Structure From Motion

Compute 3D scene structure and camera motion from a sequence of frames.

Topics:

(1) Structure from Motion Problem

(2) SFM Observation Matrix

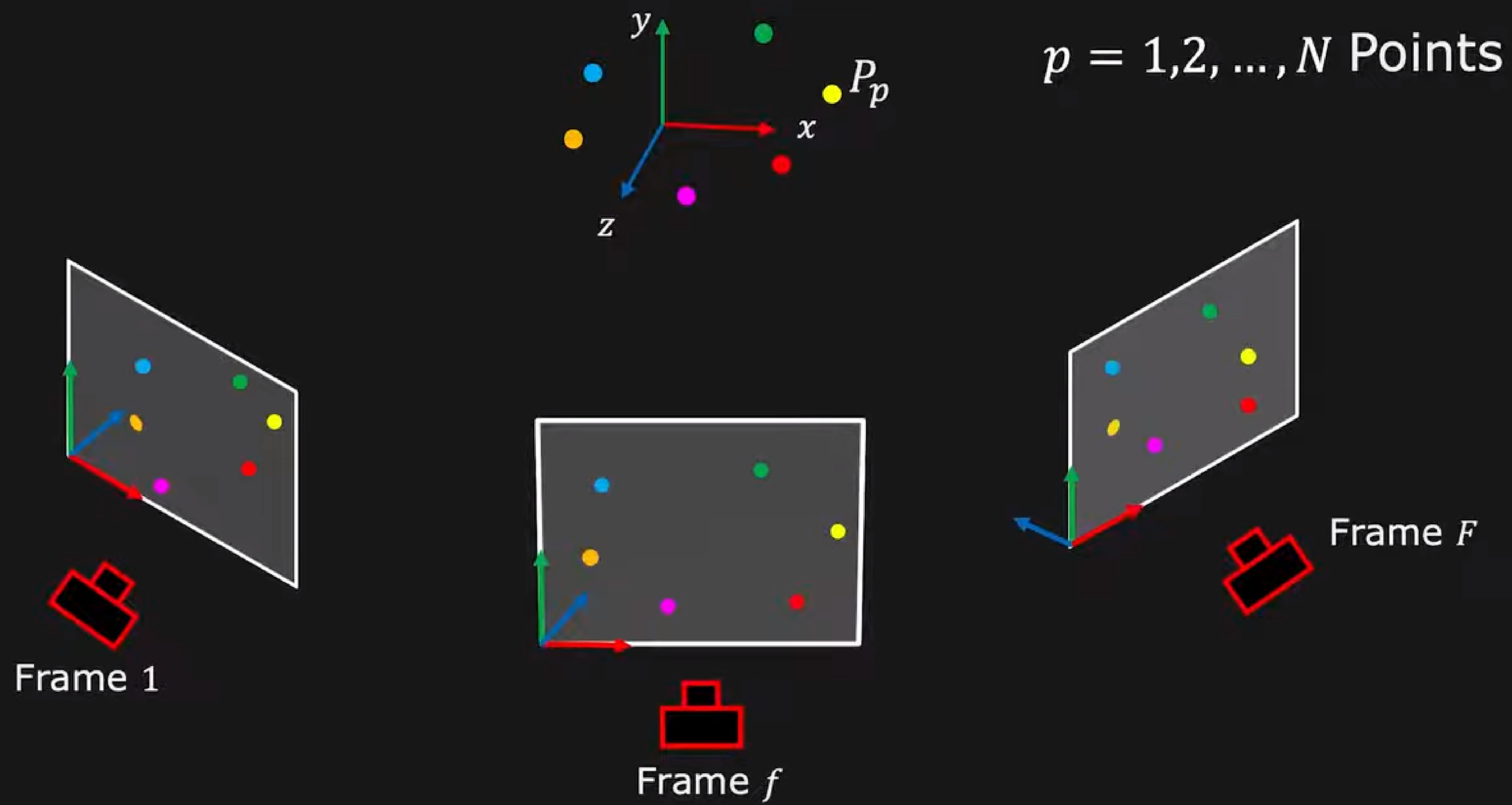
(3) Rank of Observation Matrix

(4) Tomasi-Kanade Factorization

Feature Detection and Tracking

- Detect feature points: Corners, SIFT points, …
- Track feature points: Template Matching, Optical Flow…

