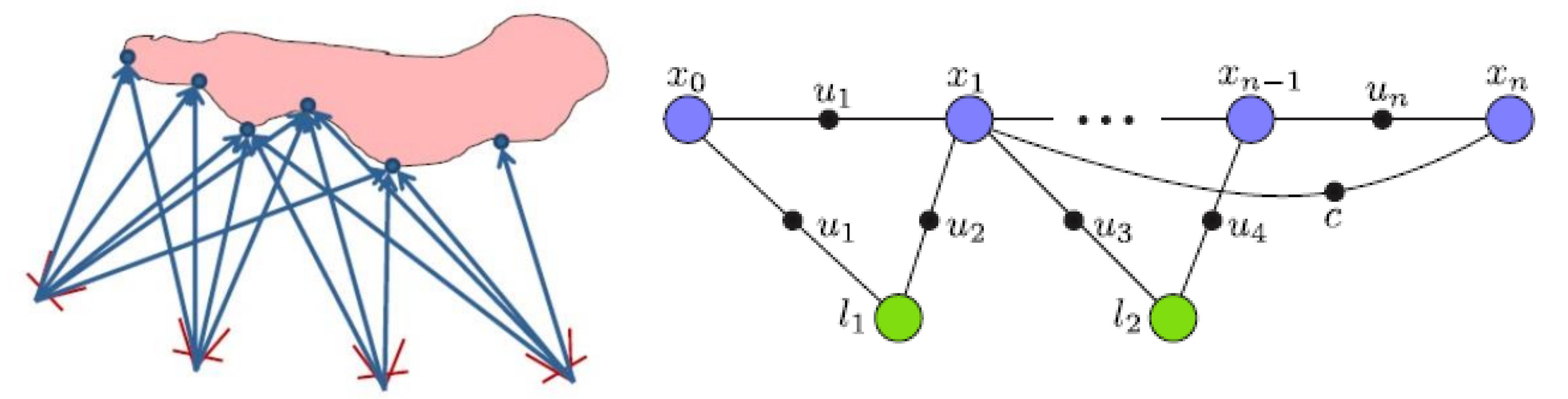
Linear RGB-D SLAM for Planar Environments

Pyojin Kim[1], Brian Coltin[2], H. Jin Kim[1]

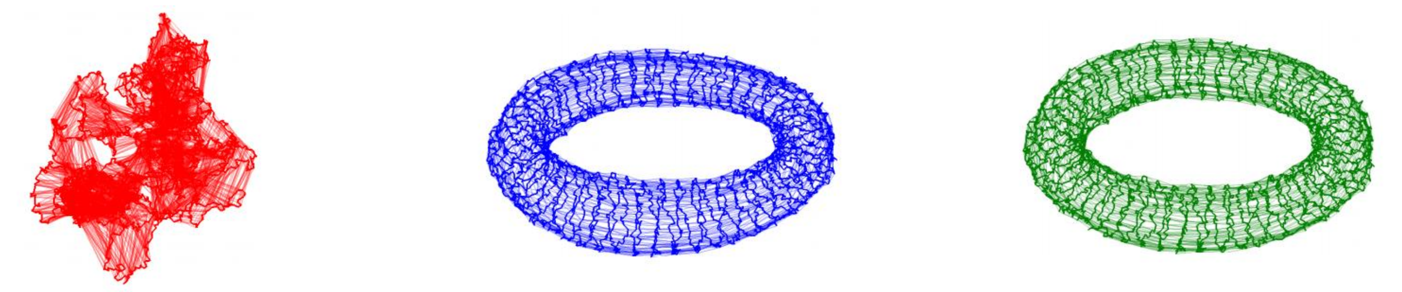
[1]Seoul National University [2]NASA Ames Research Center

