WP2: Visuo-conceptual modeling for situation awareness

CTU, ETHZ, Fraunhofer, ...

Tomas Svoboda, CTU

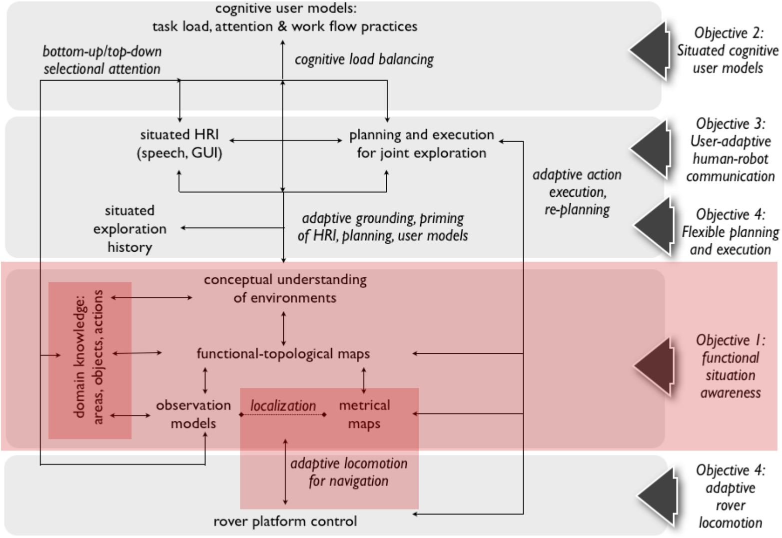
April 23, 2013

WP2 objective and project context

- Overall WP objective
  - provides observations about static and dynamic aspects of the robot's surroundings
- Year 3 objective
  - vision (perception) for shared situation awareness
- Contribution to project
  - contributes to the objective 1: Function models of dynamic environments

Natural Human-Robot Cooperation in Dynamic Environments

contribution to the integrated system



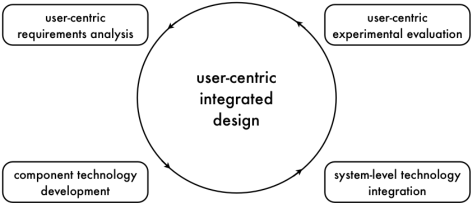
user centric approach

Challenges in the railway use case:

- uneven terrain
- large changes in visual conditions

Integrated system evaluation:

- human in the loop
- repeatability (understandability)



Data gathering in field experiments

- images (visual, thermo)
- 3D laser scans
- IMU, . . .

Integrated system functionality:

- Objects on map (WP1) and on user screen (WP3)
- Visual odometry, IMU filtering in joint position estimate (WP1)
- Human-in-the-loop (WP3)

overview

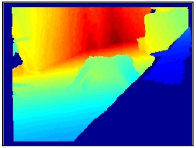
- introduction
- addressing previous comments
- progress in year 3
  - generic virtual PTZ (operator comfort)
  - higher order MRF (theoretical contribution)
  - deformable visual object detector (and fused with 3D)
  - thermal imaging - (fusing modalities towards a victim detector)
  - visual based localization
  - terrain perception/classification (for robot morphology)
- plans for year 4
- conclusions

NIFTi
Natural Human-Robot Cooperation in Dynamic Environments

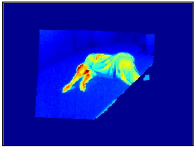
ARC1: challenges in detection of human bodies

Certain challenges remain for WP2 including the detection of human bodies. . . . bodies are not always visible with faces, . . .

