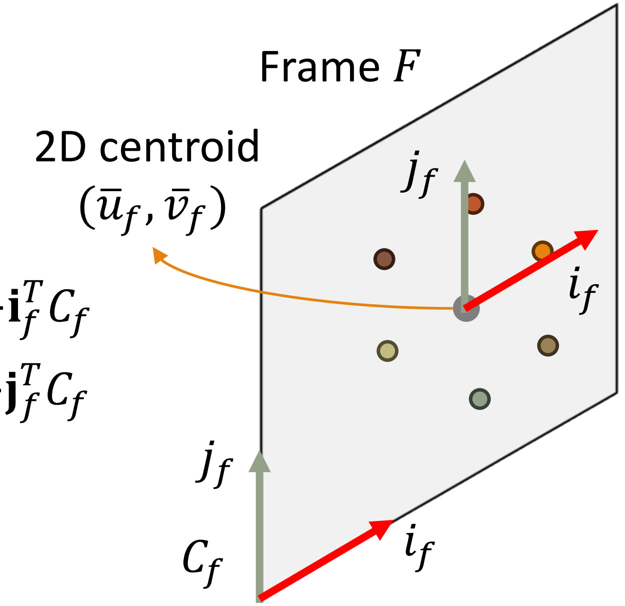
$$\bar{v}_f = -\mathbf{j}_f^T C_f$$

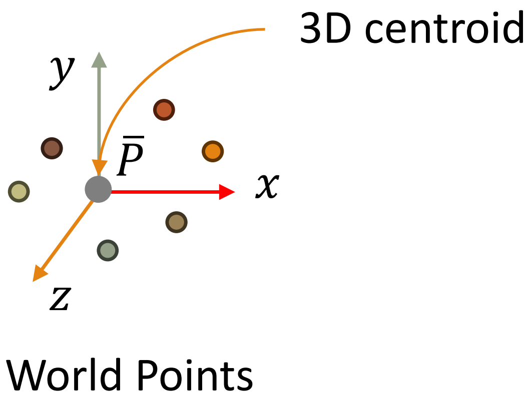
The trick is to express the 2D points' coordinates w.r.t. the 2D camera centroid.

Localization and Mapping for Autonomous Mobile Systems

Structure from Motion – SfM: The early-days approach

$$\bar{u}_f = -\mathbf{i}_f^T C_f$$
$$\bar{v}_f = -\mathbf{j}_f^T C_f$$





The 2D point's coordinates $(\tilde{u}_{f,p}, \tilde{v}_{f,p})$ w.r.t. the 2D camera centroid for each frame f:

$$\left.\begin{aligned} \tilde{u}_{f,p} &= u_{f,p} - \bar{u}_f = \mathbf{i}_f^T(P_p - C_f) - \bar{u}_f = \mathbf{i}_f^T(P_p - C_f) + \mathbf{i}_f^T C_f = \mathbf{i}_f^T P_p \\ \tilde{v}_{f,p} &= v_{f,p} - \bar{v}_f = \mathbf{j}_f^T(P_p - C_f) - \bar{v}_f = j_f^T(P_p - C_f) + \mathbf{j}_f^T C_f = \mathbf{j}_f^T P_p \end{aligned}\right\} \quad C_f \text{ removed}$$

