Indoor Place Recognition Using RGB-D Camera Based on Planar Surfaces and Line Segments

Robert Cupec, Emmanuel Karlo Nyarko, Damir Filko

J. J. Strossmayer University of Osijek
Faculty of Electrical Engineering, Computer Science and Information Technologies Osijek, Croatia

Andrej Kitanov, Ivan Petrović

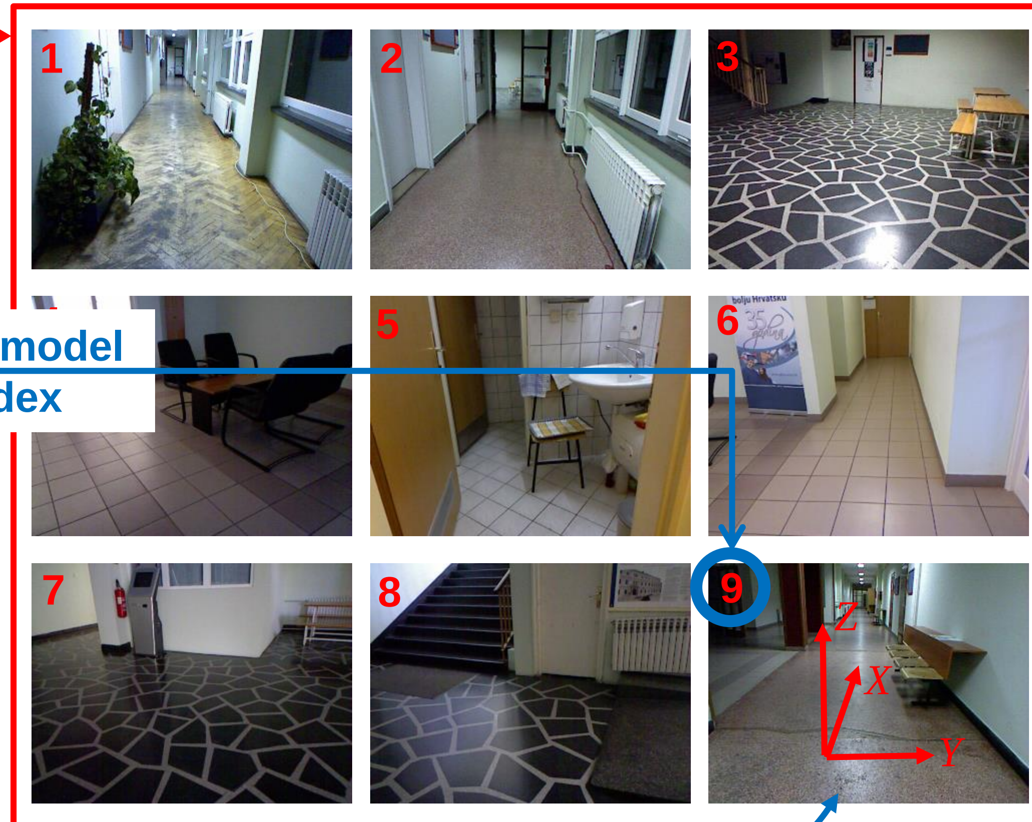
University of Zagreb
Faculty of Electrical Engineering and Computing, Zagreb, Croatia

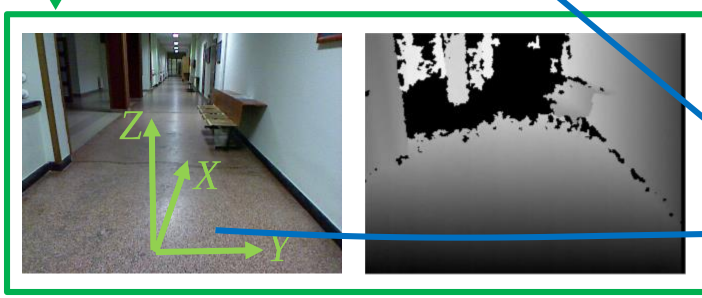
Global Localization

- Given an environment map and a camera image acquired somewhere in the considered environment, identify the camera pose at which the image is acquired.

Problem: Global Localization

- Given an environment map and a camera image acquired somewhere in the considered environment, identify the camera pose at which the image is acquired.





