Through-Water Stereo SLAM With Refraction Correction for AUV Localization

Robotics Institute, Carnegie Mellon University, USA

Sudharshan Suresh, Eric Westman, and ***Michael Kaess***

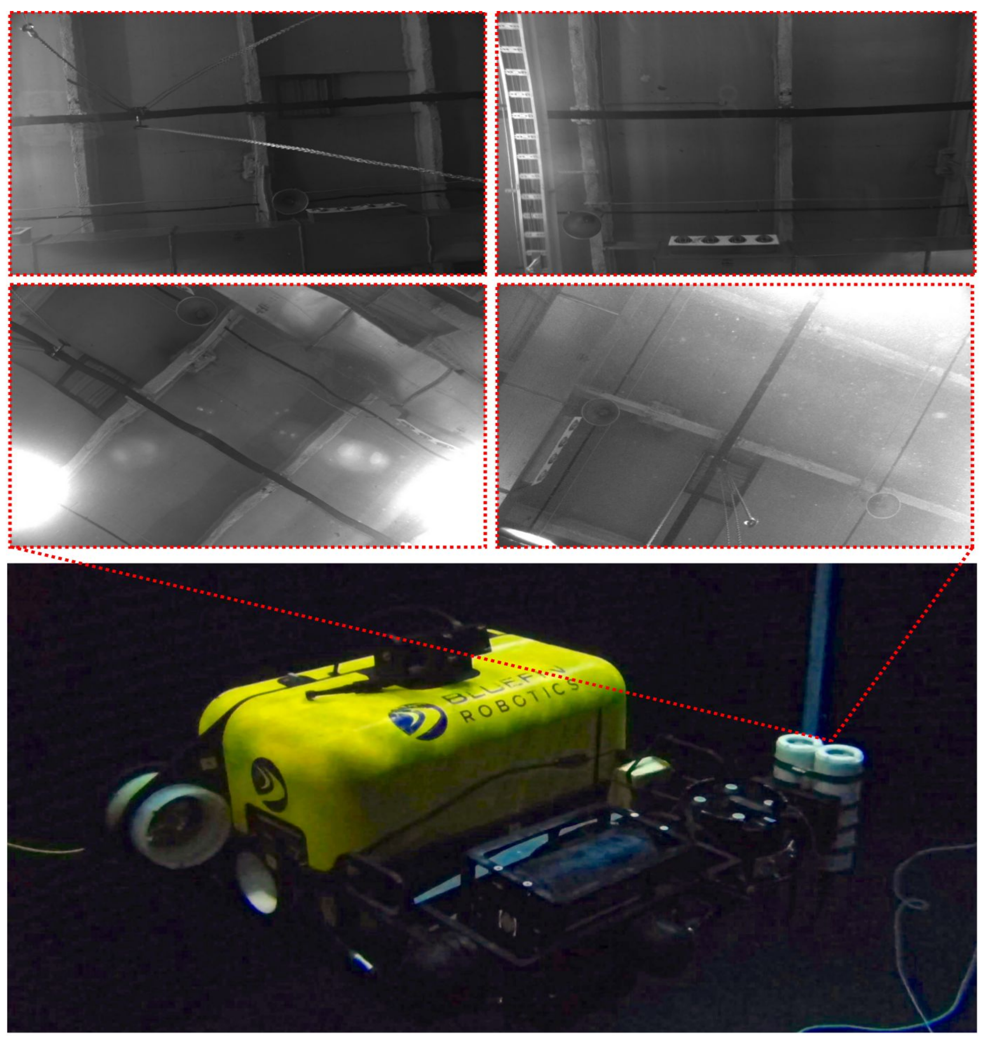

Fig. 1: **(top)** Sampling of images **(bottom)** our hovering AUV with camera and sensor payload.