Structure from Motion – SfM: The early-days approach

Observation Matrix W

$$
\begin{array}{c c}
 & \begin{array}{cccc} \text{Point 1} & \text{Point 2} & & \text{Point } N \end{array} \\
\begin{array}{c} \text{Image 1} \\ \text{Image 2} \\ \\ \text{Image } F \\ \text{Image 1} \\ \text{Image 2} \\ \\ \text{Image } F \end{array} &
\underbrace{\begin{bmatrix}
\tilde{u}_{1,1} & \tilde{u}_{1,2} & \cdots & \tilde{u}_{1,N} \\
\tilde{u}_{2,1} & \tilde{u}_{2,2} & \cdots & \tilde{u}_{2,N} \\
\vdots & \vdots & \vdots & \vdots \\
\tilde{u}_{F,1} & \tilde{u}_{F,2} & \cdots & \tilde{u}_{F,N} \\
\tilde{v}_{1,1} & \tilde{v}_{1,2} & \cdots & \tilde{v}_{1,N} \\
\tilde{v}_{2,1} & \tilde{v}_{2,2} & \cdots & \tilde{v}_{2,N} \\
\vdots & \vdots & \vdots & \vdots \\
\tilde{v}_{F,1} & \tilde{v}_{F,2} & \cdots & \tilde{v}_{F,N}
\end{bmatrix}}_{W_{2F\times N}}
\end{array}
=
\underbrace{\begin{bmatrix}
\mathbf{i}_1^T \\ \mathbf{i}_2^T \\ \vdots \\ \mathbf{i}_F^T \\ \mathbf{j}_1^T \\ \mathbf{j}_2^T \\ \vdots \\ \mathbf{j}_F^T
\end{bmatrix}}_{M_{2F\times 3}}
\cdot
\begin{array}{c}
\begin{array}{cccc} \text{Point 1} & \text{Point 2} & & \text{Point } N \end{array} \\
\underbrace{[P_1 \quad P_2 \quad \cdots \quad P_N]}_{S_{3\times N}}
\end{array}
$$

$$\tilde{u}_{f,p} = \mathbf{i}_f^T P_p$$

$$\tilde{v}_{f,p} = \mathbf{j}_f^T P_p$$

$$\Rightarrow \quad \begin{bmatrix} \tilde{u}_{f,p} \\ \tilde{v}_{f,p} \end{bmatrix} = \begin{bmatrix} \mathbf{i}_f^T \\ \mathbf{j}_f^T \end{bmatrix} P_p$$

$W_{2F\times N}$: Observation Matrix (known)

$M_{2F\times 3}$: Camera Motion (unknown)

$S_{3\times N}$: Scene Structure (unknown)

SfM: Find M and S from W

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

How can we solve for M and S?

Answer: Exploit matrix Rank and Singular Value Decomposition

Matrix Rank and its properties:

Column Rank: The number of linearly independent columns of the matrix → $ColumnRank(A_{m\times n}) \leq n$

Row Rank: The number of linearly independent rows of the matrix → $RowRank(A_{m\times n}) \leq m$

Examples

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & 7 \\ 3 & 4 & 10 \\ 4 & 5 & 13 \end{bmatrix}$$

$A_3^c = 2A_1^c + A_2^c$

$ColumnRank(A) = 2$

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$ColumnRank(B) = 3$

Full Rank

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

How can we solve for M and S?

Answer: Exploit matrix Rank and Singular Value Decomposition

Matrix Rank and its properties:

Column Rank: The number of linearly independent columns of the matrix → $ColumnRank(A_{m\times n}) \leq n$

Row Rank: The number of linearly independent rows of the matrix → $RowRank(A_{m\times n}) \leq m$

$$ColumnRank(A_{m\times n}) = RowRank(A_{m\times n}) = Rank(A_{m\times n})$$
$$Rank(A_{m\times n}) \leq \min(m, n)$$

- $Rank(A^T) = Rank(A)$
- $Rank\left(A_{m\times n}B_{n\times p}\right) = $
  $= \min\left(Rank(A_{m\times n}), Rank\left(B_{n\times p}\right)\right)$
  $\leq \min(m, n, p)$
- $Rank(B_{m\times m}A_{m\times n}) = Rank(A_{m\times n})$,
  if B is full-Rank

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

Let's apply these Rank properties in our case:

$$Rank(W) = Rank(MS) \leq \min(3, N, 2F)$$

$$\boxed{Rank(W) \leq 3}$$

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

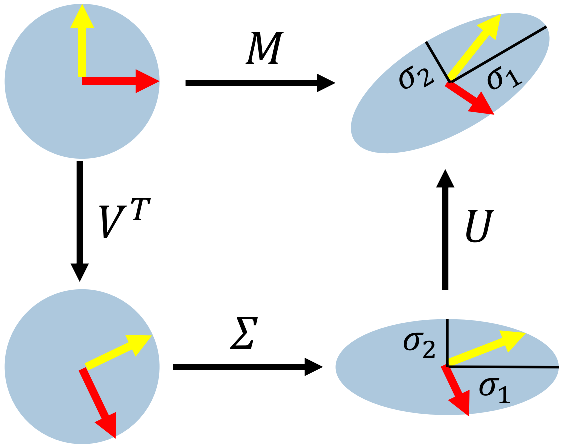
$Rank(W) \leq 3$

One more step: Singular Value Decomposition (SVD)