Related Research

- Highly efficient feature-based algorithms for localization which use 3D camera:
  - Huang et al., 2011 – visual odometry
  - Endres et al. 2012; Stückler and Behnke 2013 – SLAM
  - Fallon et al., 2012 – local pose tracking
- Appearance-based place recognition using 'standard' camera image :
  - Cummins and Newman, 2009;
  - Ciarfuglia et al., 2012;
  - Liu and Zhang 2012;
  - Milford, 2013;
  - Galvez-Lopez and Tardos, 2012
- Place recognition using 3D sensors
  - Fernandez-Moral, Mayol-Cuevas, Arevalo & Gonzalez-Jimenez, Fast place recognition with plane-based maps, in Proc. IEEE Int. Conf. on Robotics and Automation, 2013
  - Badino et al., 2012 – 'standard‘ camera image + 2 lidars
  - Granström et al., 2011 - loop closing based on features extracted from 3D point clouds
  - Cobzas and Zhang, 2001

Pipeline

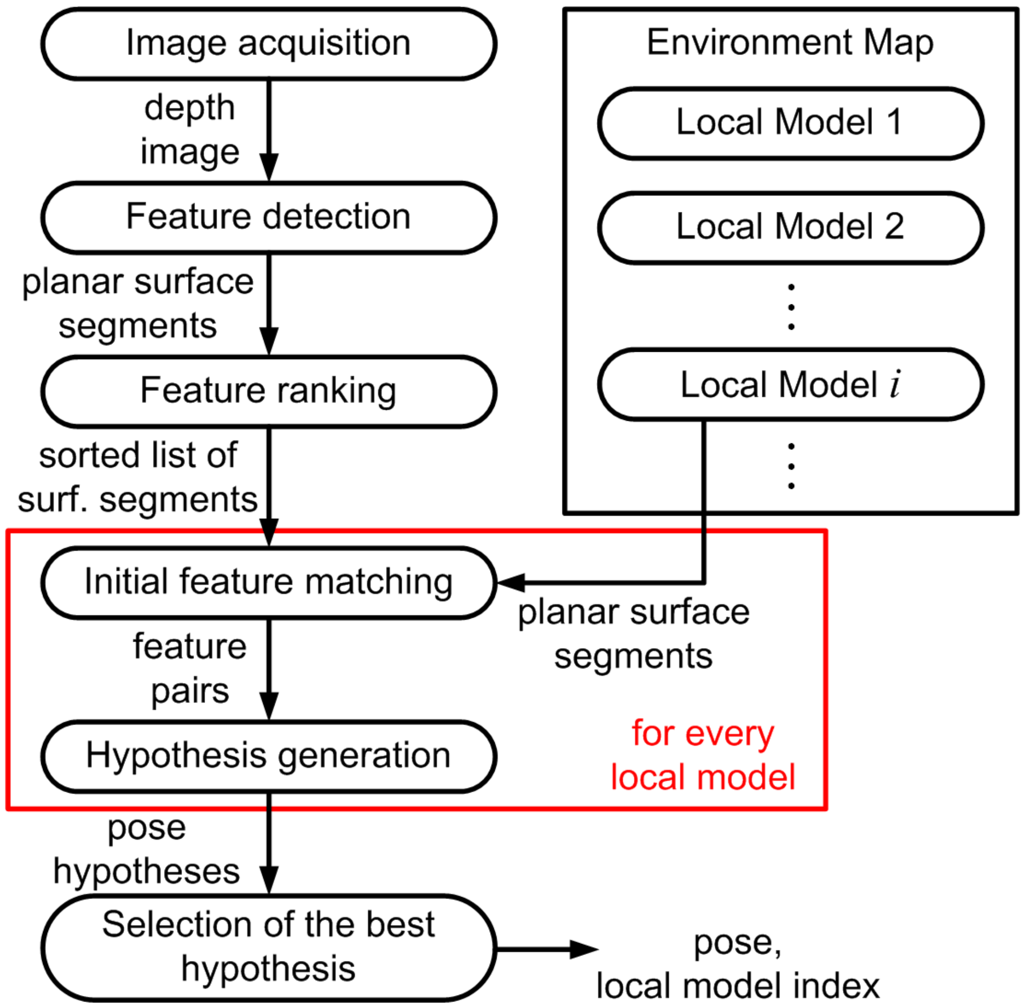
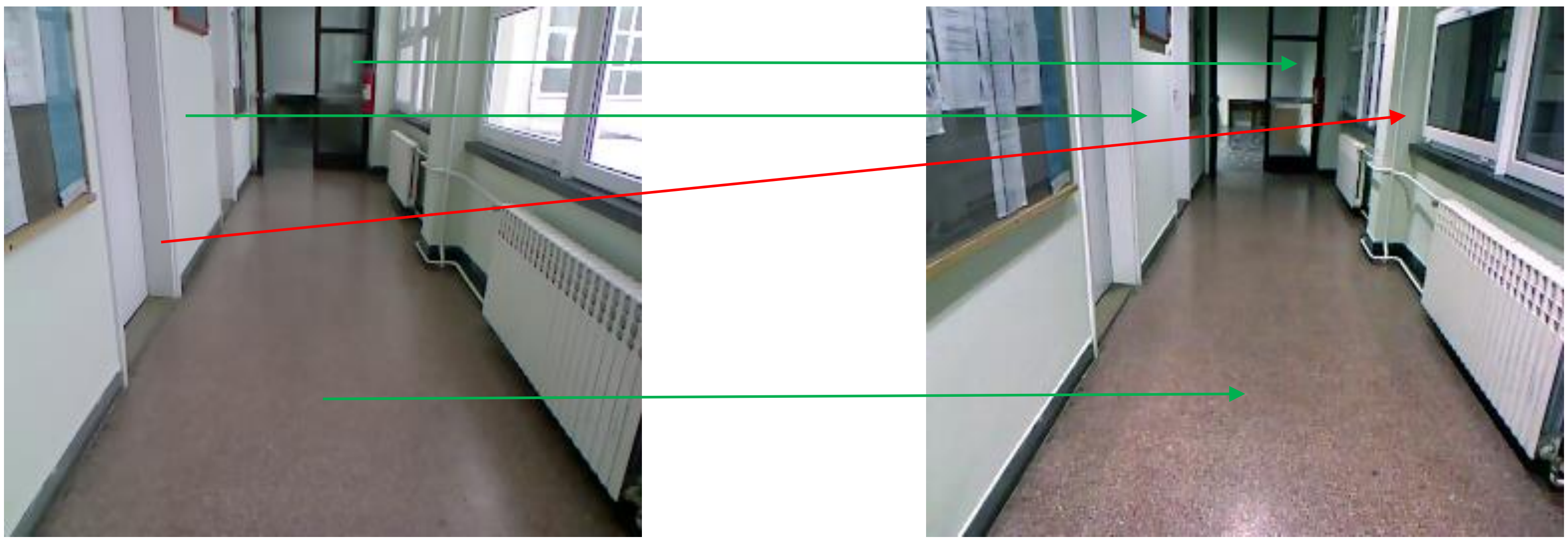