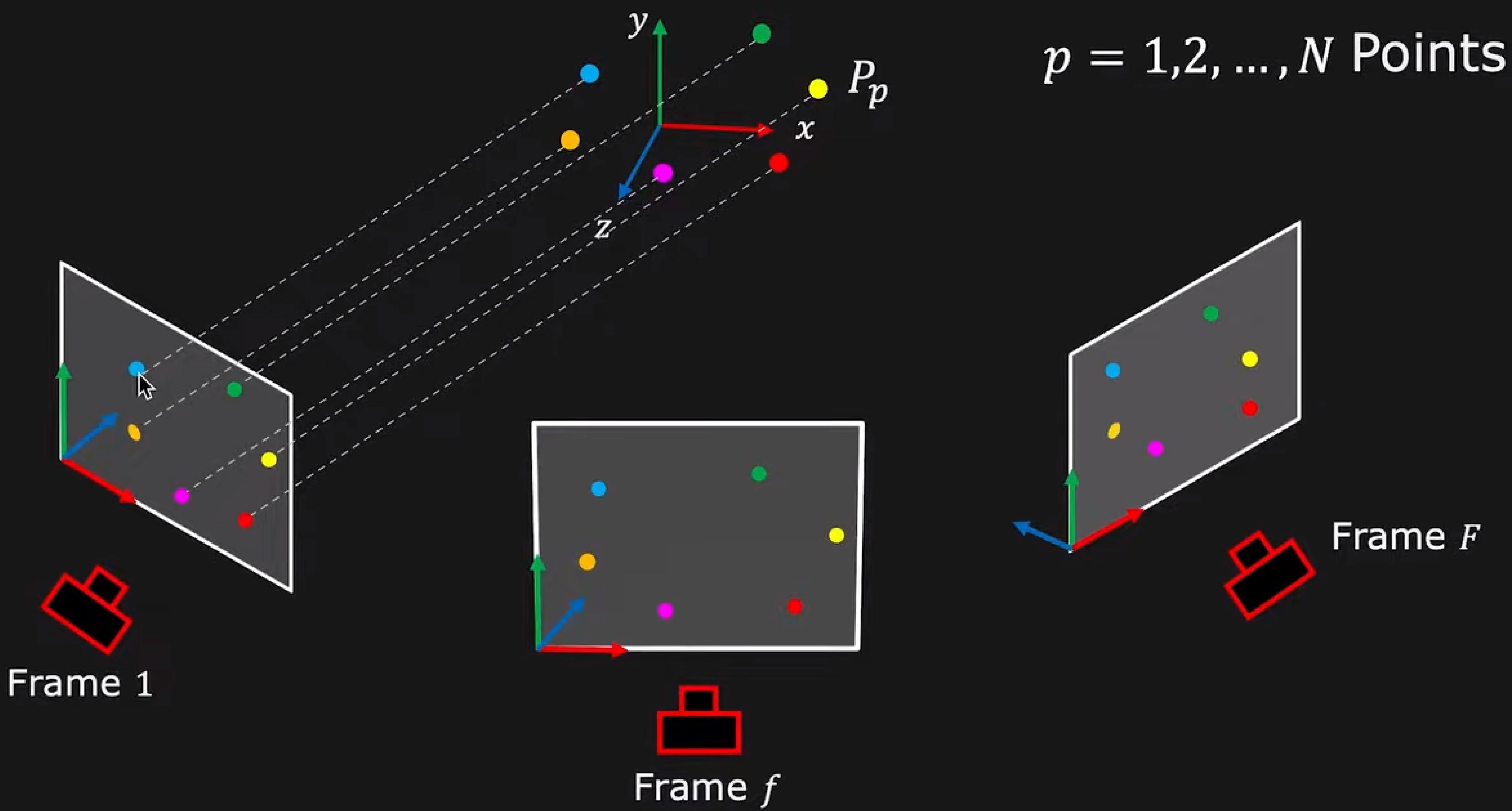
Structure From Motion



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

Find scene points (3D): P_p

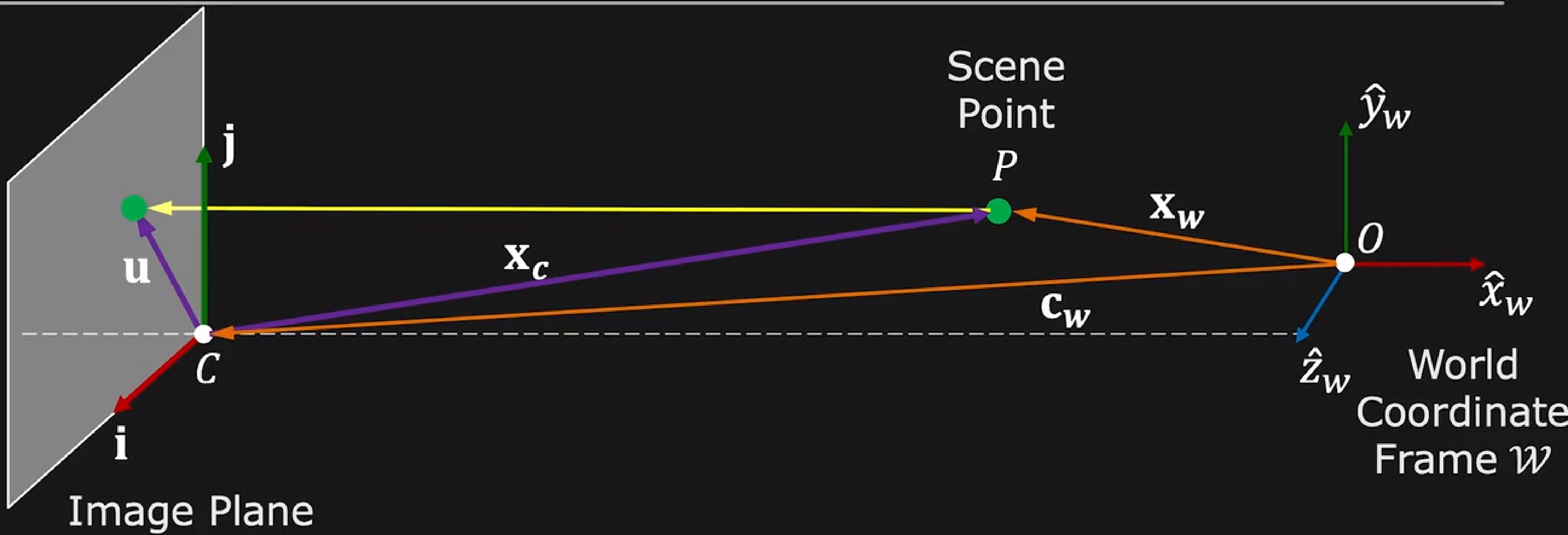
Orthographic Structure from Motion



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

Find scene points (3D) P_p, assuming orthographic camera.