For any matrix M, there exists a factorization:

$$M_{m\times n} = U_{m\times m} \cdot \Sigma_{m\times n} \cdot V_{n\times n}^T$$

where U and V^T are orthonormal and Σ is diagonal.



$$\Sigma_{m\times n} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & \sigma_n \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \end{bmatrix}$$

$\sigma_1, \dots, \sigma_n$: Singular Values
- Non-negative
- Descending order of importance (highest → lowest)

If $Rank(M) = r$, then Σ has r non-zero singular values

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

Using SVD: $W = U \cdot \Sigma \cdot V^T$

$$W = \underbrace{\begin{bmatrix} & & \\ & U & \\ & & \end{bmatrix}}_{2F\times 2F} \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3 & 0 & \dots & 0 \\ 0 & 0 & 0 & \sigma_4 & \dots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & \sigma_N \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \end{bmatrix}}_{2F\times N} \underbrace{\begin{bmatrix} & & \\ & V^T & \\ & & \end{bmatrix}}_{N\times N}$$

$\boxed{Rank(W) \leq 3}$

If $Rank(M) = r$, then Σ has r non-zero singular values

Where $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \cdots \geq \sigma_N$ are the singular values of W.

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

Using SVD: $W = U \cdot \Sigma \cdot V^T$

$$W = \underbrace{\begin{bmatrix} & & \\ & U & \\ & & \end{bmatrix}}_{2F\times 2F} \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3 & 0 & \dots & 0 \\ 0 & 0 & 0 & \sigma_4 & \dots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & \sigma_N \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \end{bmatrix}}_{2F\times N} \underbrace{\begin{bmatrix} & & \\ & V^T & \\ & & \end{bmatrix}}_{N\times N}$$

$Rank(W) \leq 3$

If $Rank(M) = r$, then Σ has r non-zero singular values

$$Rank(W) \leq 3 \Rightarrow Rank(\Sigma) \leq 3$$

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

Using SVD: $W = U \cdot \Sigma \cdot V^T$

$$W = \underbrace{\begin{bmatrix} & & \\ & U & \\ & & \end{bmatrix}}_{2F\times 2F} \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3 & 0 & \dots & 0 \\ 0 & 0 & 0 & \mathbf{0} & \dots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & \mathbf{0} \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \end{bmatrix}}_{2F\times N} \underbrace{\begin{bmatrix} & & \\ & V^T & \\ & & \end{bmatrix}}_{N\times N}$$

$Rank(W) \leq 3$

If $Rank(M) = r$, then Σ has r non-zero singular values

$$Rank(W) \leq 3 \Rightarrow Rank(\Sigma) \leq 3 \Rightarrow \sigma_4, \dots, \sigma_N = 0$$

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

Using SVD: $W = U \cdot \Sigma \cdot V^T$

$$W = \underbrace{\begin{bmatrix} \\ & U & \\ \\ \end{bmatrix}}_{2F \times 2F} \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3 & 0 & \dots & 0 \\ 0 & 0 & 0 & \sim\mathbf{0} & \dots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & \sim\mathbf{0} \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \end{bmatrix}}_{2F \times N} \underbrace{\begin{bmatrix} \\ & V^T & \\ \\ \end{bmatrix}}_{N \times N}$$

$Rank(W) \leq 3$

If $Rank(M) = r$, then Σ has r non-zero singular values

$$Rank(W) \leq 3 \Rightarrow Rank(\Sigma) \leq 3 \Rightarrow \sigma_4, \dots, \sigma_N = 0$$