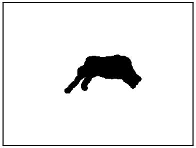
color



depth



thermo

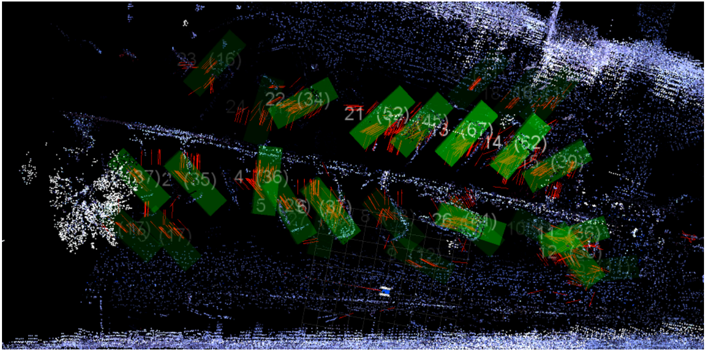


mask

NIFTi Natural Human-Robot Cooperation in Dynamic Environments

ARC2: car detection and localization

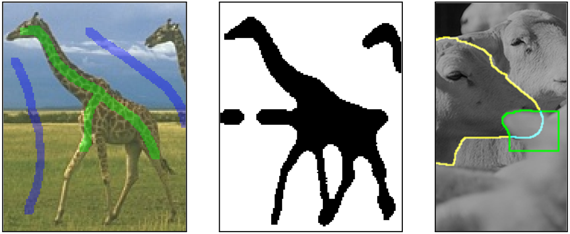
. . . performance need to be evaluated . . .



General functionalities, virtual views

Natural Human-Robot Cooperation in Dynamic Environments

Higher order MRF models

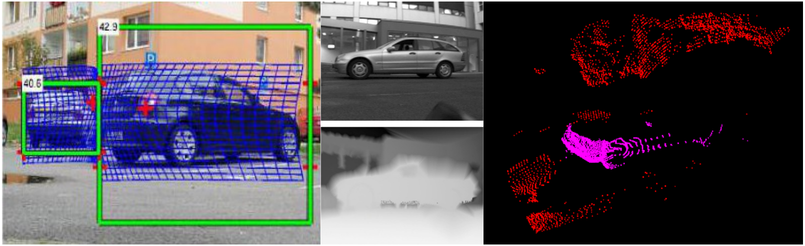


(a) (b) (c)

Figure : (a)-(b) Object segmentation with curvature prior. Given initial seeds in (a) the camouflage-colored giraffe is segmented in (b). (c) Shape inpainting. Given a missing area (green), the shape is smoothly completed (blue). [1]

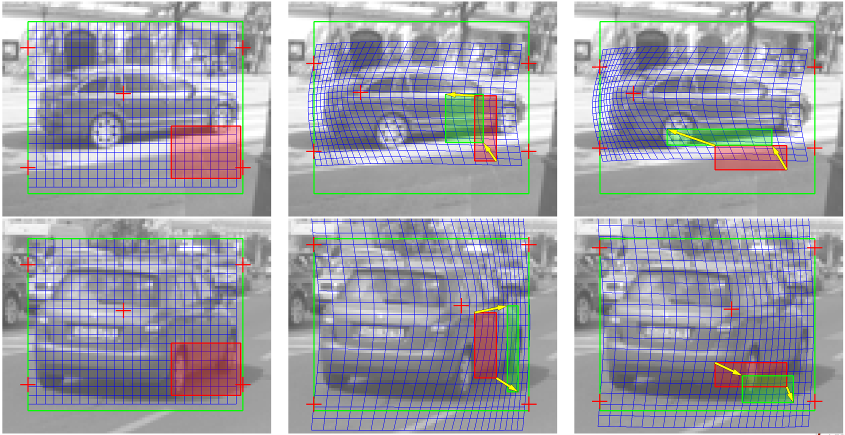
[1] [Shekhovtsov-Hlavac-DAGM2012, Shekhovtsov-Hlavac-IJCV2012]

Object detector, fusing modalities



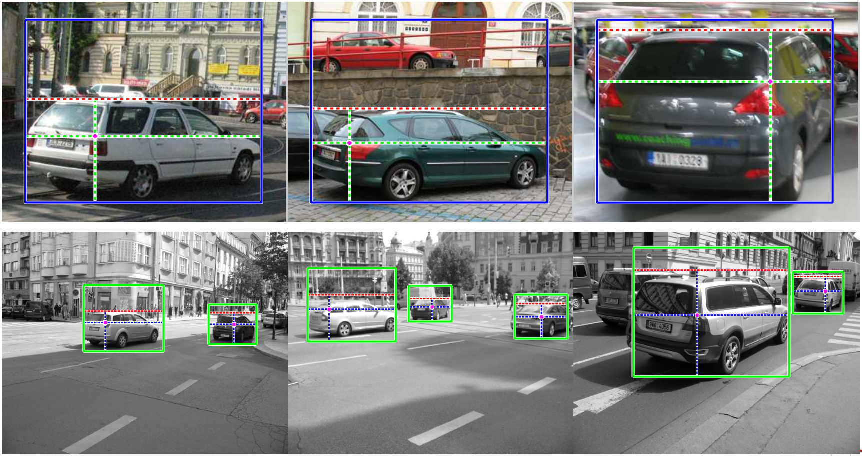
- Lidar sits low
- depth data often uncomplete for certain object parts