Image acquisition
depth image
Feature detection
planar surface segments
Feature ranking
sorted list of surf. segments
Initial feature matching
feature pairs
Hypothesis generation
pose hypotheses
Selection of the best hypothesis
pose, local model index
Environment Map
Local Model 1
Local Model 2
Local Model i
planar surface segments
for every local model

Hypothesis Generation

- 3 planes intersecting in a common point uniquely define a reference frame.
- Correspondences between 3 planar patches in the scene and 3 planar patches of a local model → robot pose w.r.t. local model reference frame.
- Many combinations!



- What about local feature descriptors ? (Common approach with point features)
- Problem: planar patches don't have distinctive stable geometric features.

Hypothesis Generation

- Solutions:
  1. Generate more probable hypotheses before less probable ones.
     - Approach proposed in [1]:
       - uses larger planar patches before smaller
       - avoids repeated combinations of features
       - computes orientation and 2 translational DoFs using correspondences of 2 features, and the remaining DoF by 1D evidence accumulation
  2. Color and texture descriptors [2]

[1] Cupec, Nyarko, Filko & Petrović, Fast Pose Tracking Based on Ranked 3D Planar Patch Correspondences. In Proc. IFAC Symposium on Robot Control, Dubrovnik, Croatia, 2012
[2] Filko, Cupec & Nyarko, Evaluation of Color and Texture Descriptors for Matching of Planar Surfaces in Global Localization Scheme. Robotics and autonomous systems, 2016

Hypothesis Probability Estimation

- Probabilistic approach [3]:

pose

Feature correspondences

H_k – hypothesis ($\boldsymbol{w}_k$, $\boldsymbol{c}_k$)

Z – measurement: surface segments F_1, F_2, ..., F_n, ...