INTRODUCTION

High-accuracy pose estimates are necessary for mapping and inspection tasks in indoor underwater environments, such as nuclear spent pools. **We propose a novel SLAM formulation for AUVs using an onboard upward-facing stereo camera.** This *through-water* method uses visual landmarks above the water surface, and models refraction at the water-air interface.

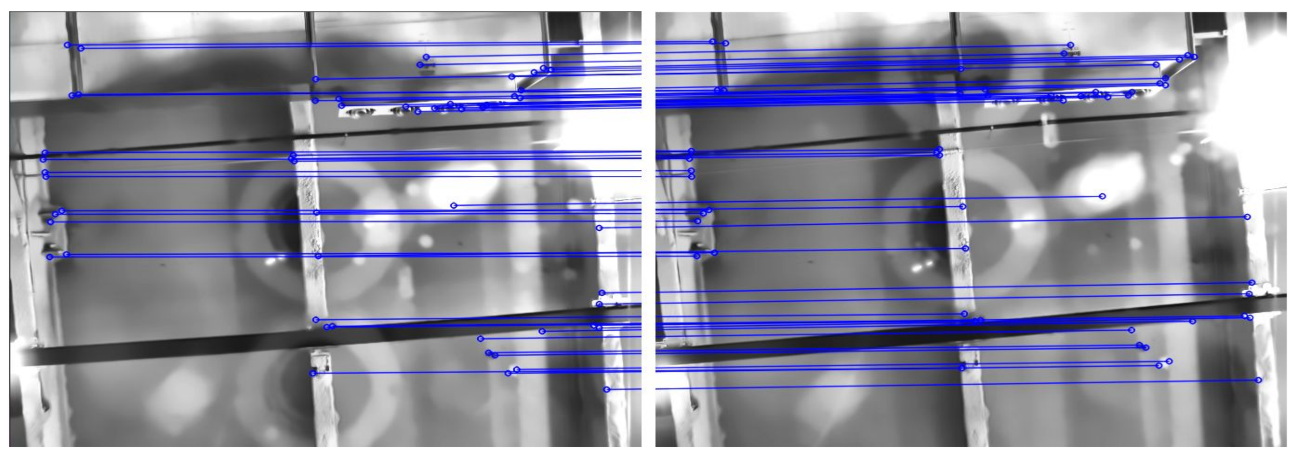

Fig. 2: Looking through the water-air interface—ORB stereo feature matching with adaptive non-maximal suppression.

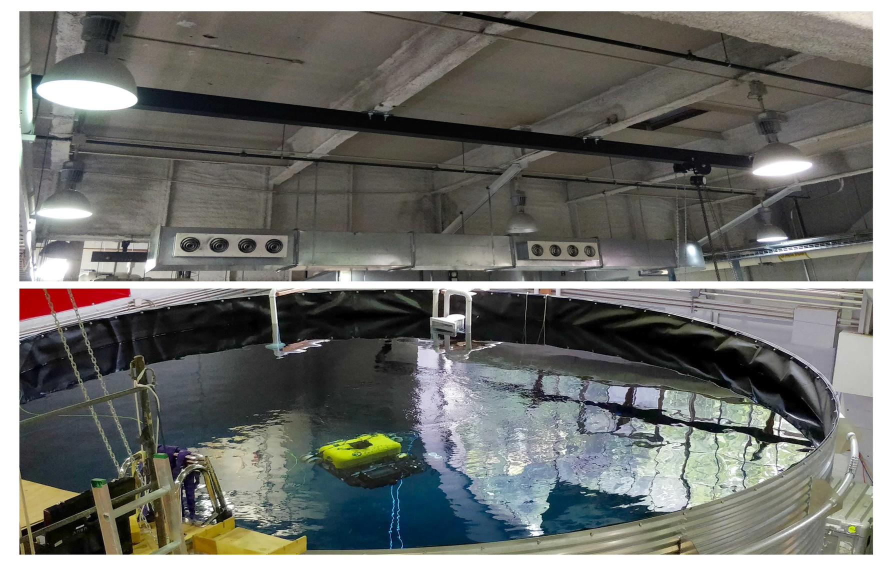

Fig. 3: Indoor underwater environment (tank and ceiling) with AUV at 1m depth.