Motivation



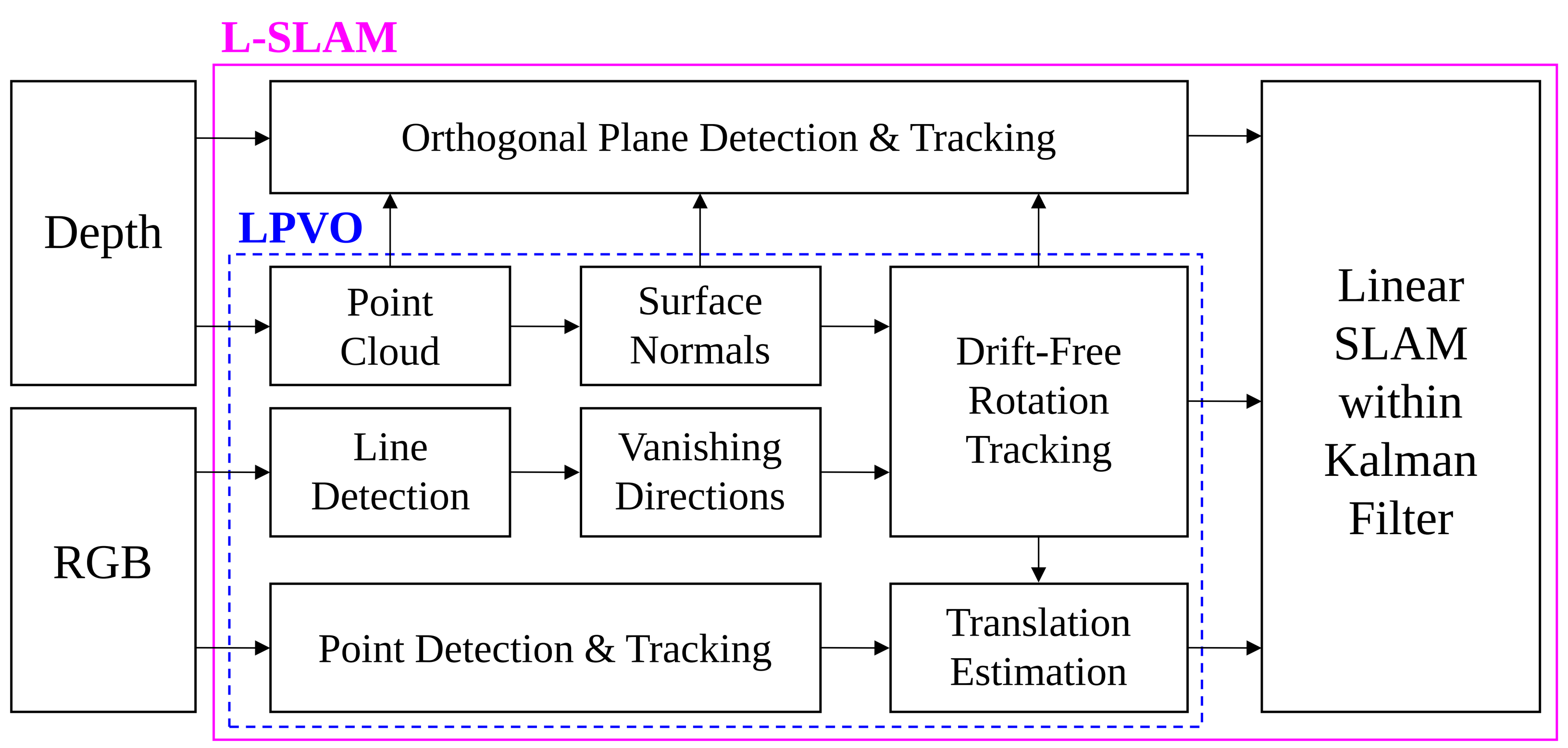
- **Visual SLAM** is a high-dimensional **non-convex & non-linear** optimization problem.

Odometry Initialization Optimum



- SLAM can be simplified as **a linear least-squares problem** *if* the camera orientations are known.