F_i – surface segment ($\boldsymbol{n}$, ρ)

distance from the origin

normal

$$P\left(H_k \mid Z\right) = \frac{p\left(Z \mid H_k\right) P\left(H_k\right)}{p\left(Z\right)}$$

[3] Cupec, Nyarko, Filko, Kitanov & Petrović, Place recognition based on matching of planar surfaces and line segments. International journal of robotics research, 2015

Hypothesis Probability Estimation

- Probabilistic approach [3]:

  - one of the hypotheses from a hypothesis set is correct
  - only one hypothesis is correct

$$P\left(H_k \mid Z\right) = \frac{p\left(Z \mid H_k\right) P\left(H_k\right)}{\sum_{H_{k'} \in \tilde{\chi}} p\left(Z \mid H_{k'}\right) P\left(H_{k'}\right)}$$

  - prior probability of all hypotheses is equal

$$P\left(H_k \mid Z\right) = \frac{p\left(Z \mid H_k\right)}{\sum_{H_{k'} \in \tilde{\chi}} p\left(Z \mid H_{k'}\right)}$$

[3] Cupec, Nyarko, Filko, Kitanov & Petrović, Place recognition based on matching of planar surfaces and line segments. International journal of robotics research, 2015

Hypothesis Probability Estimation

- Probabilistic approach [3]:

  - *independent surface model (ISM)* – errors in the parameter measurements of the surface segments are mutually independent.

$$p\left(Z \mid H_k\right)=\prod_{c_{ki}\neq 0} p\left(F_i, F_j' \mid c_{ki}=j, \boldsymbol{w}_k\right)\prod_{c_{ki}=0} p\left(F_i \mid c_{ki}=0\right)$$

hypothesis H_k assumes that i-th detected surface segment corresponds to j-th model surface segment

i-th detected surface segment don't correspond to any model surface segment

[3] Cupec, Nyarko, Filko, Kitanov & Petrović, Place recognition based on matching of planar surfaces and line segments. International journal of robotics research, 2015

Hypothesis Probability Estimation

$$p\left(Z \mid H_k\right) = \prod_{c_{ki} \neq 0} p\left(F_i, F_j' \mid c_{ki} = j, \boldsymbol{w}_k\right) \prod_{c_{ki} = 0} p\left(F_i \mid c_{ki} = 0\right)$$

- If F_i is not matched to any local model surface segment – normal is uniformly distributed

$$p\left(F_i \mid c_{ki} = 0\right) = \frac{1}{2\pi}$$

Hypothesis Probability Estimation

$$p\left(Z \mid H_k\right) = \prod_{c_{ki} \neq 0} p\left(F_i, F_j' \mid c_{ki} = j, \boldsymbol{w}_k\right) \prod_{c_{ki} = 0} p\left(F_i \mid c_{ki} = 0\right)$$

- If F_i is matched to F_j'

$$p\left(F_i, F_j' \mid c_{ki} = j, \boldsymbol{w}\right) \approx \frac{1}{2\pi\sqrt{\det\left({}^{F'}\boldsymbol{\Sigma}_{s,i} + \boldsymbol{\Sigma}_{s',j}\right)}} \exp\left(-\frac{L_{ij}}{2}\right)$$

the first two components of the normal of F_i represented in the reference frame of the corresponding model surface

$${}^{F'}\hat{\boldsymbol{s}}_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} {}^{M}\boldsymbol{R}_{F'}^{T} \boldsymbol{R}\left(\phi_k\right) {}^{C}\boldsymbol{R}_F \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