LISA: Local interleaved sequential alignment[2]

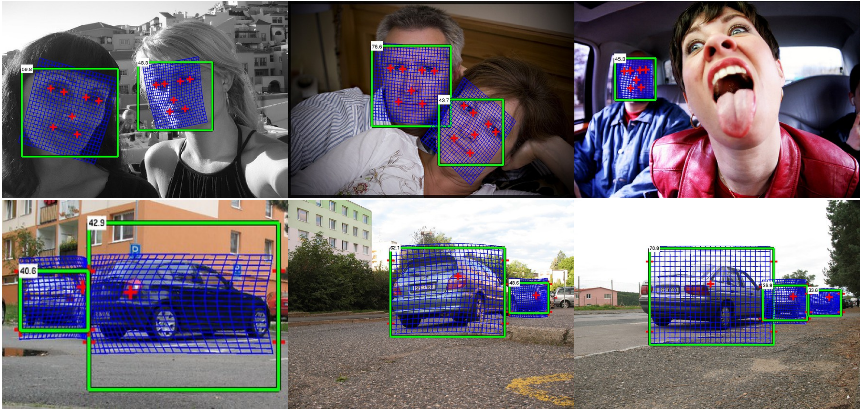


[2][Zimmermann-Hurych-Svoboda-ACCV2012, IEEE-TPAMI-2013(?)]