Contributions



- Orthogonal Plane Detection in Structured Environments
- A New, Linear Kalman Filter (KF) SLAM Formulation
- Evaluation and Application to Augmented Reality (AR)

Orthogonal Plane Detection

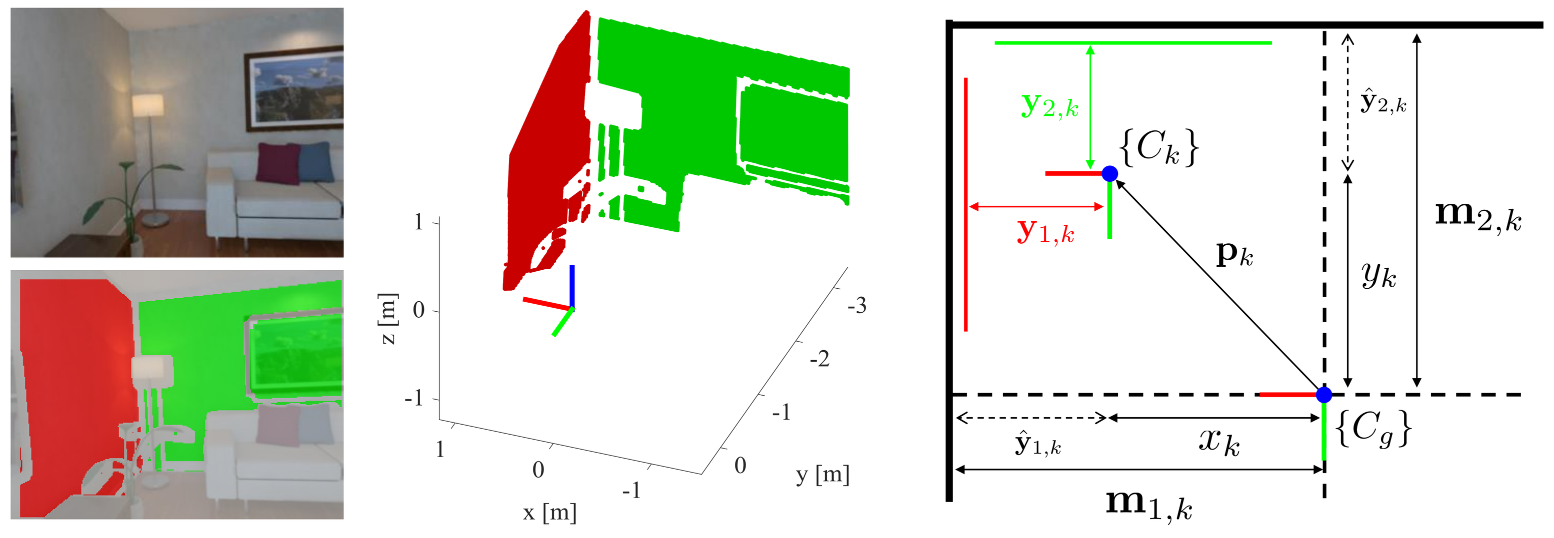
- **Modeling Planes with RANSAC**

$$n_x u + n_y v + n_z = w \qquad \left(u = \frac{X}{Z}, v = \frac{Y}{Z}, w = \frac{1}{Z}\right)$$

X, Y, Z : 3D coordinates of points

u, v : normalized image coordinates

w : measured disparity map



- **Refitting Planes to Manhattan World**

$$s^* = \arg\min_s \| s\,(r_x u + r_y v + r_z) - w \|$$

s : scale factor (reciprocal of the offset)

r_x, r_y, r_z : unit vector of the nearest Manhattan axis

Linear SLAM KF Formulation

- **State Vector**

$$\mathbf{X}_k = \left[\mathbf{p}_k^\top \quad \mathbf{m}_{1,k} \quad \cdots \quad \mathbf{m}_{n,k}\right]^\top \in \mathbb{R}^{3+n}$$

where $\mathbf{p}_k = \left[x_k \quad y_k \quad z_k\right]^\top \in \mathbb{R}^3$

$\mathbf{m}_{i,k} = \left[o_{i,k} \quad (alignment)\right] \in \mathbb{R}^1$

- **Process Model**

$$\mathbf{X}_k = \mathbf{F}\mathbf{X}_{k-1} + \boxed{\triangle \mathbf{p}_{k,k-1}} + \mathbf{w}_{k-1}$$

where $\mathbf{F} = \mathbf{I}$, $\mathbf{w}_{k-1} \sim N(0, \mathbf{Q}_{k-1})$

- 3-DoF rotational motion is compensated by LPVO.