orientation of the camera w.r.t. the local model

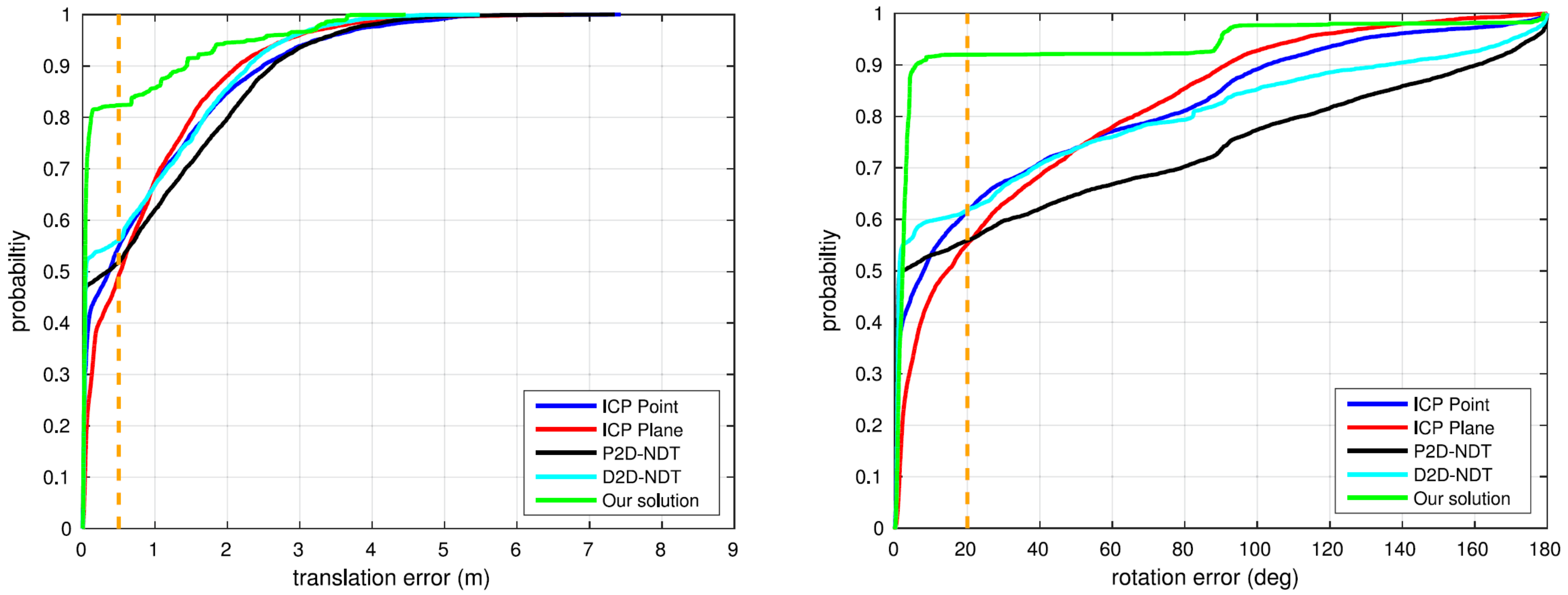
$$\hat{\boldsymbol{s}}_{ij} = \left({}^{F'}\boldsymbol{\Sigma}_{s,i}^{-1} + \boldsymbol{\Sigma}_{s',j}'^{-1}\right)^{-1} {}^{F'}\boldsymbol{\Sigma}_{s,i}^{-1} \, {}^{F'}\hat{\boldsymbol{s}}_i$$

uncertainty of the normal of F_i

$$L_{ij} = \left({}^{F'}\hat{\boldsymbol{s}}_i - \hat{\boldsymbol{s}}_{ij}\right)^T {}^{F'}\boldsymbol{\Sigma}_{s,i}^{-1} \left({}^{F'}\hat{\boldsymbol{s}}_i - \hat{\boldsymbol{s}}_{ij}\right) + \hat{\boldsymbol{s}}_{ij}^T \boldsymbol{\Sigma}_{s',j}'^{-1} \hat{\boldsymbol{s}}_{ij}$$

LIDAR Scan Registration in a SLAM system [4]

- “Challenging data sets for point cloud registration algorithms” [5]
  - ➢ Apartment dataset

