Exploiting semantic information in indoor environments

Mathias Fassini Mantelli

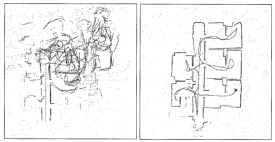
Federal University of Rio Grande do Sul
Institute of Informatics
Postgraduate Program in Computing

Advisor: Prof[a]. Mariana Luderitz Kolberg Fernandes
Co-advisor: Prof. Renan Maffei

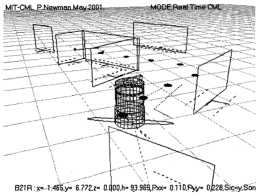
May 18, 2022

First years of mobile robotics

- Ages of mobile robotics[1]:
  - Classical age (1986-2004):
    - **Introduction** of the **main probabilistic** formulations for **SLAM**
    - **Lidar** and **sonar** sensors



(a) Online mapping.[2]



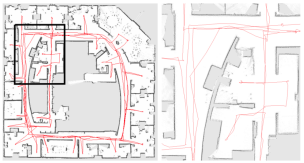
(b) Real Time CML.[3]

Figure: Initial works on SLAM.

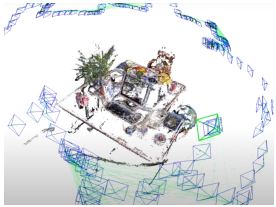
[1] Cesar, Cadena, et al. "*Simultaneous Localization And Mapping: Present Future and the Robust-Perception Age.*" arXiv preprint arXiv: 1606.05830. 2016.
[2] Thrun, Sebastian. "*An Online Mapping Algorithm for Teams of Mobile Robots*". Carnegie-Mellon Univ Pittsburgh PA School of Computer Science, 2000.
[3] Newman, Paul, et al. "*Explore and return: Experimental validation of real-time concurrent mapping and localization.*" ICRA, 2002

First years of mobile robotics

- Ages of mobile robotics[1]:
  - Classical age (1986-2004):
    - **Introduction** of the **main probabilistic** formulations for **SLAM**
    - **Lidar** and **sonar** sensors
  - Algorithmic-analysis age (2004-2015):
    - **Study** of **fundamental properties** of **SLAM**
    - **Visual** sensors



(a) GMapping.[4]



(b) ORB-SLAM.[5]

Figure: Improved SLAM and Visual SLAM.

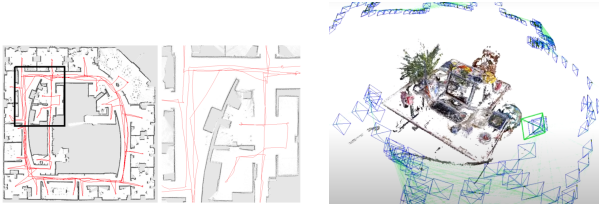
[1] Cesar, Cadena, et al. "*Simultaneous Localization And Mapping: Present Future and the Robust-Perception Age.*" arXiv preprint arXiv: 1606.05830. 2016.

[4] Grisetti, G., Stachniss, C., and Burgard, W. "*Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling.*" ICRA. 2005.

[5] Mur-Artal, R., Montiel, J.M.M., and Tardos, J.D. "*ORB-SLAM: a versatile and accurate monocular SLAM system.*" IEEE Transactions on Robotics. 2015.

First years of mobile robotics

- Ages of mobile robotics[1]:
  - Classical age (1986-2004):
    - **Introduction** of the **main probabilistic** formulations for **SLAM**
    - **Lidar** and **sonar** sensors
  - Algorithmic-analysis age (2004-2015):
    - **Study** of **fundamental properties** of **SLAM**
    - **Visual** sensors
  - **Geometric perception**



(a) GMapping.[4] (b) ORB-SLAM.[5]

Figure: Improved SLAM and Visual SLAM.

[1] Cesar, Cadena, et al. "*Simultaneous Localization And Mapping: Present Future and the Robust-Perception Age.*" arXiv preprint arXiv: 1606.05830. 2016.

[4] Grisetti, G., Stachniss, C., and Burgard, W. "*Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling.*" ICRA. 2005.

[5] Mur-Artal, R., Montiel, J.M.M., and Tardos, J.D. "*ORB-SLAM: a versatile and accurate monocular SLAM system.*" IEEE Transactions on Robotics. 2015.

