ICCV OCT 19-23, 2025 HONOLULU HAWAII

A Linear N-Point Solver for Structure and Motion from Asynchronous Tracks

ICCV 2025 Highlight

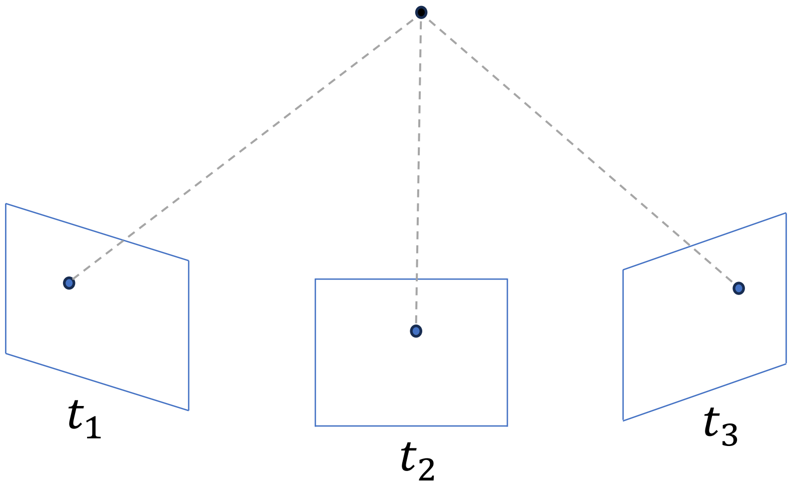
Hang Su*, Yunlong Feng*, Daniel Gehrig, Panfeng Jiang, Ling Gao, Xavier Lagorce, Laurent Kneip

上海科技大学 ShanghaiTech University

Penn Engineering UNIVERSITY of PENNSYLVANIA

Challenge

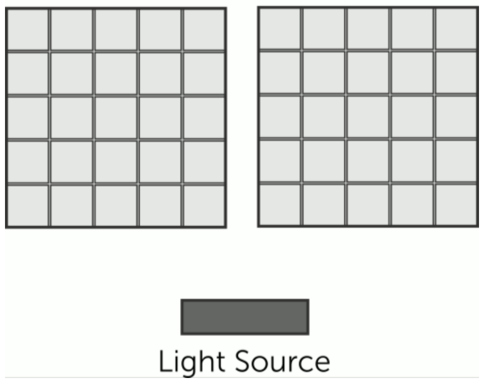
Traditional Methods



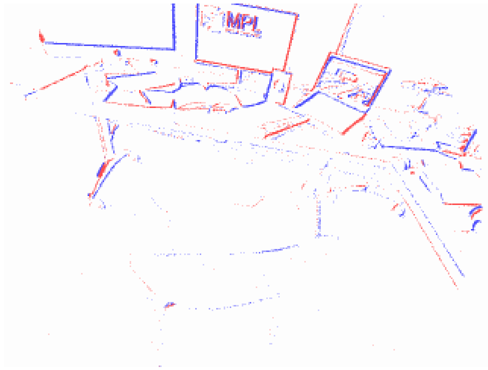
Classic solvers like the 5-point or 8-point algorithm require point track from **synchronized** views

Asynchronous Data

Rolling Shutter



Event Cameras



This assumption breaks down for rolling shutter and event cameras, which produce **asynchronous** data