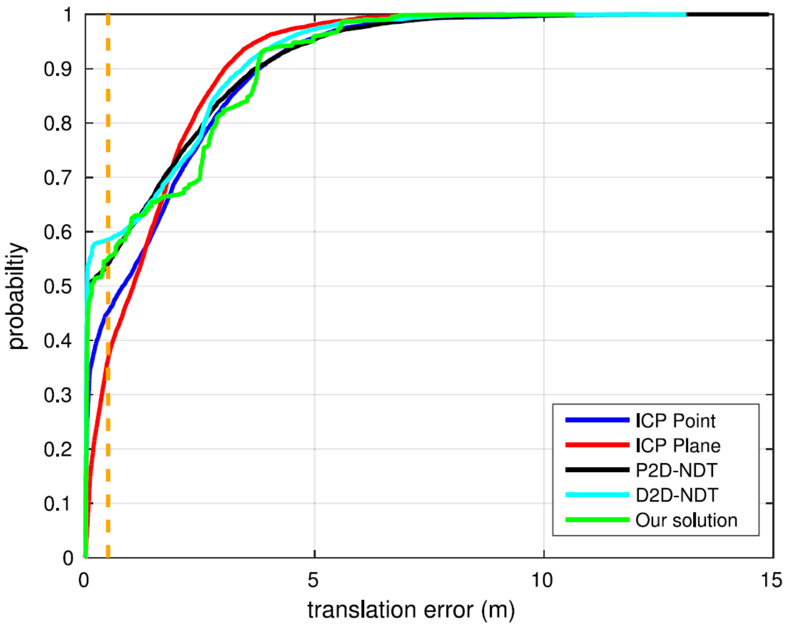


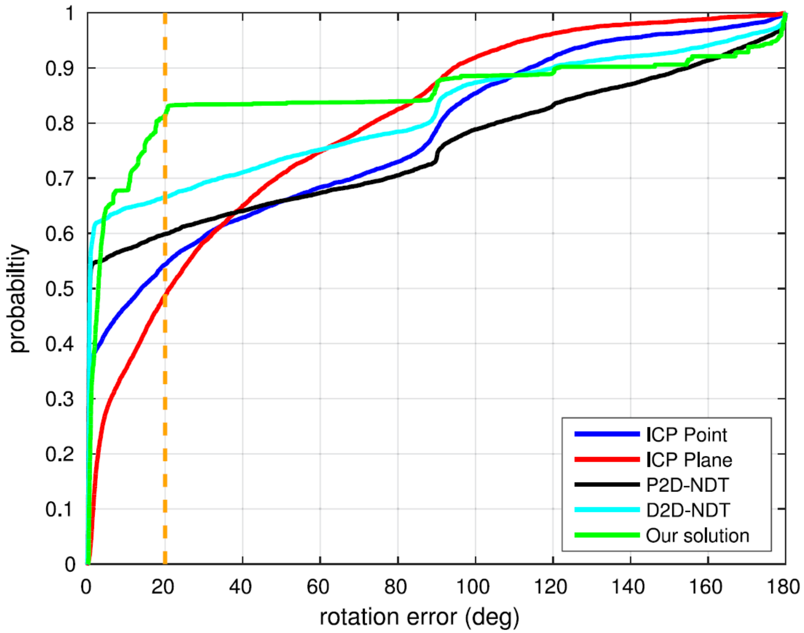
[4] Lenac, Kitanov, Cupec & Petrović, Fast planar surface 3D SLAM using LIDAR. Robotics and Autonomous Systems, 2017

[5] F. Pomerleau, F. Colas, R. Siegwart, S. Magnenat, Comparing ICP variants on real-world data sets, Auton. Robots, 2013

LIDAR Scan Registration in a SLAM system [4]

- "Challenging data sets for point cloud registration algorithms" [5]
  - Stairs dataset