From 3D to 2D: Orthographic Projection

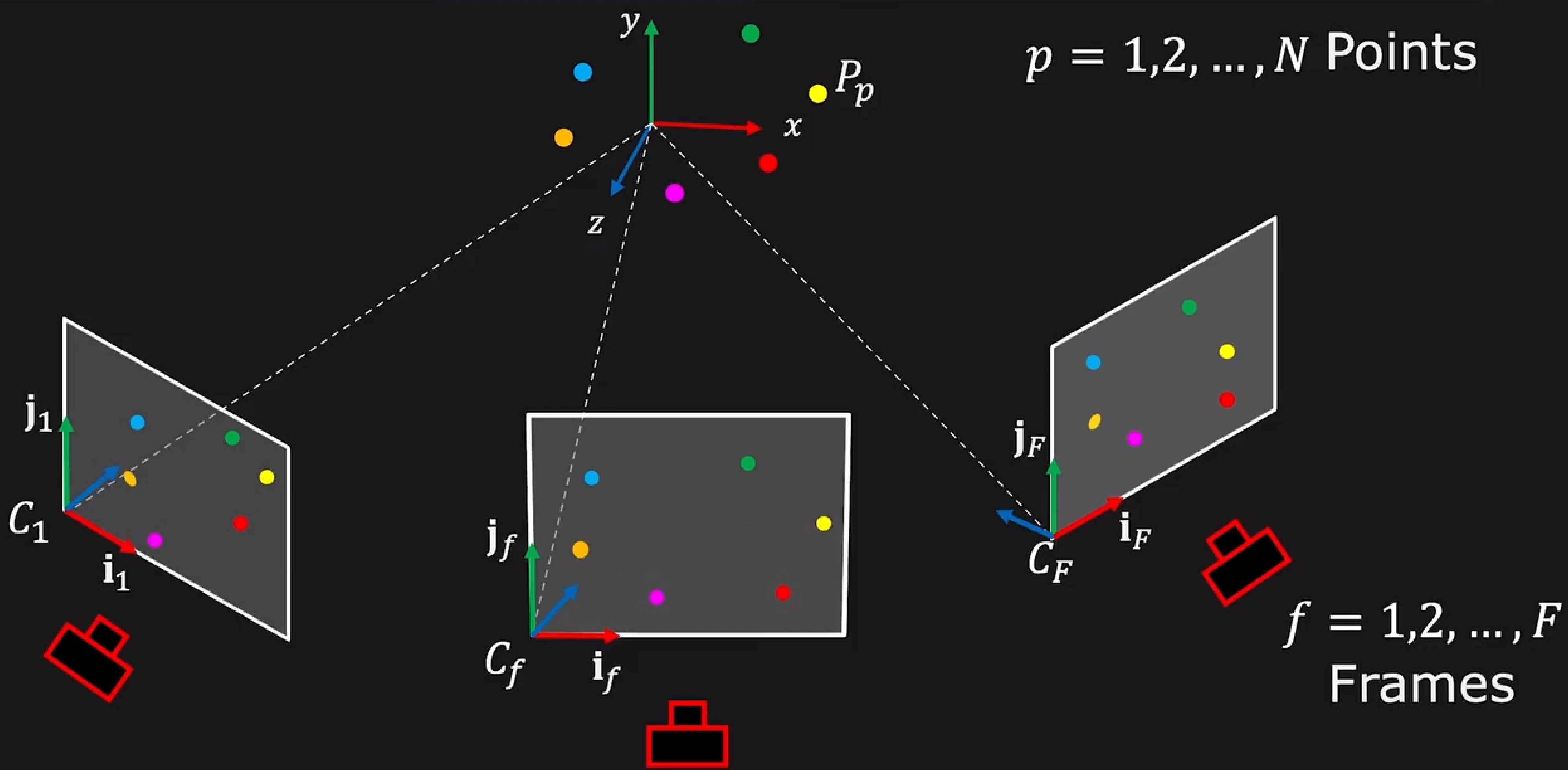


$$u = \mathbf{i}^T \mathbf{x}_c = \mathbf{i}^T(\mathbf{x}_w - \mathbf{c}_w) = \mathbf{i}^T(P - C)$$

$$v = \mathbf{j}^T \mathbf{x}_c = \mathbf{j}^T(\mathbf{x}_w - \mathbf{c}_w) = \mathbf{j}^T(P - C)$$

$$\boxed{\begin{aligned} u &= \mathbf{i}^T(P - C) \\ v &= \mathbf{j}^T(P - C) \end{aligned}}$$

Orthographic SFM



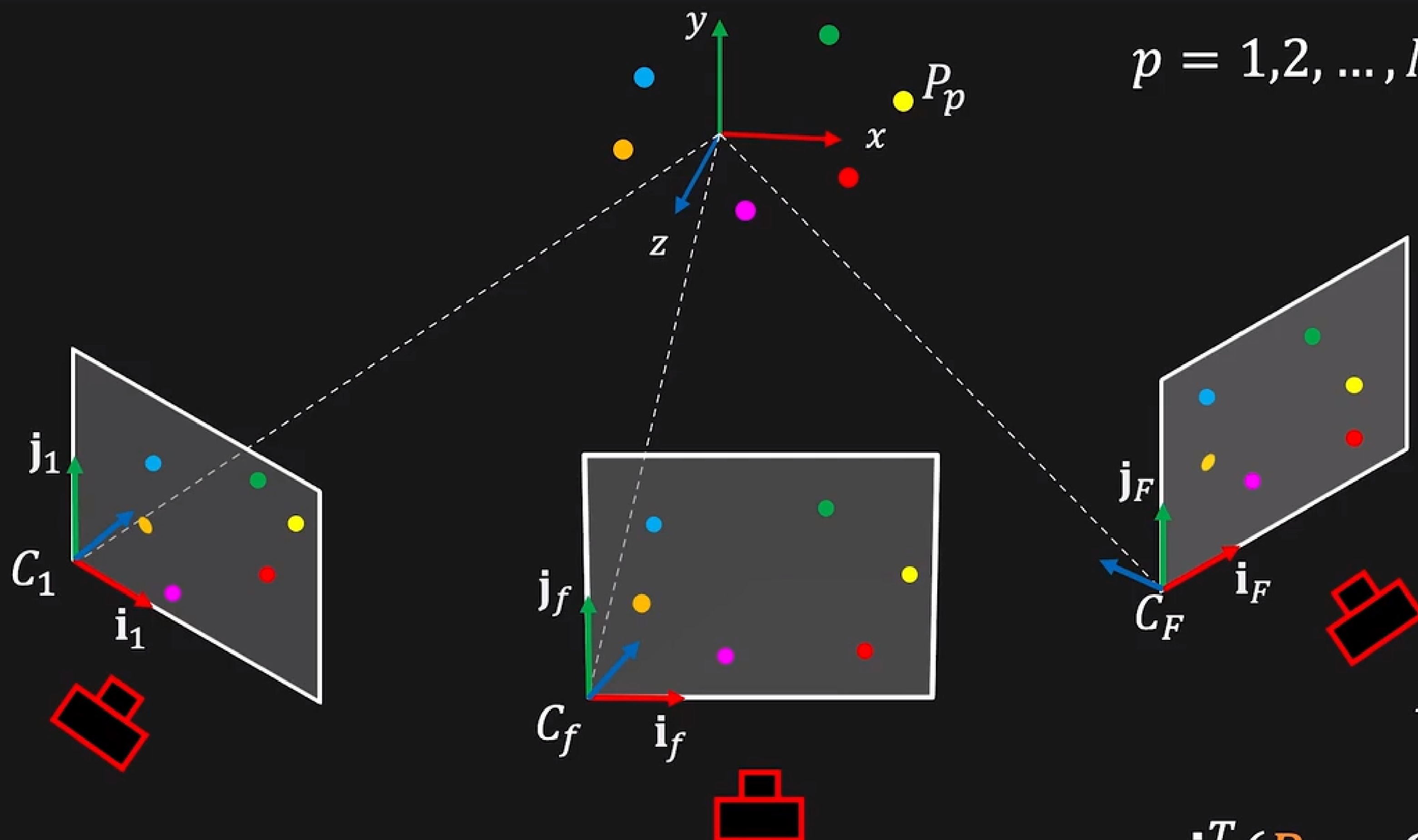
Given corresponding image points (2D) $(u_{f,p}, v_{f,p})$

Find scene points $\{P_p\}$.