- **Measurement Model**

$$\mathbf{y}_k = \begin{bmatrix} \mathbf{m}_{1,k} - x_k \\ \mathbf{m}_{2,k} - y_k \\ \mathbf{m}_{3,k} - z_k \\ \vdots \end{bmatrix} = \mathbf{H}_k \mathbf{X}_k + \mathbf{v}_k \quad \text{where} \quad \mathbf{H}_k = \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & -1 & 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & -1 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

$$\mathbf{v}_k \sim N(0, \mathbf{R}_k)$$

- Observation is an offset from the orthogonal plane.

- **Computational Complexity Analysis**

Table. Advantages of L-SLAM over Existing EKF-SLAM Methods

	L-SLAM (Ours)	[9]	[8]	[24]	[22]
State Size	$3+n$	$7+7n$	$7+9n$	$15+3n$	$12+10n$
Linearity	Linear	Nonlinear	Nonlinear	Nonlinear	Nonlinear

Evaluations

- **Video Clips**

- **ICL-NUIM Dataset**

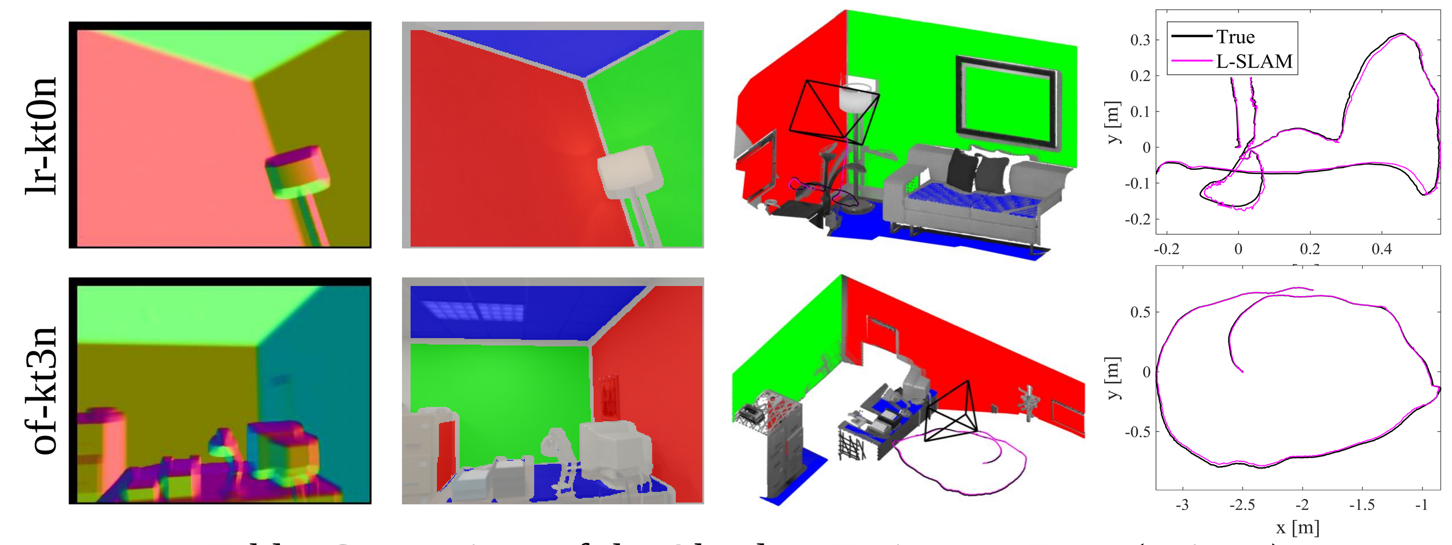
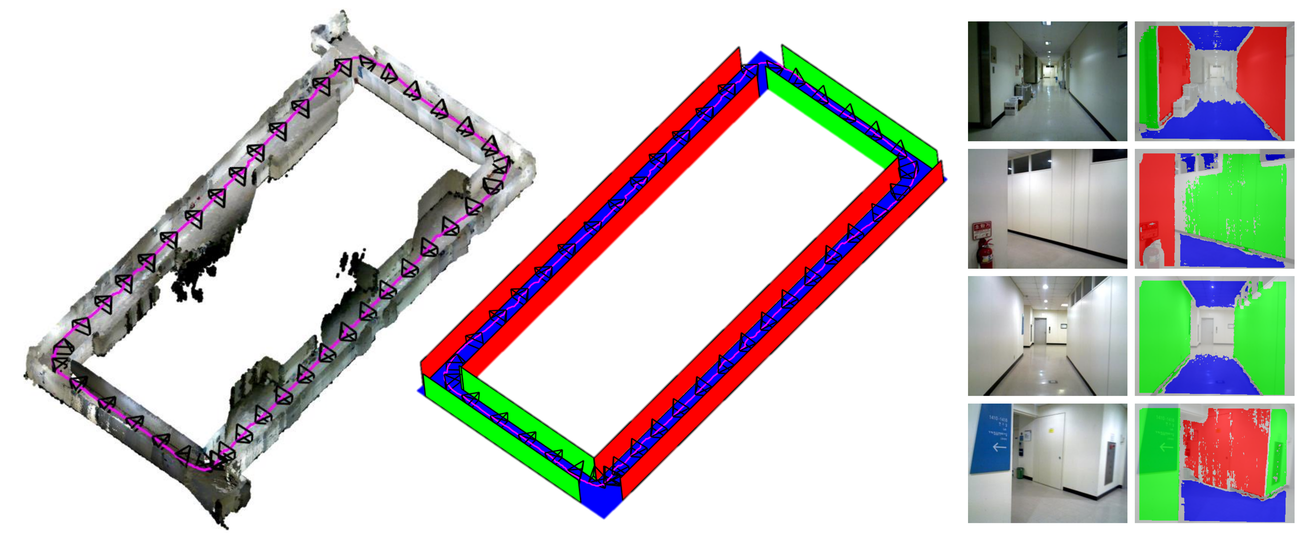