Geometric perception

- Raw sensor readings

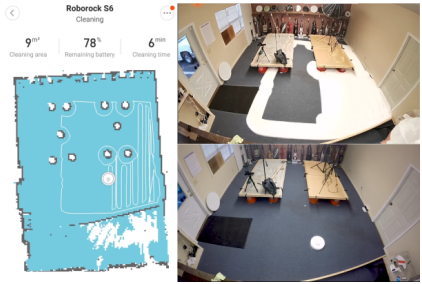


Figure: Vacuum cleaner robot in operation.[6]

[6] Extracted from youtube.com/watch?v=5O8VmDiab3w

Geometric perception

- Raw sensor readings
- Useful for many **robotic tasks**

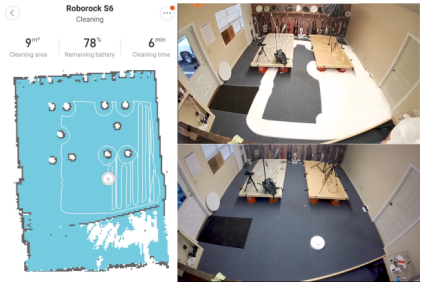


Figure: Vacuum cleaner robot in operation.[6]

[6] Extracted from youtube.com/watch?v=5O8VmDiab3w

Geometric perception

- Raw sensor readings
- Useful for many **robotic tasks**
- Efficient for **building maps** and **state estimation**

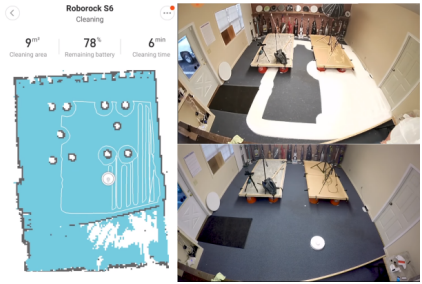


Figure: Vacuum cleaner robot in operation.[6]

[6] Extracted from youtube.com/watch?v=5O8VmDiab3w

Geometric perception

- Raw sensor readings
- Useful for many **robotic tasks**
- Efficient for **building maps** and **state estimation**
- Suitable for **path-planning** and **obstacle avoidance**

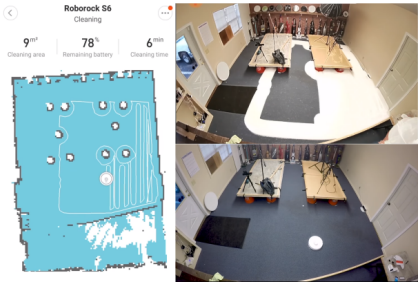


Figure: Vacuum cleaner robot in operation.[6]

[6] Extracted from youtube.com/watch?v=5O8VmDiab3w

Geometric perception

- Raw sensor readings
- Useful for many **robotic tasks**
- Efficient for **building maps** and **state estimation**
- Suitable for **path-planning** and **obstacle avoidance**
- **Limitations**:

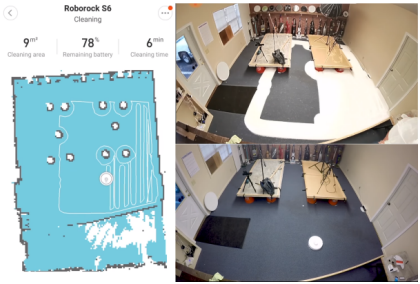


Figure: Vacuum cleaner robot in operation.[6]

[6] Extracted from youtube.com/watch?v=5O8VmDiab3w

Geometric perception

- Raw sensor readings
- Useful for many **robotic tasks**
- Efficient for **building maps** and **state estimation**
- Suitable for **path-planning** and **obstacle avoidance**
- **Limitations**:
  - Type of maps (free, occupied, unknown)

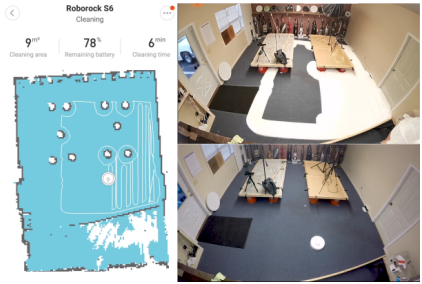


Figure: Vacuum cleaner robot in operation.[6]