LISA training and results

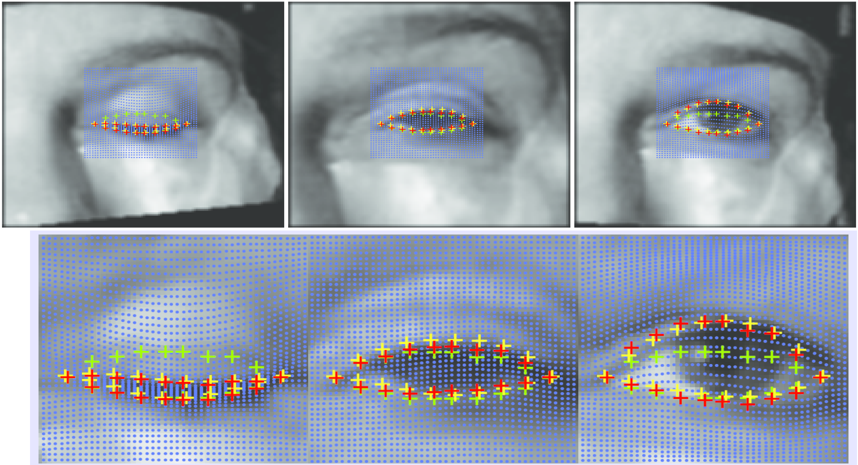


LISA results

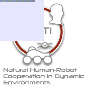


Natural Human-Robot Cooperation in Dynamic Environments

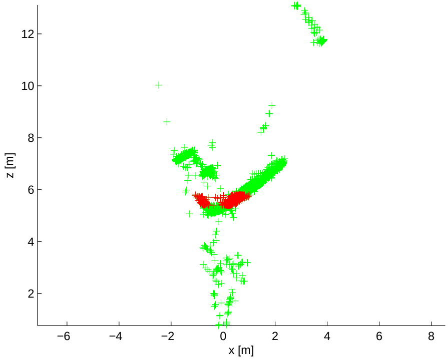
LISA non-rigid objects - eye opening

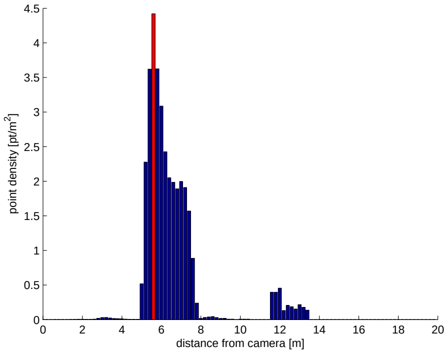


alignment in few steps

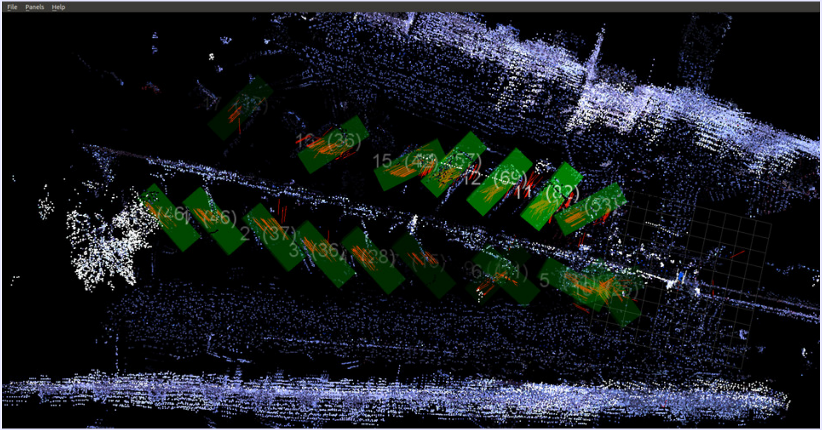


Connecting LISA and 3D



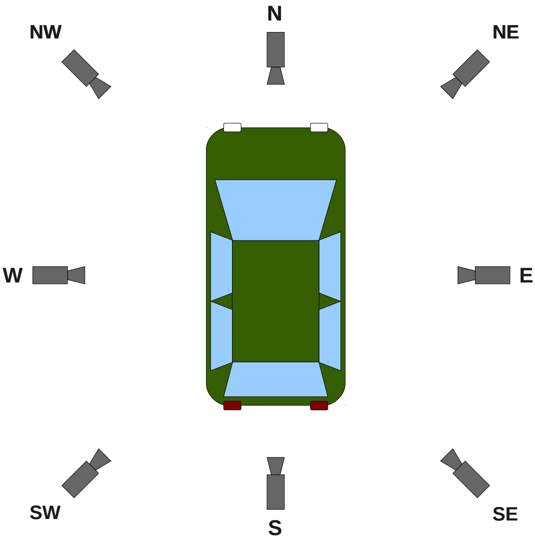


Connecting LISA and 3D



video (played faster than realtime)