Camera Positions $\{C_f\}$, camera orientations $\{(\mathbf{i}_f, \mathbf{j}_f)\}$ are unkno

Orthographic SFM

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



$f = 1, 2, \ldots, F$
Frames

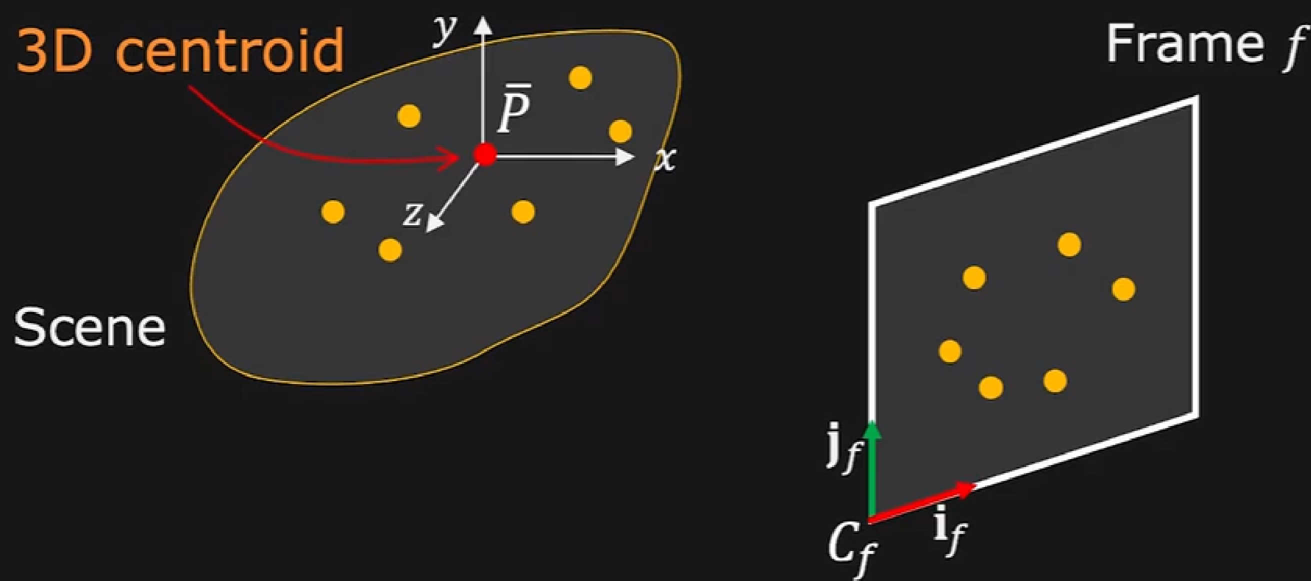
Image of point P_p in camera frame f:

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

$$\underbrace{v_{f,p}}_{\text{Known}} = \underbrace{\mathbf{j}_f^T (P_p - C_f)}_{\text{Unknown}}$$

We can remove C_f from equations to simply SFM problem.

Centering Trick

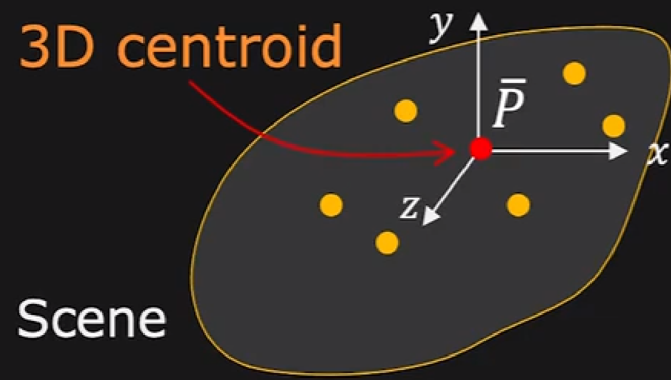


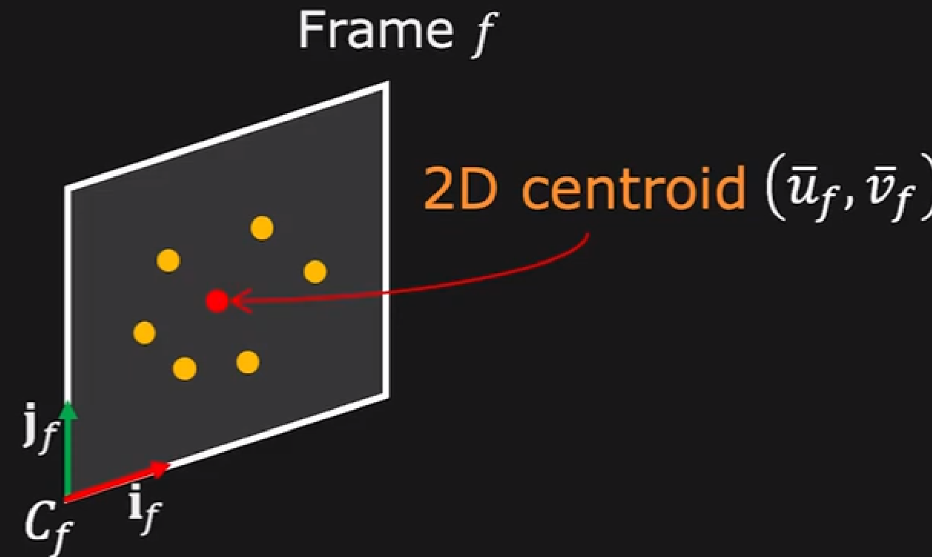
Assume origin of world at centroid of scene points:

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

We will compute scene points w.r.t their centroid!

