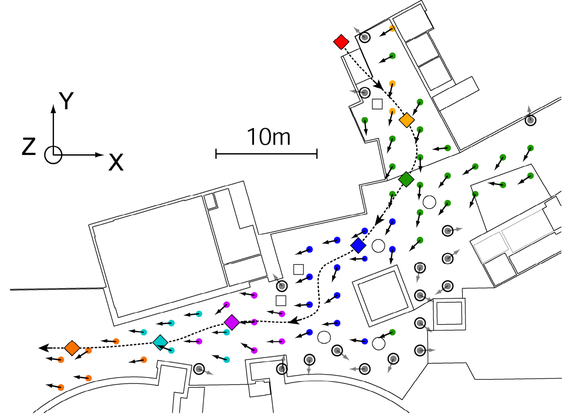
Body-relative Navigation Guidance using Uncalibrated Cameras

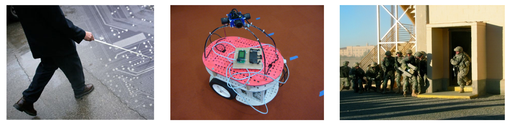
Olivier Koch and Seth Teller
MIT Computer Science and Artificial Intelligence Laboratory
{koch, teller}@csail.mit.edu



Initial trajectory (dashed) and guidance upon revisit (arrows)

Problem Statement

- Provide navigation assistance to a human in unknown, GPS-denied environments.



Applications & Scenario

- Soldiers in the field, visually impaired, robotics
- User explores the environment (*exploration phase*)
- User requests guidance (*navigation phase*)

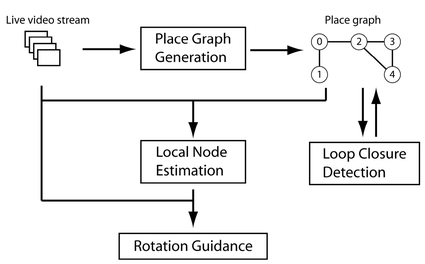
Why uncalibrated cameras?

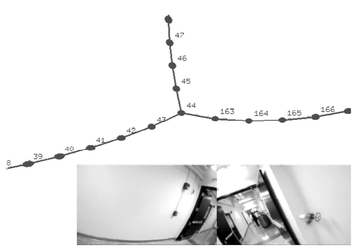
Four Pointgrey Firefly MV; FOV: 360x90 deg

- calibration is tedious
- automated calibration in the field is non-trivial
- calibration parameters change in real conditions

Method Overview

- User's exploration path = *place graph*
- Pb 1: user localization (*progress guidance)*
- Pb 2: directional guidance (*rotation guidance*)





Generating the place graph

- Analyze visual appearance variability
- Ψ-similarity between two sets of features based upon feature matching and L2 distance

Localization in the place graph

- Recursive Bayesian Estimation
- Unknown state *(x_k):* position in the graph
- Measurements *(z_k):* visual observations

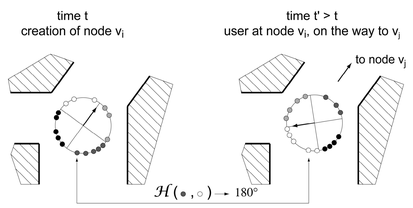
$$p(x_{k+1} \mid z_k) = \sum p(x_{k+1} \mid x_k) p(x_k \mid z_k)$$

$$p(x_{k+1} \mid z_{k+1}) = \lambda p(z_{k+1} \mid x_{k+1}) p(x_{k+1} \mid z_k)$$

$$p(z_k \mid x_k) = 1/\Psi(z_k, z_{x_k})$$

Body-relative rotation guidance

- Voting-scheme using feature matches across cameras (*match matrix*)
- Relies upon coarse geometric camera configuration
- Central Limit Theorem (*large n*): rotation error is normally distributed with $\sigma \sim 1/\sqrt{n}$
- Systems *learns* match matrix from training

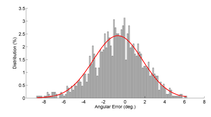


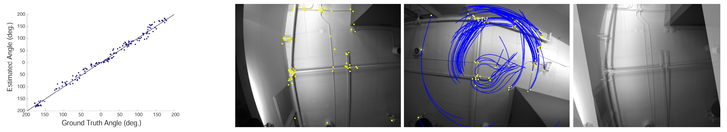
Online loop-closure detection

- Pair-wise node similarity using ``bags of words"
- Extraction of similar node sequences

Method validation

Error vs IMU: 8°
Error vs visual ground truth: 11°



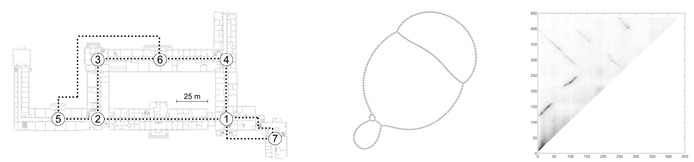


Rotation guidance ground-truth using upward facing camera

Real-world walking excursions

Name	Scenario	Duration	Length	# frames	# nodes	# checkpoints
MEZZANINE	replay	10 min.	400m	6,000	91	36
GALLERIA	homing	15 min.	700m	9,000	154	150
CORRIDORS	point-to-point	30 min.	1,500m	18,000	197	0

SIFT features (left); body-relative guidance output (right)



Loop closure (CORRIDORS dataset). 30 min walk across MIT corridors.

Discussion

- ✓ Robust guidance in user's body frame
- ✓ Single assumption on the environment (distinctive visual features)
- Fails in dark or featureless environments
- Assumes smooth user motion

Acknowledgments

This project supported by Draper Laboratory.