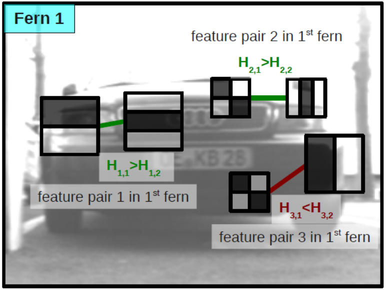
Fern detector 3D and intensities

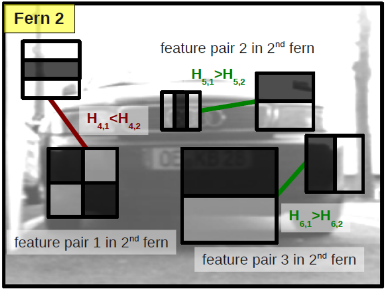


Fern detector 3D and intensities

W S NE

Learning FERNs





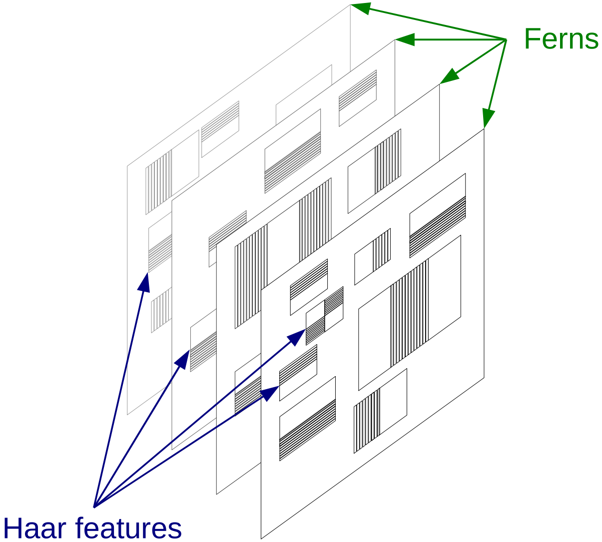
$F_1 \sim \{1,1,0\}$ $\quad F_1 = 2^0\cdot 1 + 2^1\cdot 1 + 2^2\cdot 0 = \boxed{3}$

$F_2 \sim \{0,1,1\}$ $\quad F_2 = 2^0\cdot 0 + 2^1\cdot 1 + 2^2\cdot 1 = \boxed{6}$

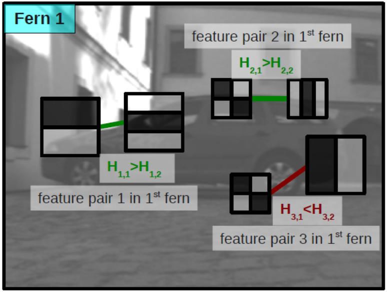
$F_1 \backslash c_i$	N	E	S	W
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	**1**	0	0	0
4	0	0	0	0
5	0	0	0	0
6	0	0	0	0
7	0	0	0	0

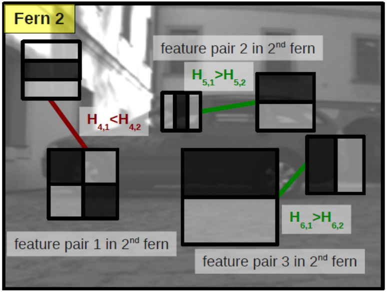
$F_2 \backslash c_i$	N	E	S	W
0	0	0	0	0
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	0	0	0	0
6	**1**	0	0	0
7	0	0	0	0

FERNs



Classifying FERNs





$F_1 \sim \{1,1,0\}$ $\quad F_1 = 2^0\cdot 1 + 2^1\cdot 1 + 2^2\cdot 0 = \boxed{3}$

$F_2 \sim \{0,1,1\}$ $\quad F_2 = 2^0\cdot 0 + 2^1\cdot 1 + 2^2\cdot 1 = \boxed{6}$

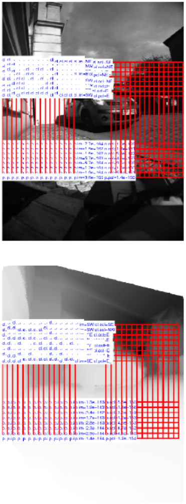
$F_1 \backslash c_i$	N	E	S	W
0	0.12	0.08	0.09	0.11
1	0.14	0.23	0.08	0.02
2	0.21	0.11	0.04	0.09
3	0.04	0.05	0.19	0.33
4	0.01	0.24	0.51	0.07
5	0.35	0.10	0.02	0.21
6	0.04	0.12	0.04	0.01
7	0.09	0.07	0.03	0.16

c_i	N	E	S	W
$p_i = F_1 \cdot F_2$	0.04× 0.03= **0.001**	0.05× 0.09= **0.005**	0.19× 0.03= **0.006**	0.33× 0.07= **0.023**

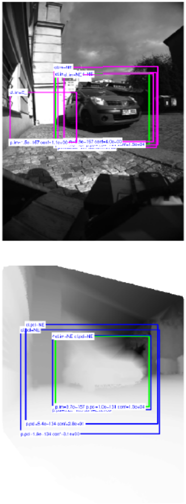
$F_2 \backslash c_i$	N	E	S	W
0	0.02	0.21	0.06	0.21
1	0.13	0.19	0.12	0.09
2	0.07	0.01	0.03	0.02
3	0.29	0.15	0.17	0.23
4	0.06	0.20	0.23	0.09
5	0.31	0.14	0.17	0.19
6	0.03	0.09	0.03	0.07
7	0.09	0.01	0.19	0.10