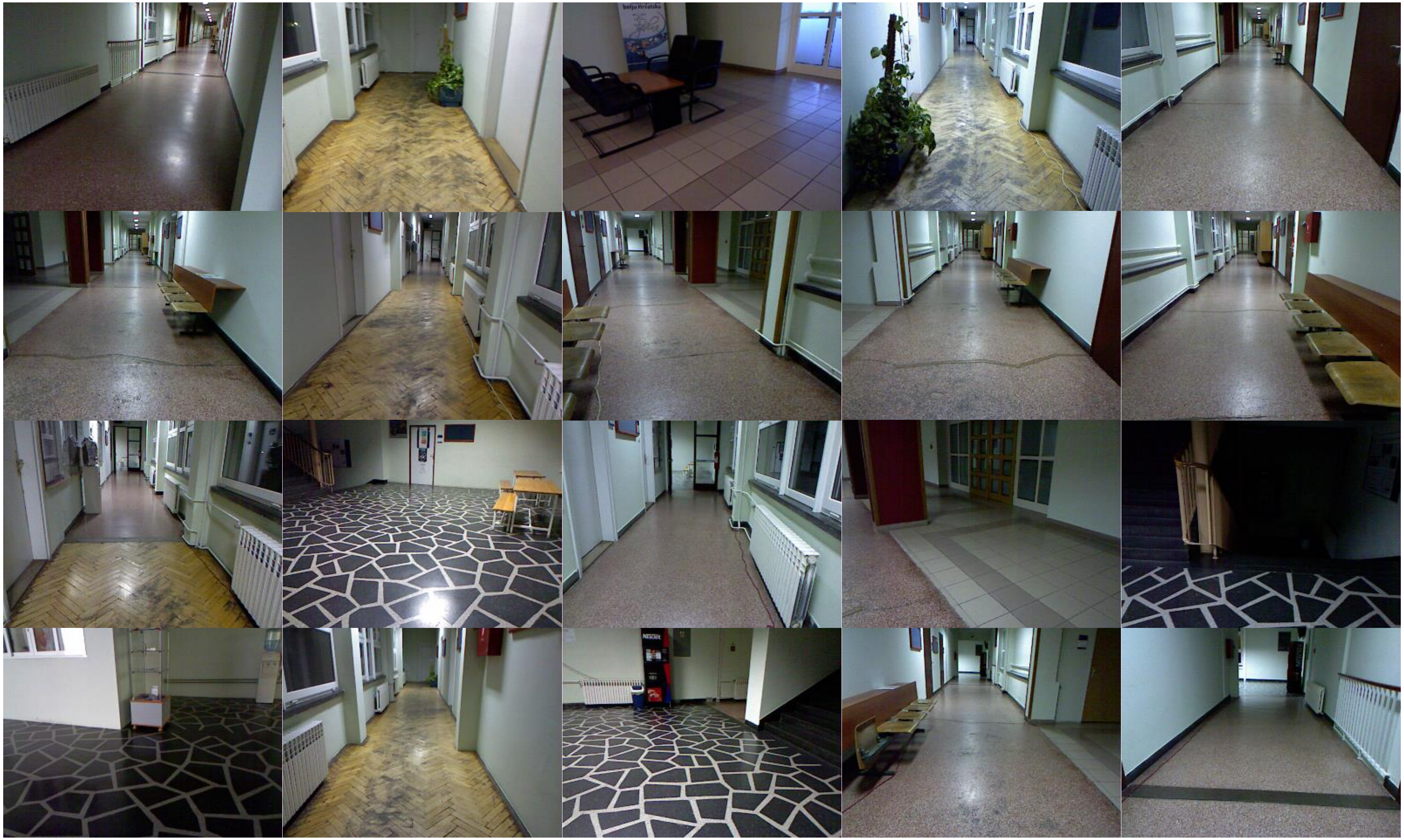


[4] Lenac, Kitanov, Cupec & Petrović, Fast planar surface 3D SLAM using LIDAR. Robotics and Autonomous Systems, 2017

[5] F. Pomerleau, F. Colas, R. Siegwart, S. Magnenat, Comparing ICP variants on real-world data sets, Auton. Robots, 2013

Place Recognition with RGB-D Camera

Place Recognition with RGB-D Camera



Place Recognition with RGB-D Camera

Experimental Evaluation: Dataset

subset	different lighting conditions then in the corresponding reference image	dynamic objects
1	-	-
2	-	+
3	+	-
4	+	+

Experimental Evaluation: Dataset

subset 1

subset 3

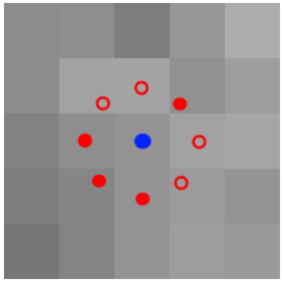
subset 4

Comparison to Related Research

subset	total	correct recognitions					
		FAB-MAP		DLoopDetector		proposed approach	
		number	%	number	%	number	%
1	68	66	**97**	56	82	63	93
2	676	574	85	577	85	606	**90**
3	246	114	46	6	2	204	**83**
4	1175	453	39	49	4	984	**84**
Σ	2165	1207	56	688	32	1857	**86**

Color and Texture Features for Surface Matching

1. Color/texture histogram is assigned to each planar surface segment F
    - RGB
    - normalized rgb
    - XYZ
    - CIE L*a*b*
    - HSV
    - $O_1O_2O_3$ (opponent color space)
    - LBP (Local Binary Patterns)
    - LBP RIU (rotational invariant version – number of transitions)
    - LBP RIU VAR (variance information integrated)



Color and Texture Features for Surface Matching

1. Color histogram is assigned to each planar surface segment F
2. Matching of scene surface samples with local model surface samples according to normalized histogram intersection

$$h\left(\mathbf{I},\mathbf{M}\right)=\frac{\sum_{i=1}^{n}\min\left(I_i,M_i\right)}{\sum_{i=1}^{n}M_i}$$