REFRACTION OBSERVATION MODEL

How does one deal with refraction at the water-air interface? There is a systematic error in landmark triangulation when an underwater camera views points in air. We adapt prior work from aerial photogrammetry to address this:

(A) Refraction-corrected stereo triangulation (RCST)

Fig. 5: Given stereo pixel correspondences, we compute a landmark's true position. **(closed form operation)**

(B) Refraction-corrected projection (RCP)

Fig. 6: Given the true position of a landmark, we compute its projection to image pixels. **(iterative procedure)**

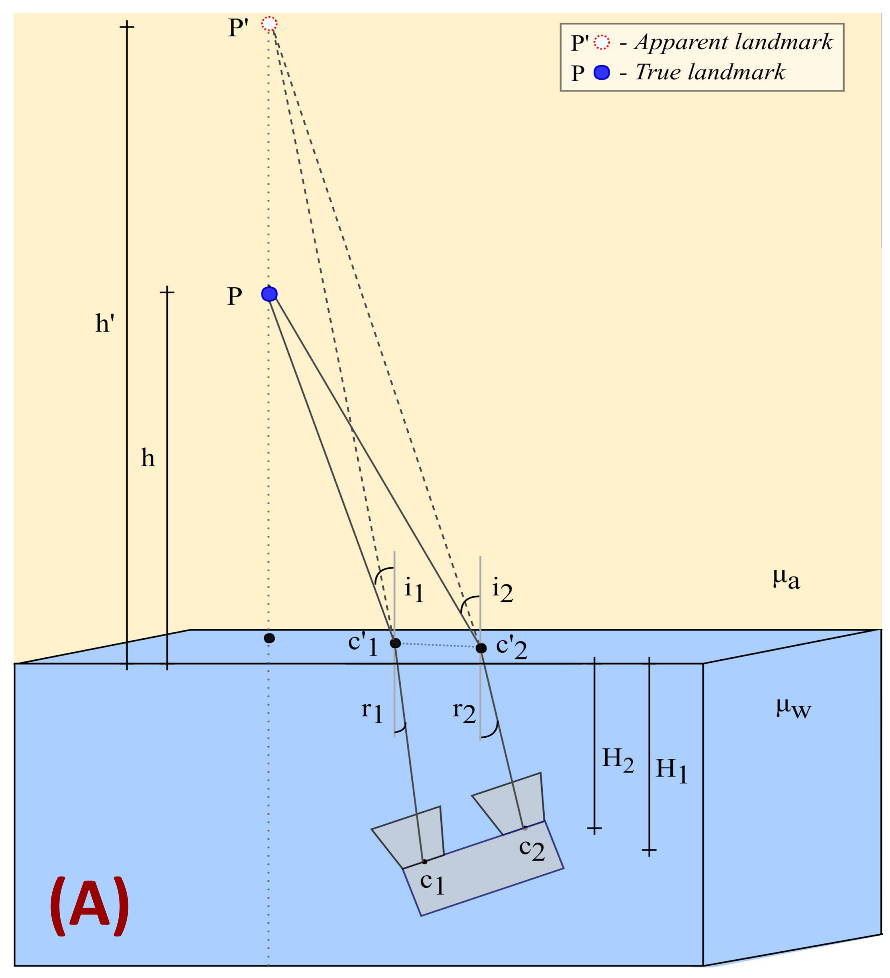

Fig. 4: **RCST** for a landmark in air. The camera incorrectly triangulates a measurement to *apparent* position P', we correct it to obtain the *true* position P.