Centering Trick





Centroid $(\bar{u}_f, \bar{v}_f)$ of the image points in frame f:

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

$$\bar{v}_f = \frac{1}{N}\sum_{p=1}^{N} v_{f,p} = \frac{1}{N}\sum_{p=1}^{N} \mathbf{j}_f^T(P_p - C_f)$$

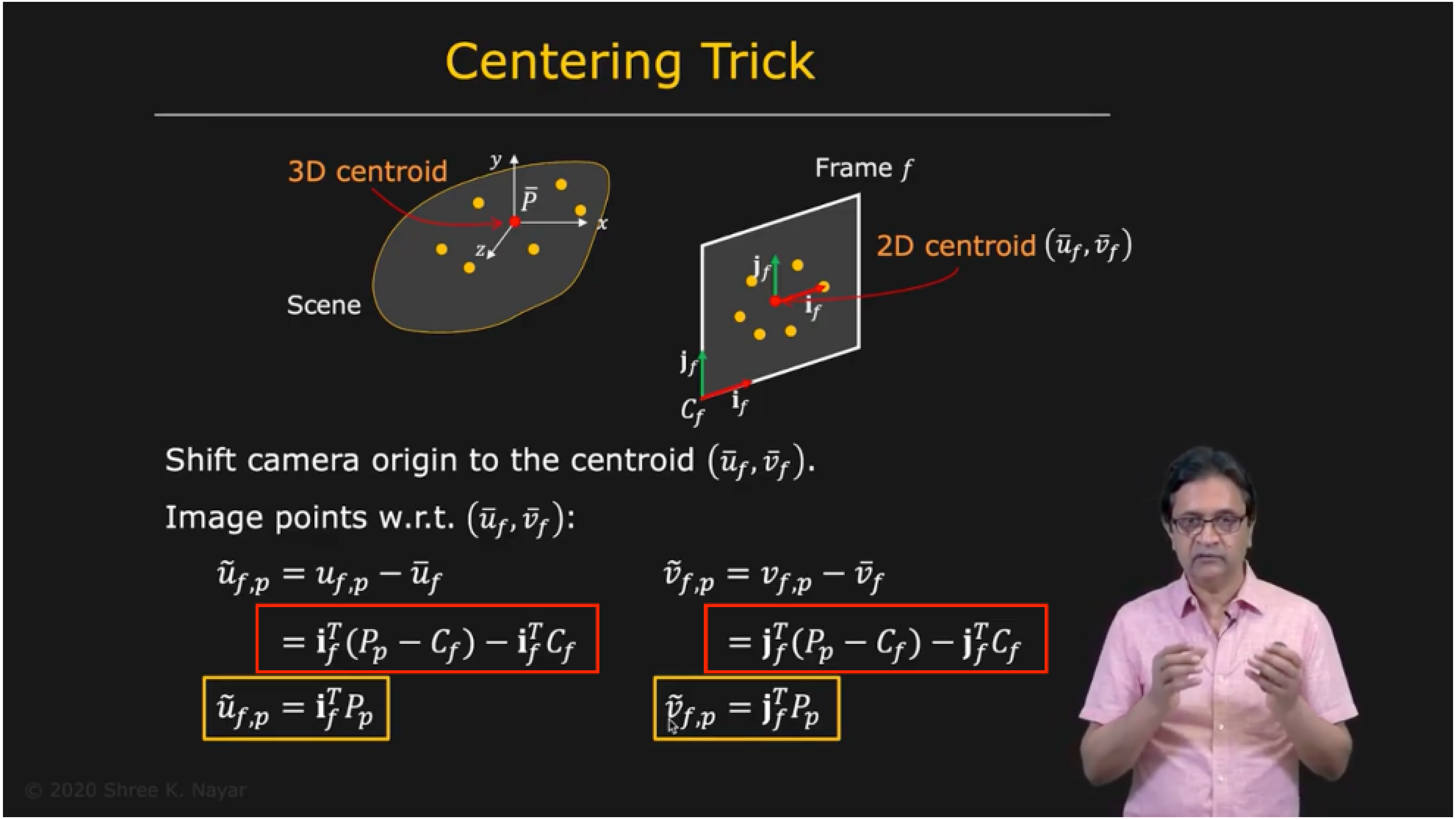
$$\bar{u}_f = \cancel{\frac{1}{N}\mathbf{i}_f^T\sum_{p=1}^{N} P_p} - \frac{1}{N}\sum_{p=1}^{N} \mathbf{i}_f^T C_f$$

$$\bar{v}_f = \cancel{\frac{1}{N}\mathbf{j}_f^T\sum_{p=1}^{N} P_p} - \frac{1}{N}\sum_{p=1}^{N} \mathbf{j}_f^T C_f$$

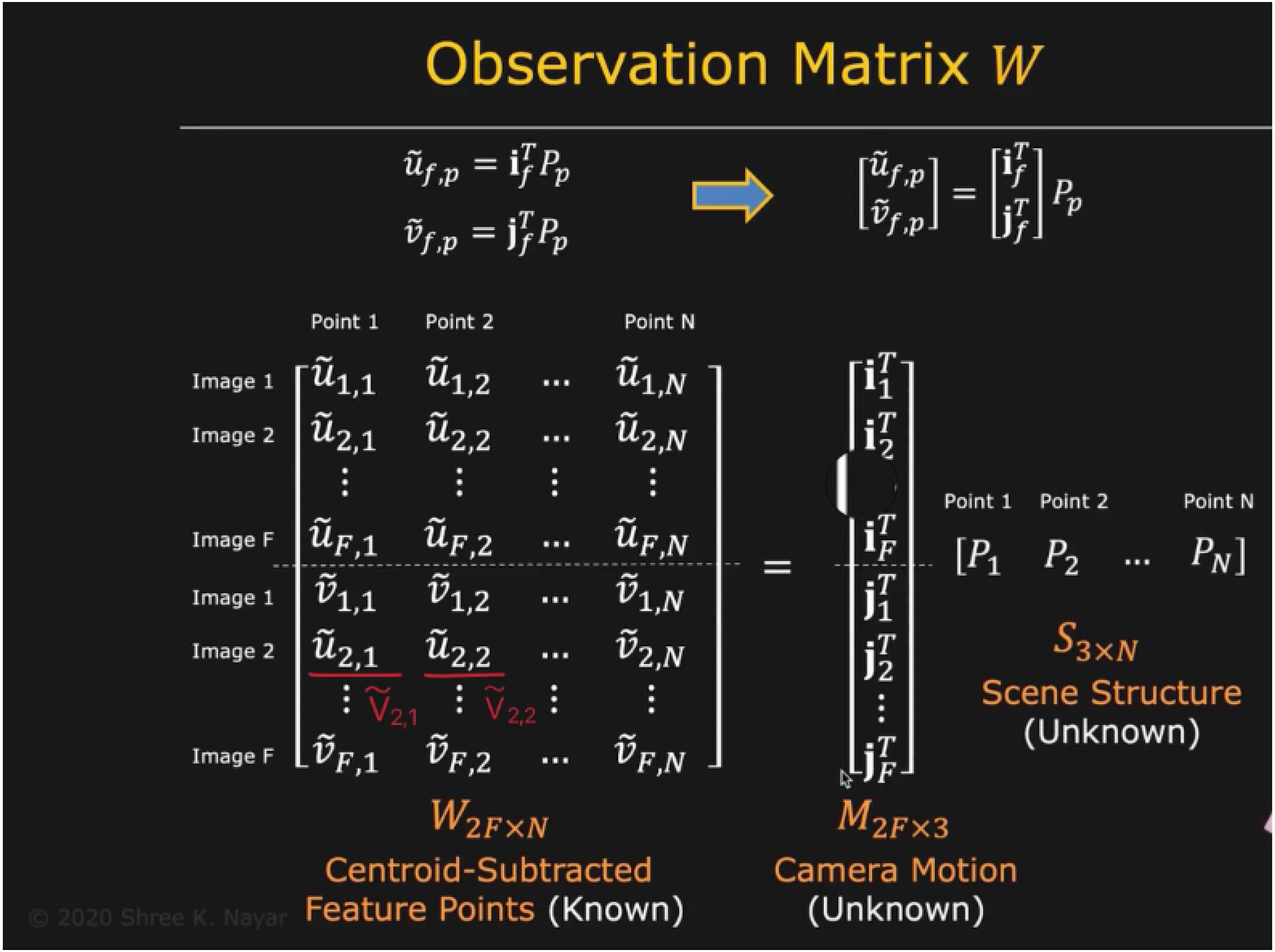
$$\boxed{\bar{u}_f = -\mathbf{i}_f^T C_f}$$