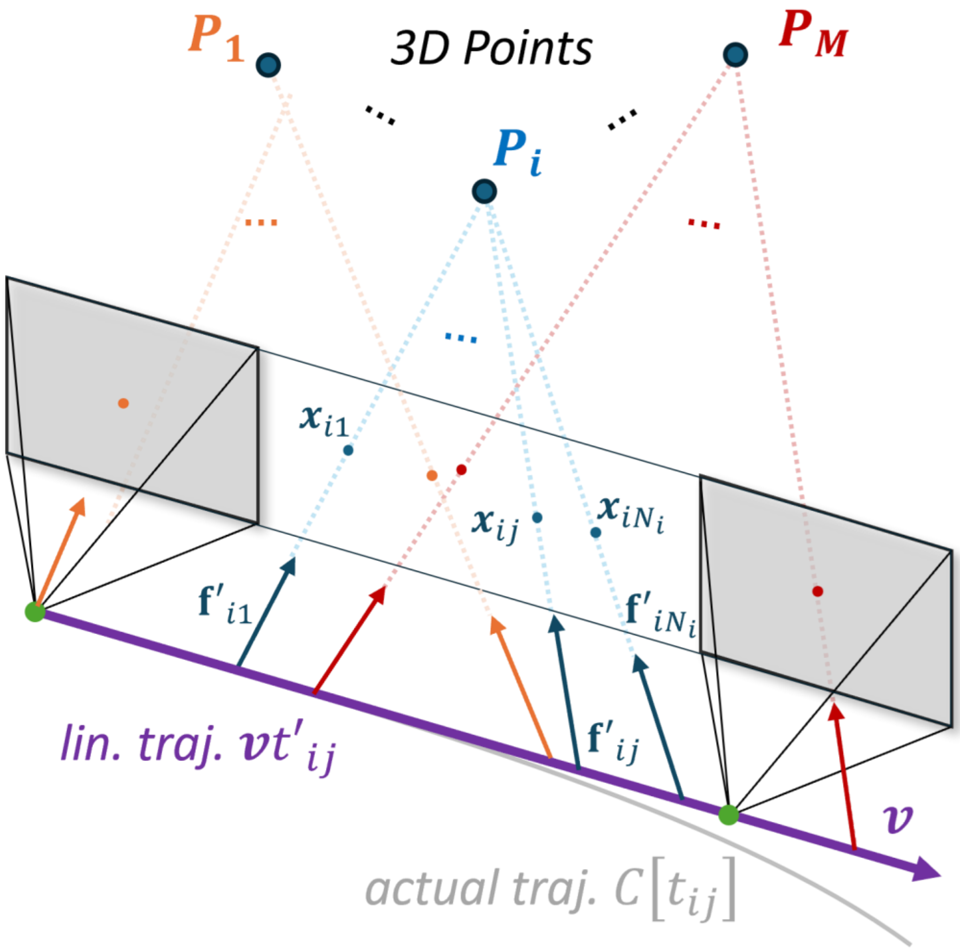
Our Goal

Develop a unified solver to estimate 3D structure and velocity directly from these asynchronous point tracks

Contributions

- A novel linear N-Point solver that uses an arbitrary number of asynchronous point feature tracks
- A general formulation that handles various sensors like global shutter, rolling shutter, and event cameras – and can even combine them
- A highly efficient method that employs Schur complement to reduce time complexity.