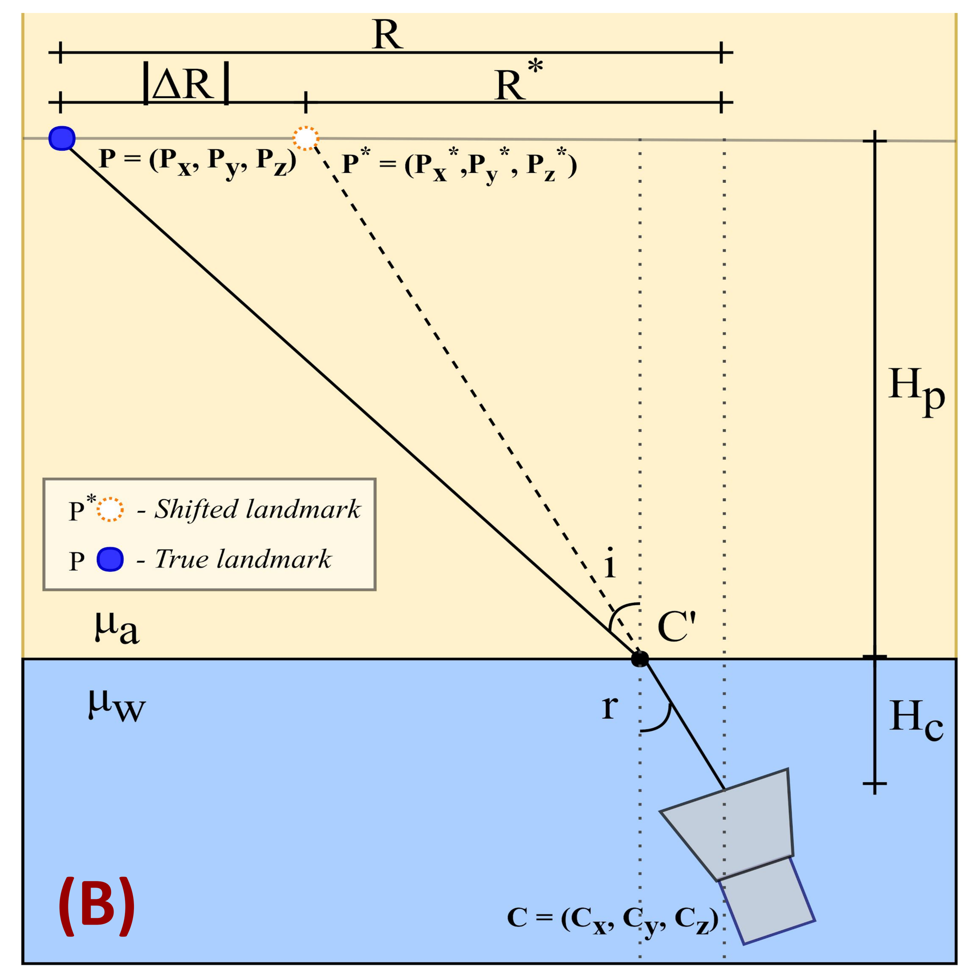

Fig. 5: Iterative **RCP** procedure shifts the landmark radially to P*, after which we trivially project it into the camera similar to the single-medium setting.

SLAM FORMULATION

The vehicle state at any timestep is:

$$x_i = [\underbrace{t_{i,x}, t_{i,y}, t_{i,z}}_{\text{translational components}},\ \underbrace{\phi_i, \theta_i, \psi_i}_{\text{yaw, pitch and roll angles}}]^T$$