Yet, there is noise

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

Using SVD: $W = U \cdot \Sigma \cdot V^T$

$$W = \underbrace{\begin{array}{cc} \left[\begin{array}{c|c} U_1 & \cancel{U_2} \end{array}\right] \\ \;3 \qquad 2F-3 \end{array}}_{2F\times 2F} \underbrace{\begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & 0 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3 & 0 & \dots & 0 \\ 0 & 0 & 0 & \mathbf{0} & \dots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & \mathbf{0} \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \end{bmatrix}}_{2F\times N} \underbrace{\begin{bmatrix} V_1^T \\ \cancel{V_2^T} \end{bmatrix}}_{N\times N} \begin{array}{l} 3 \\ N-3 \end{array}$$

$Rank(W) \leq 3$

If $Rank(M) = r$, then Σ has r non-zero singular values

$$\underset{(2F\times N)}{W} = \underset{(2F\times 3)(3\times 3)(3\times N)}{U_1 \cdot \Sigma_1 \cdot V_1^T}$$

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

$$W = U_1 \cdot \Sigma_1 \cdot V_1^T \Rightarrow$$

$$W = \underbrace{\boxed{U_1 \cdot {\Sigma_1}^{1/2}}}_{2F\times 3} \cdot \underbrace{\boxed{{\Sigma_1}^{1/2} \cdot V_1^T}}_{3\times N}$$

Is this M? Is this S?

Not necessarily! Factorization is not unique.

$$W = \underbrace{\boxed{U_1 \cdot {\Sigma_1}^{1/2} \cdot Q}}_{\substack{2F\times 3 \\ = M}} \cdot \underbrace{\boxed{Q^{-1} \cdot {\Sigma_1}^{1/2} \cdot V_1^T}}_{\substack{3\times N \\ = S}}$$ is also valid.

We need to find $Q_{3\times 3}$ and we are finally done.

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

Exploit the structure of motion matrix M:

$$\underset{2F\times 3}{M} = \underset{2F\times 3}{\begin{bmatrix} \mathbf{i}_1^T \\ \mathbf{i}_2^T \\ \vdots \\ \mathbf{i}_F^T \\ \mathbf{j}_1^T \\ \mathbf{j}_2^T \\ \vdots \\ \mathbf{j}_F^T \end{bmatrix}} = \underset{\substack{\text{Computed} \\ 2F\times 3}}{\boxed{U_1 {\Sigma_1}^{1/2}}} Q = \underset{\substack{\text{Computed} \\ 2F\times 3}}{\boxed{\begin{bmatrix} \hat{\mathbf{i}}_1^T \\ \hat{\mathbf{i}}_2^T \\ \vdots \\ \hat{\mathbf{i}}_F^T \\ \hat{\mathbf{j}}_1^T \\ \hat{\mathbf{j}}_2^T \\ \vdots \\ \hat{\mathbf{j}}_F^T \end{bmatrix}}} Q = \underset{2F\times 3}{\begin{bmatrix} \hat{\mathbf{i}}_1^T Q \\ \hat{\mathbf{i}}_2^T Q \\ \vdots \\ \hat{\mathbf{i}}_F^T Q \\ \hat{\mathbf{j}}_1^T Q \\ \hat{\mathbf{j}}_2^T Q \\ \vdots \\ \hat{\mathbf{j}}_F^T Q \end{bmatrix}}$$

Orthonormality Constrains:

$$\begin{aligned} \mathbf{i}_f \cdot \mathbf{i}_f &= \mathbf{i}_f^T \mathbf{i}_f = 1 \\ \mathbf{j}_f \cdot \mathbf{j}_f &= \mathbf{j}_f^T \mathbf{j}_f = 1 \\ \mathbf{i}_f \cdot \mathbf{j}_f &= \mathbf{i}_f^T \mathbf{j}_f = 0 \end{aligned} \quad \Rightarrow \quad \begin{aligned} \hat{\mathbf{i}}_f^T Q Q^T \hat{\mathbf{i}}_f &= 1 \\ \hat{\mathbf{j}}_f^T Q Q^T \hat{\mathbf{j}}_f &= 1 \\ \hat{\mathbf{i}}_f^T Q Q^T \hat{\mathbf{j}}_f &= 0 \end{aligned}$$