From Geometry to Linear System



1. Single 2D Observation

$$\mathbf{f}_{ij} = \mathbf{K}^{-1}\tilde{\mathbf{x}}_{ij}$$

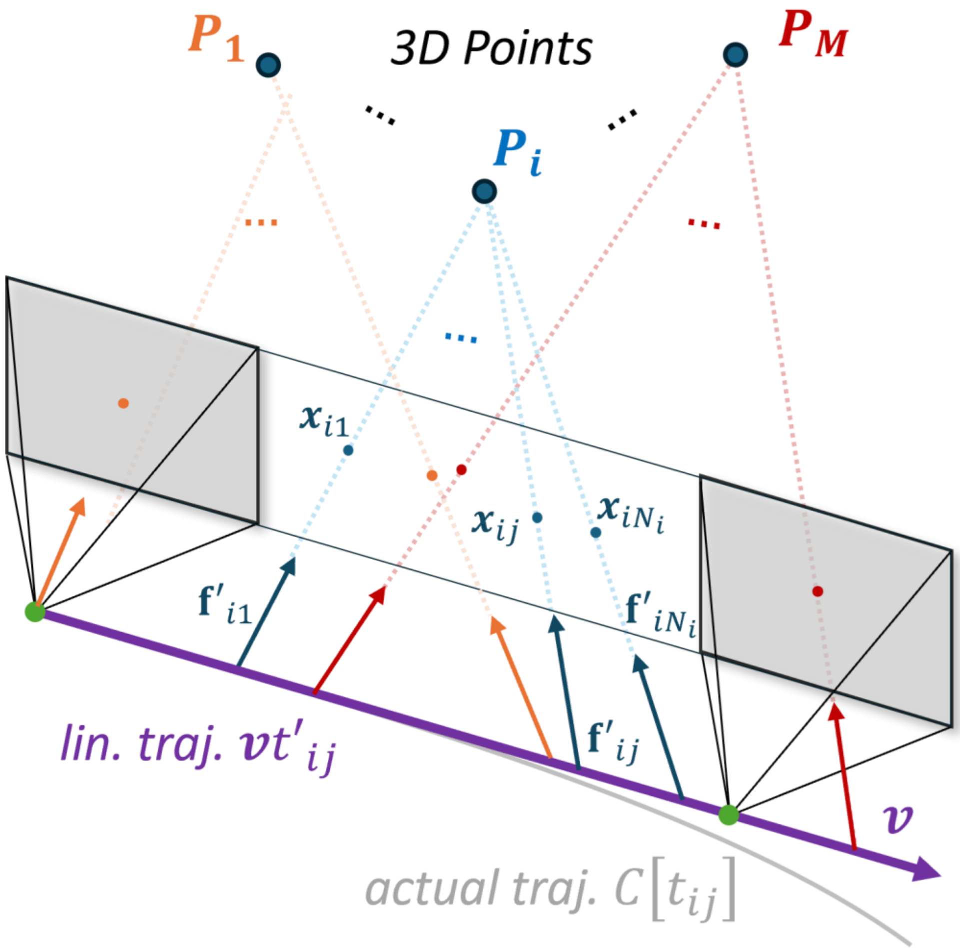
$$\mathbf{f}_{ij} \times \left(\mathbf{R}^\intercal(t_{ij})\left(\mathbf{P}_i - \mathbf{p}(t_{ij})\right)\right) = \mathbf{0}_{3\times 1}$$

Bearing vector — *Rotation* — *3D Point* — *Camera Position*

$$[\mathbf{f}'_{ij}]_\times \mathbf{P}_i - t'_{ij}\,[\mathbf{f}'_{ij}]_\times \mathbf{v} = \mathbf{0}_{3\times 1}$$