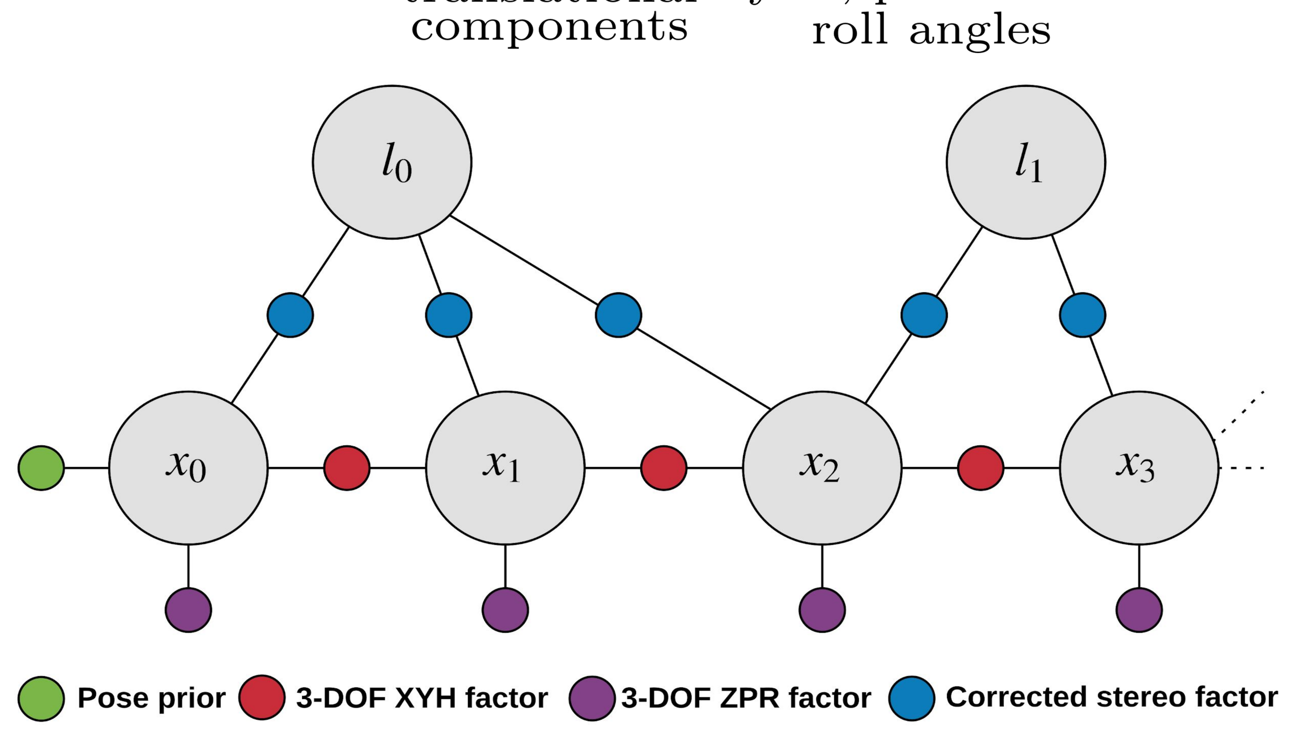

Fig. 6: Factor graph representation—large circles are **variable nodes** (poses, landmarks) and dots are **measurement factors** (odometry, corrected stereo factor, and pose prior).

The state and measurement vectors are:

$$\mathcal{X} = \{x_0, \ldots, l_0, \ldots\}$$
$$\mathcal{Z} = \{p_0, u_0, \ldots, v_0, \ldots, m_0, \ldots\}$$

The *maximum a posteriori* estimate gives us the variable values that maximally agree with measurements:

$$\mathcal{X}^* = \underset{\mathcal{X}}{\operatorname{argmax}}\ p(\mathcal{X}|\mathcal{Z}) = \underset{\mathcal{X}}{\operatorname{argmax}}\ p(\mathcal{X})p(\mathcal{Z}|\mathcal{X})$$
$$= \underset{\mathcal{X}}{\operatorname{argmax}}\ \underbrace{p(x_0)}_{\text{prior}} \prod_{i=1}^{n} \underbrace{p(u_i|x_{i-1}, x_i)}_{\text{XYH}}\ \underbrace{p(v_i|x_i)}_{\text{ZPR}} \prod_{k=1}^{m} \underbrace{p(m_k|x_i, l_j)}_{\text{corr. stereo factor}}$$

We use an online incremental solver for nonlinear least squares optimization (`GTSAM` with `iSAM2`).