Finding objects

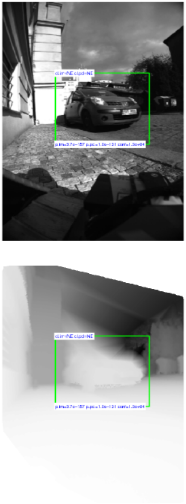
1. Classify all windows



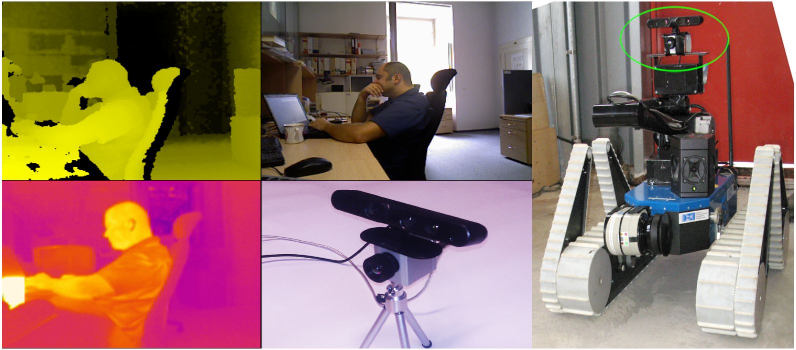
2. Thresholding



3. Non-maxima-suppression



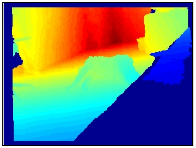
Combining modalities

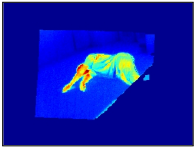


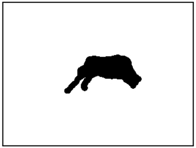
Thermal imaging

Registering modalities


color

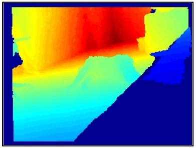

depth


thermo

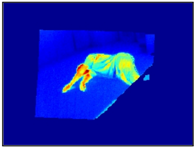

mask

Registering modalities

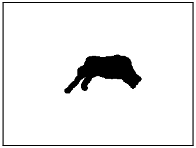
color



depth



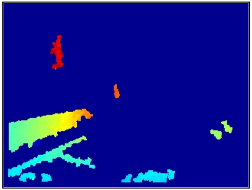
thermo



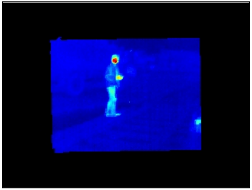
mask

Outdoor:

color



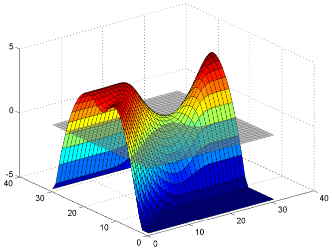
depth??



thermo?

Pixel-wise classifier

body temperature p_F is a mixture of two fixed μ_1 and σ_F (*skin*)
and a Gaussian with dynamically estimated mean μ_2, and a fixed variance σ_F (*clothes*)

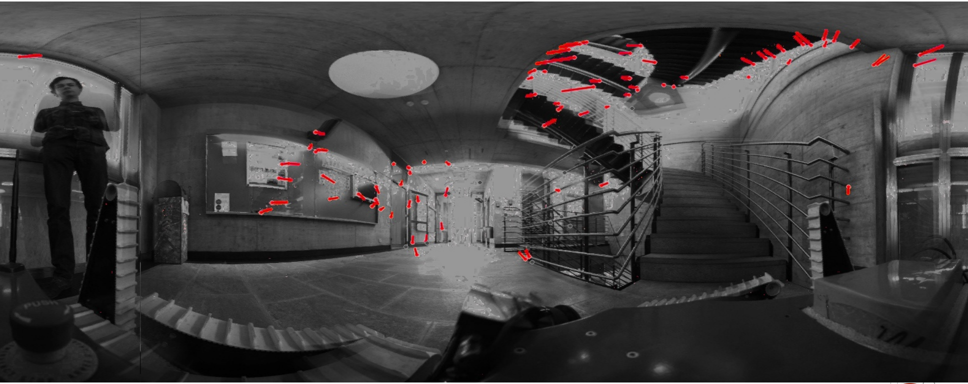


Classifier: $\log \frac{p_F(x;\mu_B)}{p_B(x;\mu_B)}$
Pre-computed table

$$\mu_2 = \begin{cases} \mu_1 & \text{if } \mu_B > \mu_1 \\ \alpha * \mu_B + (1-\alpha)\mu_1 & \text{if } \mu_B < \mu_1, \end{cases}$$