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

In Video 5.3 Observation Matrix, from 10:40 to 11:13, the equations of $\tilde{u}_{f,p}$ and $\tilde{v}_{f,p}$ should be replaced by $= i_f^T(P_p - C_f) - (-i_f^T C_f)$ and $= j_f^T(P_p - C_f) - (-j_f^T C_f)$.



From 12:05 to 14:08, point 1 and point 2 in the penultimate row of the matrix form in the slide should be $\tilde{v}_{2,1}$ and $\tilde{v}_{2,2}$ instead of $\tilde{u}_{2,1}$ and $\tilde{u}_{2,2}$.



Observation Matrix W

$$
\begin{array}{rcccc}
 & \text{Point 1} & \text{Point 2} & & \text{Point N} \\
\begin{array}{r} \text{Image 1} \\ \text{Image 2} \\ \\ \text{Image F} \\ \text{Image 1} \\ \text{Image 2} \\ \\ \text{Image F} \end{array} &
\multicolumn{4}{c}{\begin{bmatrix}
\tilde{u}_{1,1} & \tilde{u}_{1,2} & \cdots & \tilde{u}_{1,N} \\
\tilde{u}_{2,1} & \tilde{u}_{2,2} & \cdots & \tilde{u}_{2,N} \\
\vdots & \vdots & \vdots & \vdots \\
\tilde{u}_{F,1} & \tilde{u}_{F,2} & \cdots & \tilde{u}_{F,N} \\
\tilde{v}_{1,1} & \tilde{v}_{1,2} & \cdots & \tilde{v}_{1,N} \\
\tilde{u}_{2,1} & \tilde{u}_{2,2} & \cdots & \tilde{v}_{2,N} \\
\vdots & \vdots & \vdots & \vdots \\
\tilde{v}_{F,1} & \tilde{v}_{F,2} & \cdots & \tilde{v}_{F,N}
\end{bmatrix}}
\end{array}
=
\begin{bmatrix}
\mathbf{i}_1^T \\ \mathbf{i}_2^T \\ \vdots \\ \mathbf{i}_F^T \\ \mathbf{j}_1^T \\ \mathbf{j}_2^T \\ \vdots \\ \mathbf{j}_F^T
\end{bmatrix}
\begin{array}{cccc}
\text{Point 1} & \text{Point 2} & & \text{Point N} \\
\multicolumn{4}{c}{\begin{bmatrix} P_1 & P_2 & \cdots & P_N \end{bmatrix}}
\end{array}
$$

$W_{2F\times N}$

Centroid-Subtracted Feature Points (Known)

$M_{2F\times 3}$

Camera Motion (Unknown)

$S_{3\times N}$

Scene Structure (Unknown)

Can we find M and S from W?

Linear Independence of Vectors

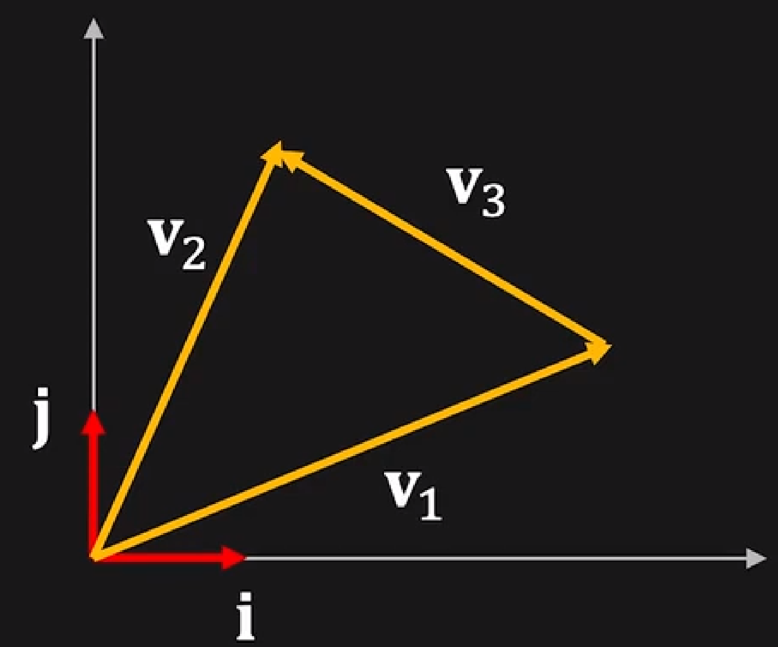
A set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is said to be linearly independent if no vector can be represented as a weighted linear sum of the others.

$\{\mathbf{i}, \mathbf{j}\}$ is linearly independent.

$\{\mathbf{i}, \mathbf{j}, \mathbf{v}_1\}$ is linearly dependent.

$\{\mathbf{i}, \mathbf{j}, \mathbf{v}_3\}$ is linearly dependent.

$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly dependent.



Rank of a Matrix

Column Rank: The number of linearly independent columns of the matrix.