RESULTS

Simulation: Square/corkscrew trajectories, known data association, noisy odometry and pixel measurements (Fig. 7).

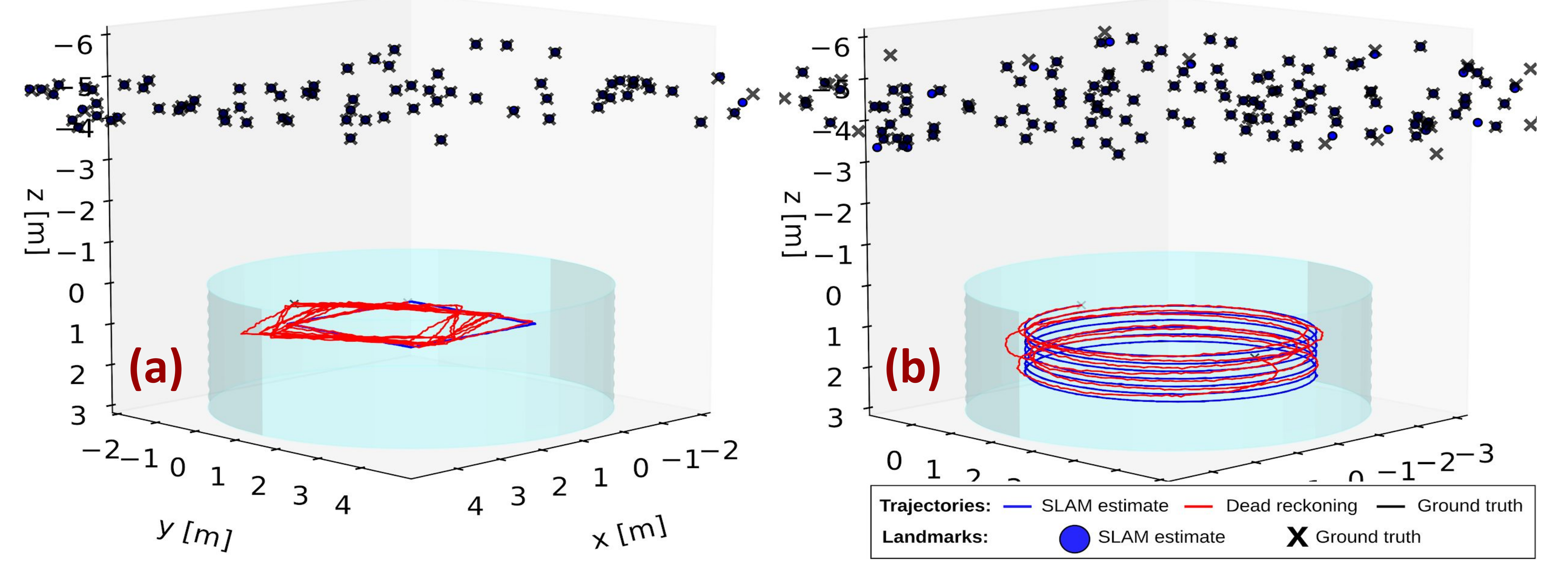

Fig 7: Trajectory and landmark comparison in (a) `square` and (b) `corkscrew` dataset.

Real-world: 12 datasets across different conditions, with synthetically added odometry noise (Fig. 8).