Applying constant velocity motion model

From Geometry to Linear System

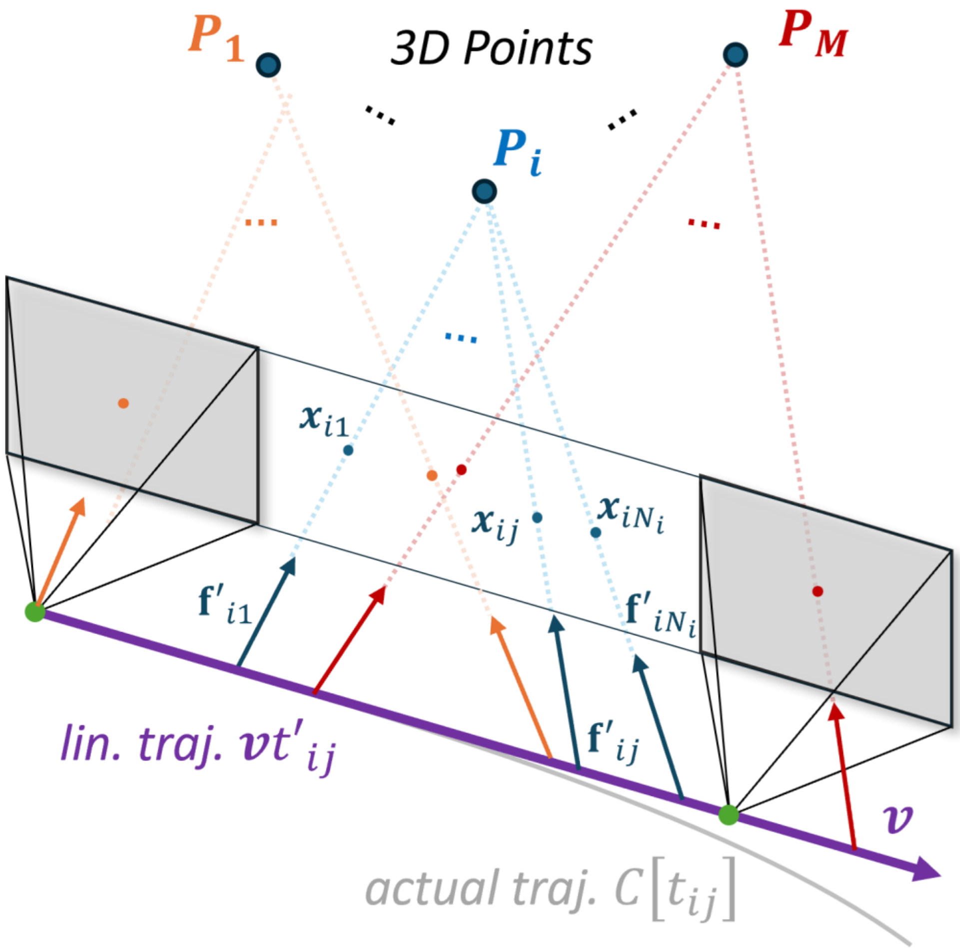


2. Multiple observations from a single 3D point

$$\underbrace{\begin{bmatrix} [\mathbf{f}'_{i1}]_\times & -t'_{i1}\,[\mathbf{f}'_{i1}]_\times \\ \vdots & \vdots \\ [\mathbf{f}'_{iN_i}]_\times & -t'_{iN_i}\,[\mathbf{f}'_{iN_i}]_\times \end{bmatrix}}_{\doteq\left[\mathbf{F}_i \quad \mathbf{G}_i\right]\in\mathbb{R}^{3N_i\times 6}} \begin{bmatrix} \mathbf{P}_i \\ \mathbf{v} \end{bmatrix} = \mathbf{0}_{3N_i\times 1}$$