Structure from Motion – SfM: The early-days approach

$$W_{2F\times N} = M_{2F\times 3} \cdot S_{3\times N}$$

$W_{2F\times N}$: Observation Matrix
$M_{2F\times 3}$: Camera Motion
$S_{3\times N}$: Scene Structure

We have computed $(\hat{\mathbf{i}}_f^T, \hat{\mathbf{j}}_f)$ for frame $f = 1, \ldots, F$:

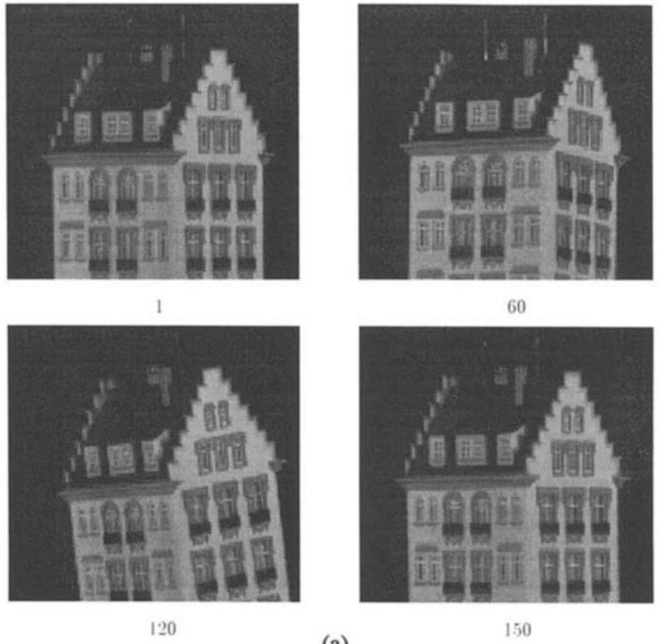
$$\left.\begin{aligned} \hat{\mathbf{i}}_f^T Q Q^T \hat{\mathbf{i}}_f &= 1 \\ \hat{\mathbf{j}}_f^T Q Q^T \hat{\mathbf{j}}_f &= 1 \\ \hat{\mathbf{i}}_f^T Q Q^T \hat{\mathbf{j}}_f &= 0 \end{aligned}\right\} \quad Q \text{ is unknown.}$$

- Q is a 3×3 matrix with 9 variables
- We have $3F$ equations that involve Q
- As long as we have 3 or more frames ($F \geq 3$) this can be solved (e.g., with Newton's method)

Final solution:

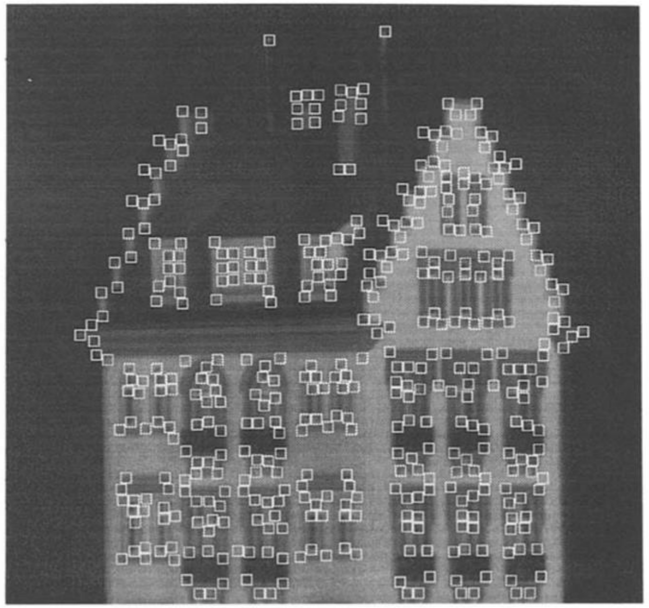
$$M = U_1 \cdot {\Sigma_1}^{1/2} \cdot Q$$
$$S = Q^{-1} \cdot {\Sigma_1}^{1/2} \cdot V_1^T$$