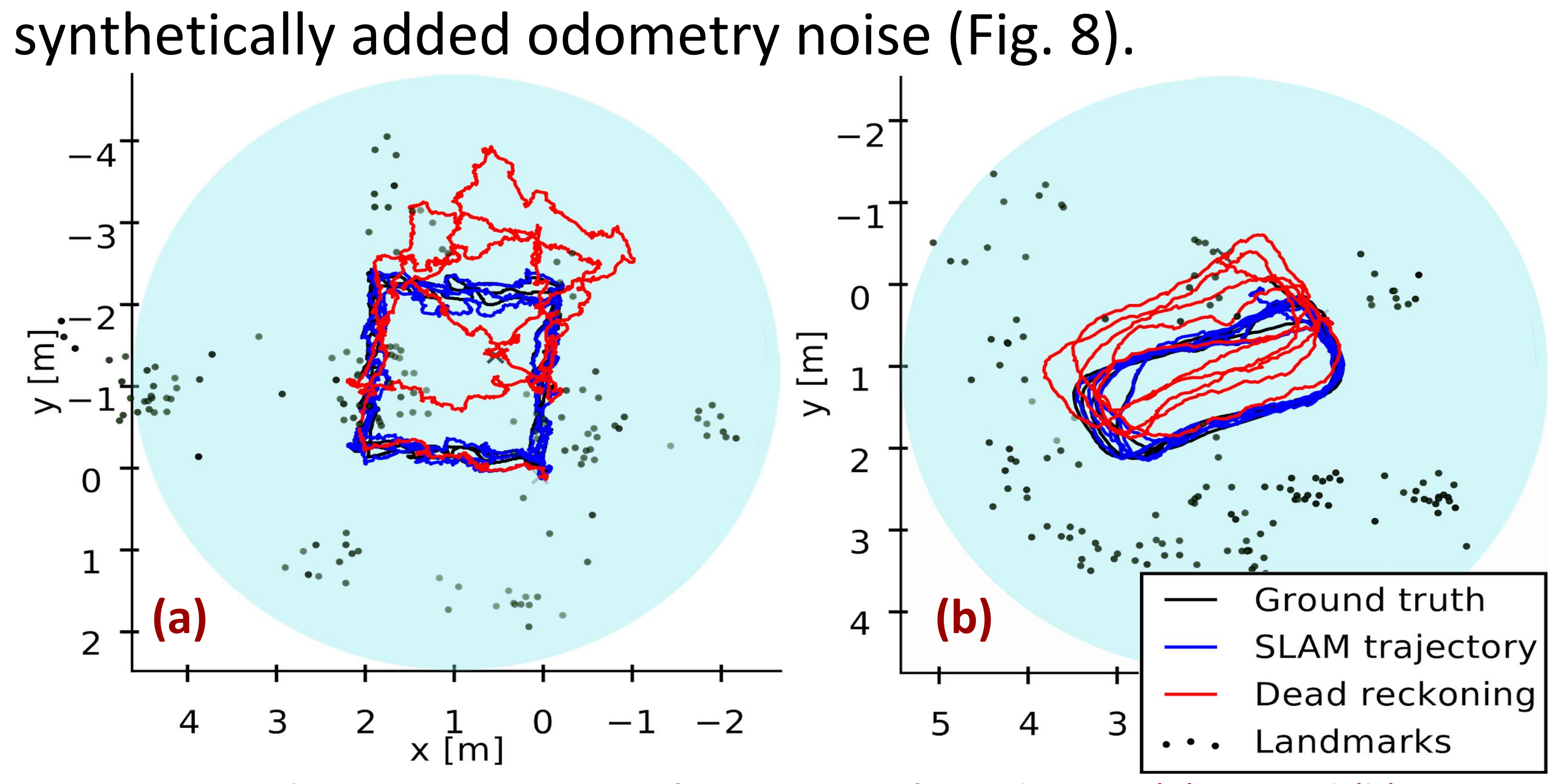

Fig 8: Qualitative comparison of trajectories from dataset (a) 08 and (b) 09.

Dataset			Dead reckoning		SLAM solution	
#	duration (s)	solve time (s)	ATE (m)	RPE_{trans} (m)	ATE (m)	RPE_{trans}
08 (1m)	724.0	884.3	0.568	0.818	0.073	0.098
09 (2m)	260.0	449.2	0.265	0.343	0.086	0.105

Table 1: Absolute trajectory error (ATE) and relative pose error (RPE) decreases as compared to dead reckoning for the 12 datasets (2 shown here).