Different timestamps to avoid degeneracy

From Geometry to Linear System

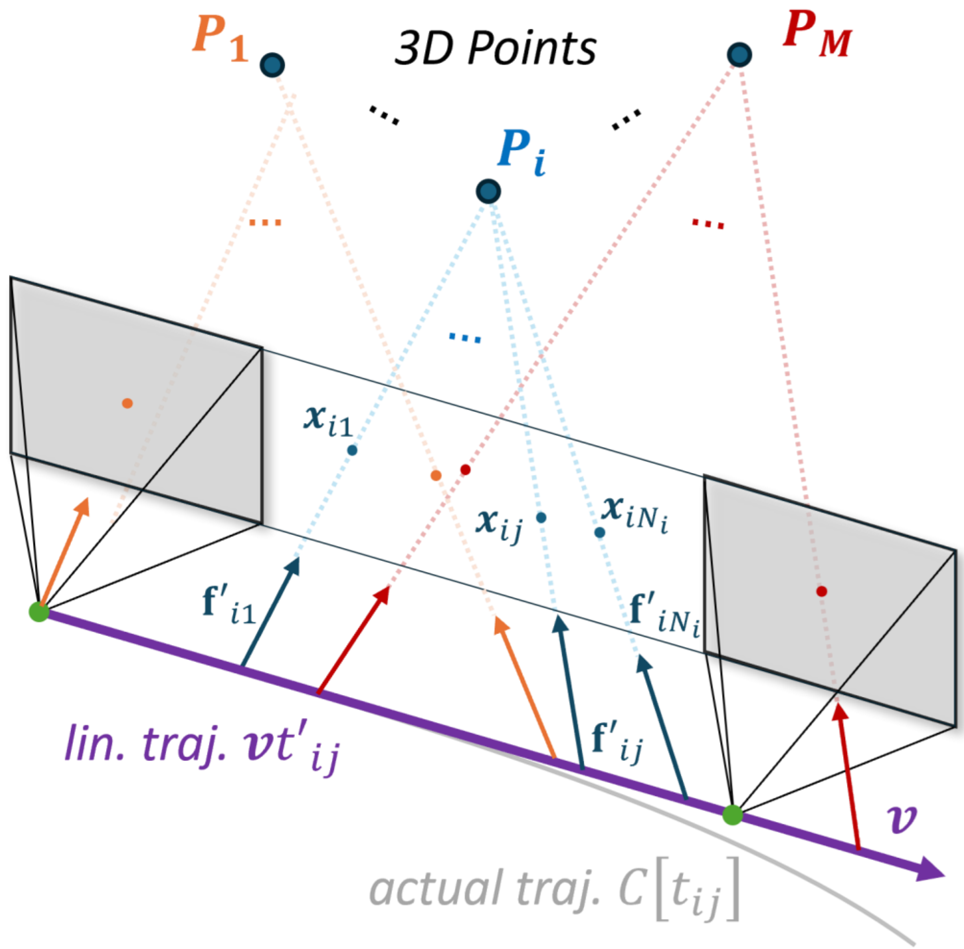


3. Multiple 3D points

$$\underbrace{\begin{bmatrix} \mathbf{F}_1 & & & & \mathbf{G}_1 \\ & \mathbf{F}_2 & & & \mathbf{G}_2 \\ & & \ddots & & \vdots \\ & & & \mathbf{F}_M & \mathbf{G}_M \end{bmatrix}}_{\doteq \mathbf{A} \in \mathbb{R}^{3N \times (3M+3)}} \underbrace{\begin{bmatrix} \mathbf{P}_1 \\ \mathbf{P}_2 \\ \vdots \\ \mathbf{P}_M \\ \mathbf{v} \end{bmatrix}}_{\doteq \mathbf{x} \in \mathbb{R}^{3M+3}} = \mathbf{0}_{3N \times 1}$$

Stacking constraints from multiple tracks

From Geometry to Linear System



4. Schur Complement

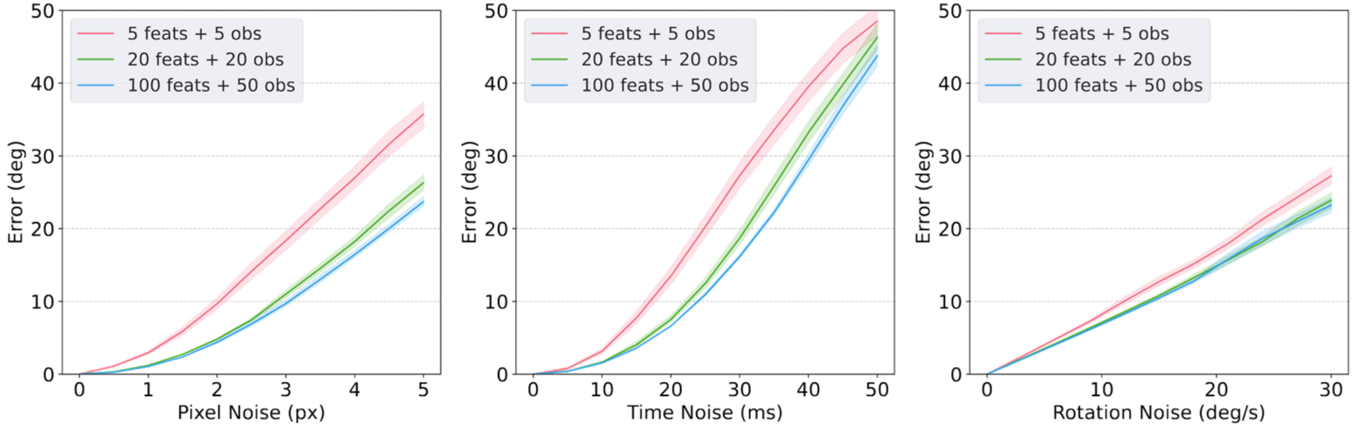
$$\underbrace{\mathbf{A}^\intercal\mathbf{A}}_{\doteq\mathbf{M}}\mathbf{x} = \begin{bmatrix}\mathbf{M}_A & \mathbf{M}_B \\ \mathbf{M}_B^\intercal & \mathbf{M}_D\end{bmatrix}\begin{bmatrix}\mathbf{P}_{1:M} \\ \mathbf{v}\end{bmatrix} = \mathbf{0}_{(3M+3)\times 1}$$

$$\underbrace{\left(\mathbf{M}_D - \mathbf{M}_B^\intercal\mathbf{M}_A^{-1}\mathbf{M}_B\right)}_{\doteq\mathbf{B}\in\mathbb{R}^{3\times 3}}\mathbf{v} = \mathbf{0}$$