Row Rank: The number of linearly independent rows of the matrix.

$$
\begin{array}{cc} m & \begin{bmatrix} \\ \quad A \quad \\ \\ \end{bmatrix} \\ & n \end{array} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \vdots \\ \mathbf{r}_m^T \end{bmatrix}
$$

$$ColumnRank(A) \leq n \qquad\qquad RowRank(A) \leq m$$

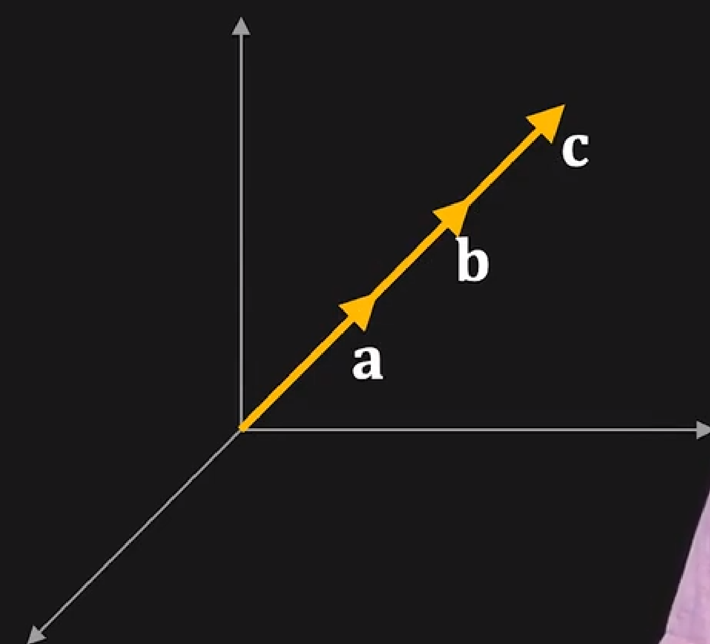
$$\boxed{\begin{gathered} ColumnRank(A) = RowRank(A) = Rank(A) \\ Rank(A) \leq \min(m, n) \end{gathered}}$$

Geometric Meaning of Matrix Rank

Rank is the dimensionality of the space spanned by column or row vectors of the matrix.

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = [\mathbf{a} \quad \mathbf{b} \quad \mathbf{c}]$$

$$Rank(A) = 1$$

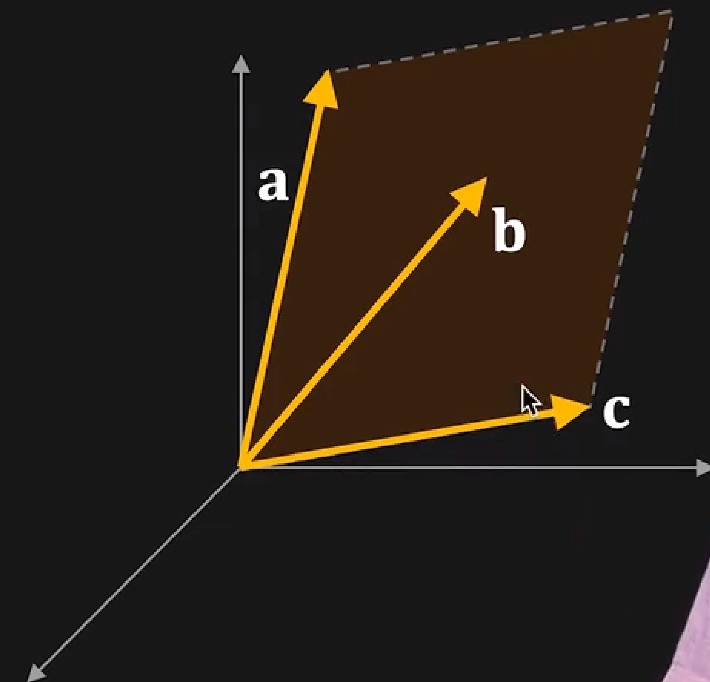


Geometric Meaning of Matrix Rank

Rank is the dimensionality of the space spanned by column or row vectors of the matrix.

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = [\mathbf{a} \quad \mathbf{b} \quad \mathbf{c}]$$

$$Rank(A) = 2$$

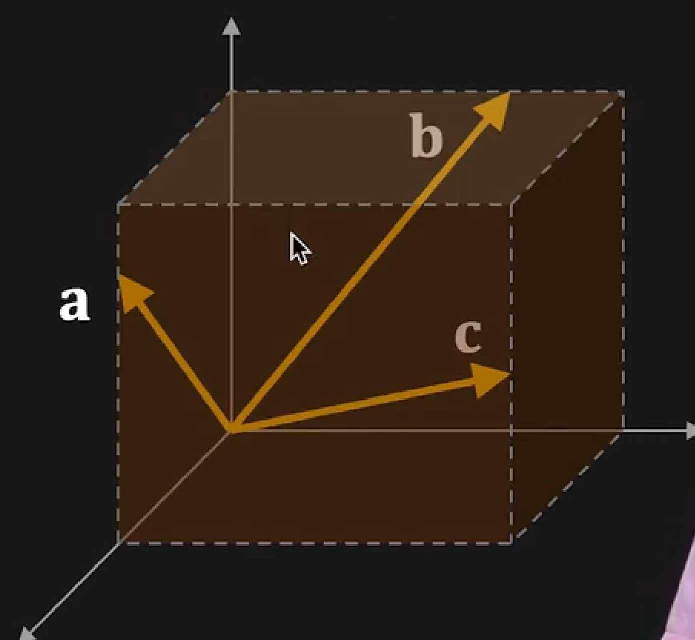


Geometric Meaning of Matrix Rank

Rank is the dimensionality of the space spanned by column or row vectors of the matrix.

$$A = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = [\mathbf{a} \quad \mathbf{b} \quad \mathbf{c}]$$

$$Rank(A) = 3 \quad \Leftrightarrow$$



Important Properties of Matrix Rank

- $Rank(A^T) = Rank(A)$

- $$Rank\left(A_{m\times n}B_{n\times p}\right) = \min\left(Rank(A_{m\times n}), Rank\left(B_{n\times p}\right)\right)$$
  $$\leq \min(m, n, p)$$