Detecting bodies

Visual odometry – correspondences (disparities)



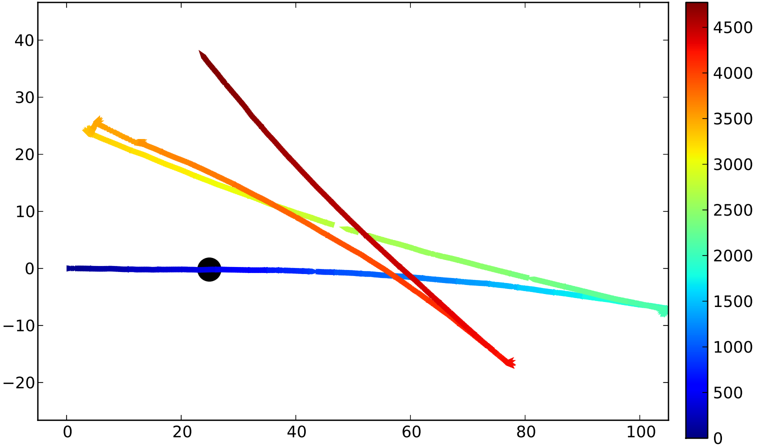
Visual odometry, along rails

video (played much faster than realtime)

- complement the existing methods
- **Plan:** ICP mapping, INSO, VO, together → *large scale* robot localization

Natural Human-Robot Cooperation in Dynamic Environments

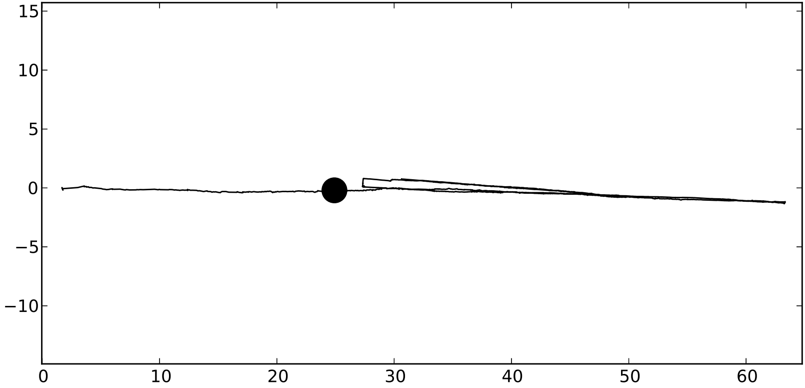
INSO trajectory



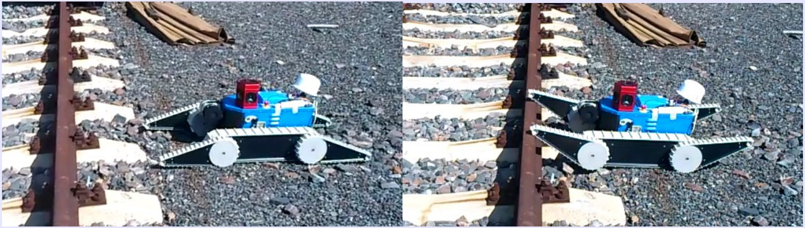
- local method supporting ICP mapping[3]

[3] [Kubelka-Reinstein-ICRA2012]

VO trajectory



Terrain perception/classification



video (stepping up sleeper)

video (crossing rails)

Natural Human-Robot Cooperation in Dynamic Environments

plans for year 4

- vision for localization (visual odometry, compass)
- terrain perception: machine learning approach[4]
- victim detector

[4] [Reinstein-Kubelka-Zimmermann-ICRA2013]

thank you for your attention