$\mathbf{I}$ – color histogram assigned to scene surface F

$\mathbf{M}$ – color histogram assigned to model surface F'

I_i – i-th bin of scene surface color histogram $\mathbf{I}$

M_i – i-th bin of model surface color histogram $\mathbf{M}$

Surface segment pair (F, F') is considered in the hypothesis generation process only if

$$h\left(\mathbf{I},\mathbf{M}\right)\geq\varepsilon_s$$

Experimental Evaluation: Color and Texture Features

Descriptor	ΔHyp. [%]	The top hyp. is corect [%]	Localization time change [%]	Image processing time change [%]	Total time change [%]	No. of generated hypotheses [%]	Change of FCHI [%]
RGB32S	-4,04	22,98	214,07	323,75	218,98	-49,55	-62,35
RGB32S (*)	-1,29	22,77	211,55	323,75	216,56	-29,49	-50,39
RGB8 (5)	-5,68	12,34	-59,81	43,19	-55,20	-54,56	-56,27
RGB8 (5*)	-2,07	14,26	-15,72	43,19	-13,08	-37,43	-43,67
XYZ8S	-2,93	9,57	-13,80	95,23	-8,92	-38,78	-36,19
XYZ8S (*)	-1,03	9,79	14,94	95,23	18,53	-22,76	-25,21
XYZ8 (5)	-3,44	12,34	-49,21	44,98	-45,00	-46,58	-41,01
XYZ8 (5*)	-1,03	12,55	-10,79	44,98	-8,29	-29,99	-35,09
HSV8	-1,55	25,11	-35,98	46,71	-32,29	-42,94	-51,76
HSV8 (*)	0,00	22,55	-1,21	46,71	0,93	-25,76	-48,56
HSV8 (5)	-3,01	16,17	-53,23	46,71	-48,76	-47,53	-50,80
HSV8 (5*)	-0,77	16,17	-15,46	46,71	-12,68	-30,74	-47,12
Lab32	-2,41	21,06	-36,81	46,81	-33,07	-46,48	-51,52
Lab32 (*)	-0,52	19,79	-1,86	46,81	0,32	-30,46	-45,94
Lab8 (5)	-2,84	18,30	-40,15	40,21	-36,56	-41,52	-45,71
Lab8 (5*)	-0,69	15,53	-6,21	40,21	-4,14	-26,55	-32,45
RGC16S	-4,22	16,17	-32,52	62,95	-28,25	-41,72	-42,62
RGC16S (*)	-1,20	17,23	-1,55	62,95	1,33	-28,17	-35,63
RGC16 (5)	-3,44	9,15	-39,48	41,86	-35,85	-39,27	-23,50
RGC16 (5*)	-0,77	11,70	-11,48	41,86	-9,09	-26,69	-7,73
Opp16S	-3,10	19,15	-42,33	62,71	-37,63	-48,43	-48,81
Opp16S (*)	-1,12	18,94	-10,19	62,71	-6,93	-34,84	-41,55
Opp16S (5)	-3,10	13,62	-46,81	62,71	-41,92	-45,55	-39,65
Opp16S (5*)	-1,12	12,98	-16,26	62,71	-12,73	-32,55	-41,56
LBP16	-5,42	6,38	-23,77	252,92	-11,40	-31,79	-24,34
LBP16 (*)	-0,95	10,21	17,98	252,40	28,47	-5,90	-15,24
LBP RIU 24	-10,33	-0,43	-43,14	349,26	-25,59	-34,90	-21,44
LBP RIU 24 (*)	-2,50	2,55	3,45	349,08	18,90	-5,89	2,56
LBP RIU VAR 24	-6,28	7,23	-32,06	405,95	-12,47	-45,72	-30,97
LBP RIU VAR 24 (*)	-0,52	10,21	19,65	405,99	36,93	-11,21	-2,73

Conclusion

- A place recognition system based on planar surface segments and edge line segments is presented.
- The advantage of using depth information over some state-of-the-art methods which use RGB image only is demonstrated.
- The advantage of using color and texture information in addition to 3D geometry is shown.
- The applicability of the presented approach for visual odometry using LIDAR is investigated and a comparison to 4 state-of-the-art methods is made.
- Expanding the camera FoV by taking a series of images from diffrent viewing angles in the same viewpoint improves the performance of the system.

Thank you for your attention.