- $Rank(AA^T) = Rank(A^T A) = Rank(A^T) = Rank(A)$

- $A_{m\times m}$ is invertible iff $Rank(A_{m\times m}) = m$

...Back to Observation Matrix W

$$
\begin{array}{rc}
 & \begin{array}{cccc} \text{Point 1} & \text{Point 2} & & \text{Point N} \end{array} \\
\begin{array}{r} \text{Image 1} \\ \text{Image 2} \\ \\ \text{Image F} \\ \text{Image 1} \\ \text{Image 2} \\ \\ \text{Image F} \end{array} &
\begin{bmatrix}
\tilde{u}_{1,1} & \tilde{u}_{1,2} & \cdots & \tilde{u}_{1,N} \\
\tilde{u}_{2,1} & \tilde{u}_{2,2} & \cdots & \tilde{u}_{2,N} \\
\vdots & \vdots & \vdots & \vdots \\
\tilde{u}_{F,1} & \tilde{u}_{F,2} & \cdots & \tilde{u}_{F,N} \\
\tilde{v}_{1,1} & \tilde{v}_{1,2} & \cdots & \tilde{v}_{1,N} \\
\tilde{u}_{2,1} & \tilde{u}_{2,2} & \cdots & \tilde{v}_{2,N} \\
\vdots & \vdots & \vdots & \vdots \\
\tilde{v}_{F,1} & \tilde{v}_{F,2} & \cdots & \tilde{v}_{F,N}
\end{bmatrix}
\end{array}
=
\begin{bmatrix}
\mathbf{i}_1^T \\ \mathbf{i}_2^T \\ \vdots \\ \mathbf{i}_F^T \\ \mathbf{j}_1^T \\ \mathbf{j}_2^T \\ \vdots \\ \mathbf{j}_F^T
\end{bmatrix}
\begin{array}{c}
\begin{array}{cccc} \text{Point 1} & \text{Point 2} & & \text{Point N} \end{array} \\
\begin{bmatrix} P_1 & P_2 & \cdots & P_N \end{bmatrix}
\end{array}
$$

$S_{3\times N}$

Scene Structure (Unknown)

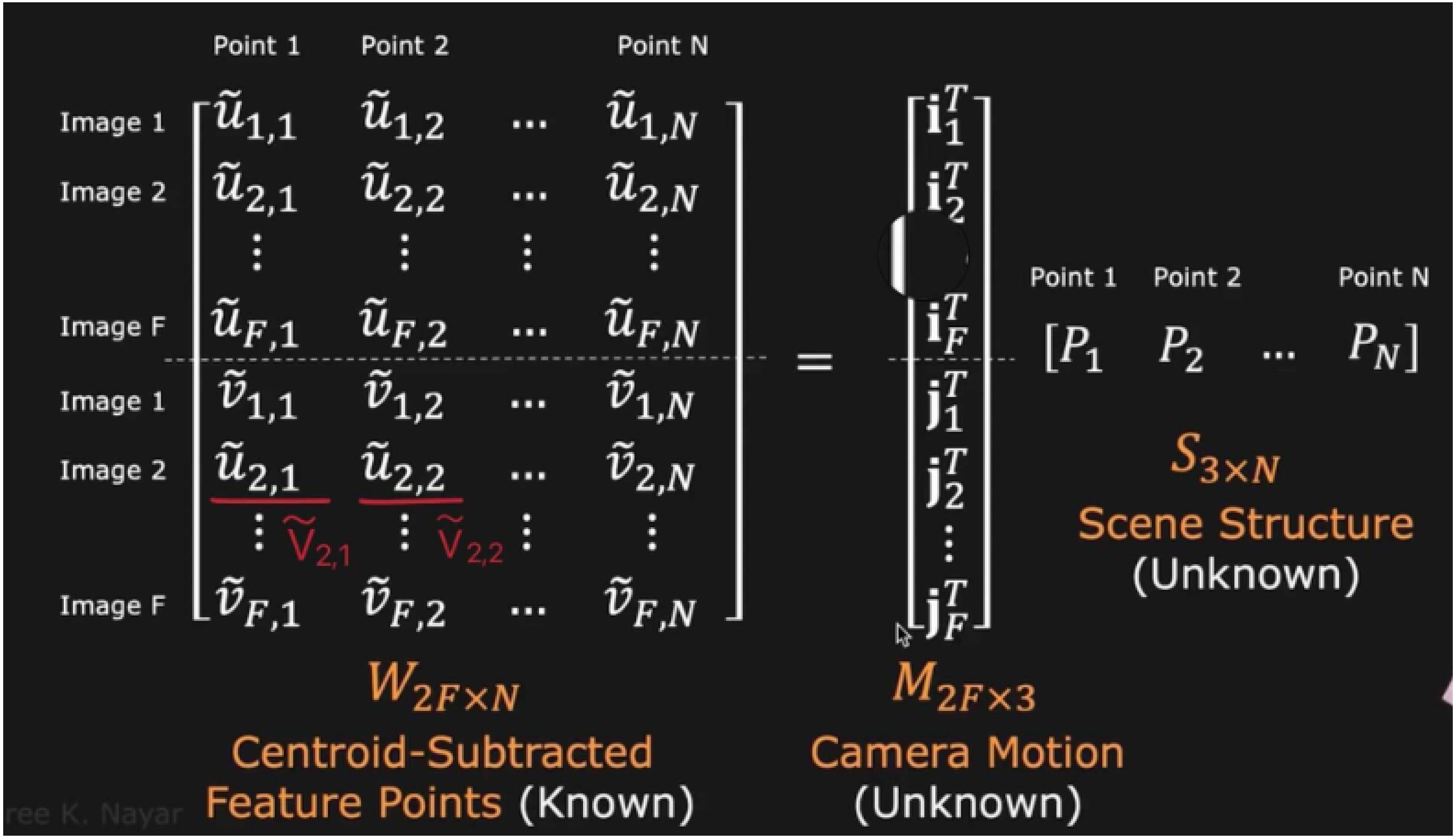
$W_{2F\times N}$

Centroid-Subtracted Feature Points (Known)

$M_{2F\times 3}$

Camera Motion (Unknown)

In [Video 5.4 Rank of Observation Matrix](), from 7:02 to 8:12, the point 1 and point 2 in the penultimate row of the matrix form in slide should be $\tilde{v}_{2,1}$ and $\tilde{v}_{2,2}$ instead of $\tilde{u}_{2,1}$ and $\tilde{u}_{2,2}$.



Rank of Observation Matrix

$$\underset{2F\times N}{W} = \underset{2F\times 3}{M} \times \underset{3\times N}{S}$$

We know:

$$Rank(MS) \le Rank(M) \qquad Rank(MS) \le Rank(S)$$

$$\Rightarrow \quad Rank(MS) \le \min(3,2F) \qquad Rank(MS) \le \min(3,N)$$

$$\Rightarrow \quad Rank(W) = Rank(MS) \le \min(3,N,2F)$$

Rank Theorem: $Rank(W) \le 3$

We can "factorize" W into M and S!

[Tomasi 1992]

e K. Nayar