Structure from Motion – SfM: The early-days approach



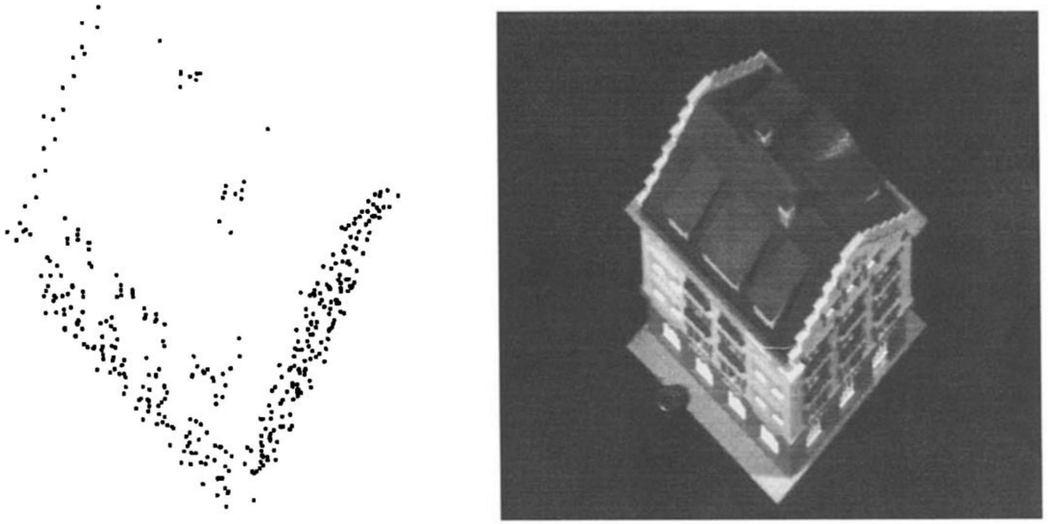
(a)

Fig. 2a. The “Hotel” stream: four of the 150 frames.



(b)

Fig. 2b. The “Hotel” stream: the 430 features selected by the automatic detection method



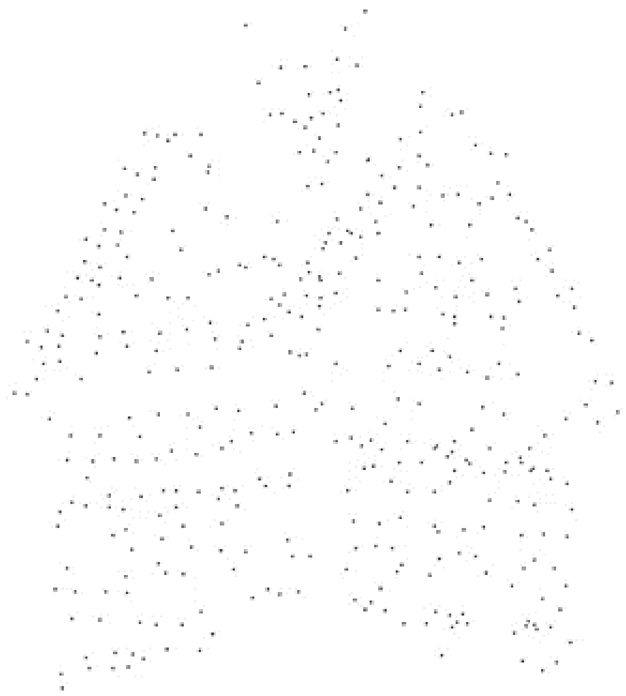
(a) (b)

Fig. 4. Qualitative shape results for the “Hotel” stream: top view of the (a) computed and (b) actual shape.

Carlo Tomasi, Takeo Kanade, **Shape and Motion from Image Streams under Orthography: a Factorization Method** International Journal of Computer Vision, IJCV, 1992, pages 137-154, Springer

Localization and Mapping for Autonomous Mobile Systems

Structure from Motion – SfM: The early-days approach



Carlo Tomasi, Takeo Kanade, **Shape and Motion from Image Streams under Orthography: a Factorization Method** International Journal of Computer Vision, IJCV, 1992, pages 137-154, Springer

SLAM:: The online and real-time version

Now, what’s the issue with that? Why not to use such approaches in robotics applications?

- We cannot wait for all the frames to be captured
- Even then, we cannot expect to process all these frames in real-time