Table. Comparison of the Absolute Trajectory Error (unit: m)

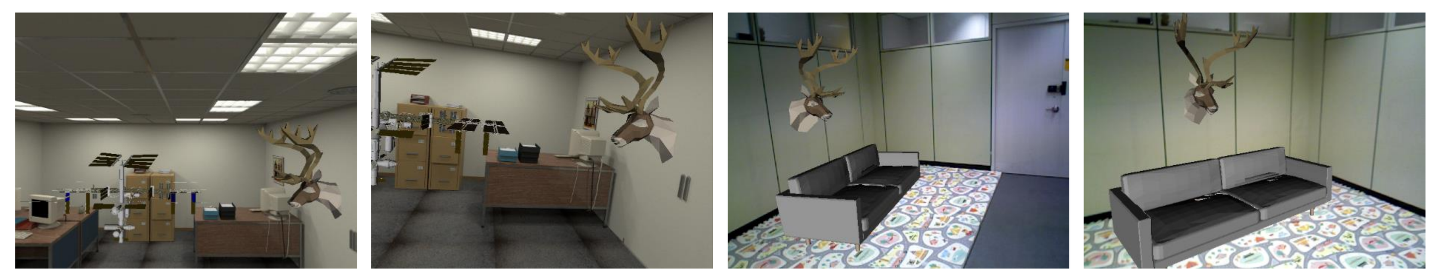
Sequence	lr-kt0n	lr-kt1n	lr-kt2n	lr-kt3n	of-kt0n	of-kt1n	of-kt2n	of-kt3n
ORB-SLAM2	0.010	0.185	**0.028**	0.014	0.049	0.079	**0.025**	0.065
DVO-SLAM	0.108	0.059	0.375	0.433	0.244	0.178	0.099	0.079
CPA-SLAM	**0.007**	**0.006**	0.089	**0.009**	–	–	–	–
KDP-SLAM	0.009	0.019	0.029	0.153	–	–	–	–
LPVO	0.015	0.039	0.034	0.102	0.061	0.052	0.039	0.030
L-SLAM (Ours)	0.012	0.027	0.053	0.143	**0.020**	**0.015**	0.026	**0.011**

- L-SLAM is **comparable** to recent SLAM approaches.

- **Author-collected RGB-D Dataset**



- **Augmented Reality (AR) Application**



- They show a **consistent** view of the 3D models.