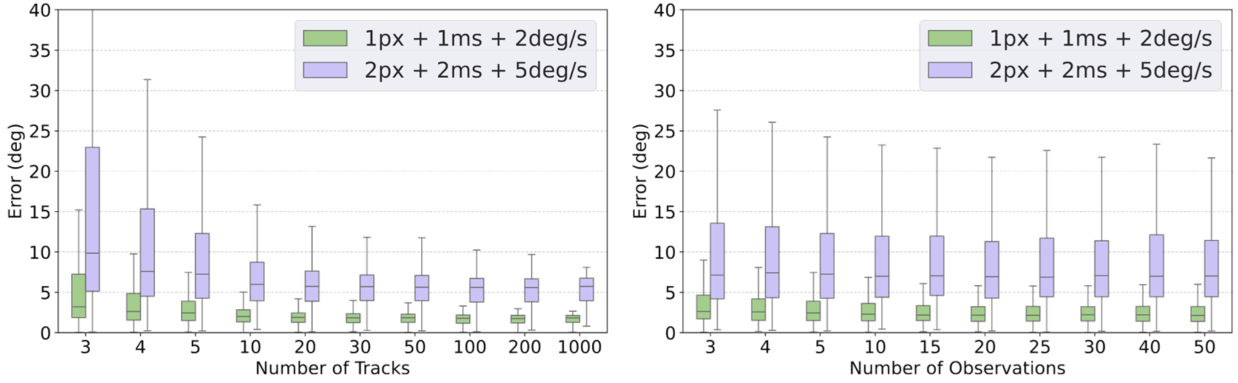
Only solve for velocity via SVD on B

Key Results: Robustness

Noise Resilience



Impacts of Spatial-temporal observations



Key Results: Real Performance

Global Shutter

VECtor dataset

Seq.	eventail [11] + E	Ours	Ours*
desk-normal	22.7 / 23.4	**15.1 / 8.5**	10.2 / 7.3
sofa-normal	21.9 / 17.6	**15.9 / 7.8**	9.8 / 6.3
mountain-normal	25.2 / 21.4	**17.1 / 7.5**	10.9 / 6.1
shapes_translation	31.8 / 32.7	**17.1 / 7.2**	9.9 / 6.2
boxes_translation	34.8 / 34.1	**16.5 / 11.6**	13.3 / 10.7

Rolling Shutter

TUM dataset

Seq.	with correction			no correction	
	eventail [11] + RS	Ours	Ours*	Ours	Ours*
Seq 4	43.8 / 40.8	**27.5 / 20.1**	22.6 / 17.4	28.1 / 22.9	22.8 / 15.7
Seq 5	45.5 / 44.8	**24.7 / 17.0**	19.3 / 13.8	27.0 / 18.4	19.2 / 14.6

Events

VECtor dataset

Seq.	eventail [11] + E	Ours + E	Ours* + E	Ours + E + GS	Ours* + E + GS
desk-normal	22.7 / 23.4	**19.3 / 17.8**	14.2 / 14.2	–	–
sofa-normal	21.9 / 17.6	**19.0 / 18.5**	16.3 / 14.9	–	–
mountain-normal	25.2 / 21.4	**17.1 / 16.1**	16.9 / 15.8	–	–
shapes_translation	31.8 / 32.7	16.8 / 10.1	13.0 / 9.1	**14.4 / 7.5**	7.0 / 6.7
boxes_translation	34.3 / 34.1	12.6 / 10.0	12.1 / 7.7	**10.3 / 8.1**	9.3 / 5.9

Error metric

$$\theta_{err} = \arccos\left(\frac{\mathbf{v}_{gt}^{\mathsf{T}}\hat{\mathbf{v}}}{\|\mathbf{v}_{true}\|\|\hat{\mathbf{v}}\|}\right)$$

- *Note: results are mean/median velocity error in degrees*
- *Ours* indicates results on subset with track inlier ratio > 0.9*

Thanks for watching

Join me at my ICCV Poster Session!

- **Date:** Tuesday, October 21st
- **Time:** 11:45 a.m. — 1:45 p.m. HST
- **Location:** Poster #407

arXiv

GitHub