[6] Extracted from youtube.com/watch?v=5O8VmDiab3w

Geometric perception

- Raw sensor readings
- Useful for many **robotic tasks**
- Efficient for **building maps** and **state estimation**
- Suitable for **path-planning** and **obstacle avoidance**
- **Limitations**:
  - Type of maps (free, occupied, unknown)
  - Does not distinguish obstacles

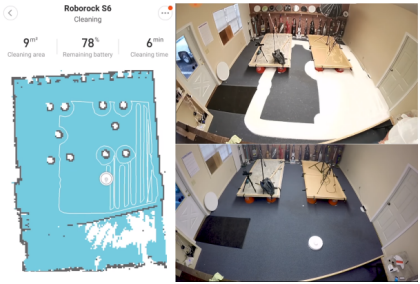


Figure: Vacuum cleaner robot in operation.[6]

[6] Extracted from youtube.com/watch?v=5O8VmDiab3w

Geometric perception

- Raw sensor readings
- Useful for many **robotic tasks**
- Efficient for **building maps** and **state estimation**
- Suitable for **path-planning** and **obstacle avoidance**
- **Limitations**:
  - Type of maps (free, occupied, unknown)
  - Does not distinguish obstacles
  - Insufficient for high-level tasks

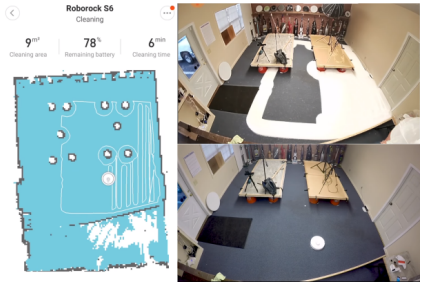


Figure: Vacuum cleaner robot in operation.[6]

[6] Extracted from youtube.com/watch?v=5O8VmDiab3w

Geometric perception

- Raw sensor readings
- Useful for many **robotic tasks**
- Efficient for **building maps** and **state estimation**
- Suitable for **path-planning** and **obstacle avoidance**
- **Limitations**:
  - Type of maps (free, occupied, unknown)
  - Does not distinguish obstacles
  - Insufficient for high-level tasks
- How to **overcome** these **limitations**?

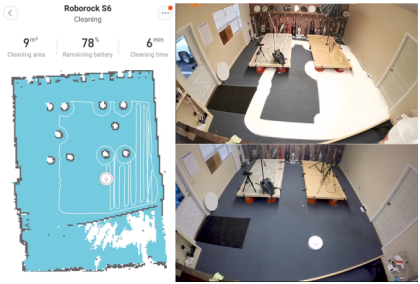


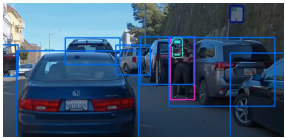
Figure: Vacuum cleaner robot in operation.[6]

[6] Extracted from youtube.com/watch?v=5O8VmDiab3w

Expand the geometric perception

- Ages of mobile robotics[1]:
  - Robust-perception age (2015-now):
    - **Understand** the **concepts** of parts of the **sensor readings** (Semantic information)

(A) The sire of the fire truck.



(B) The car door.

Figure: Self-Driving System of an autonomous driving car.[7]

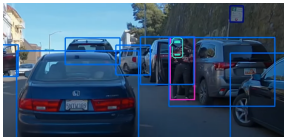
[1] Cesar, Cadena, et al. "*Simultaneous Localization And Mapping: Present Future and the Robust-Perception Age.*" arXiv preprint arXiv: 1606.05830. 2016.

[7] Extracted from `youtube.com/watch?v=BVRMh9NO9Cs`

Expand the geometric perception

- Ages of mobile robotics[1]:
  - Robust-perception age (2015-now):
    - **Understand** the **concepts** of parts of the **sensor readings** (Semantic information)
    - **Associate them** to the **map** (Semantic mapping)

(A) The sire of the fire truck.



(B) The car door.

Figure: Self-Driving System of an autonomous driving car.[7]

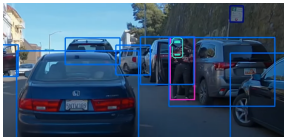
[1] Cesar, Cadena, et al. "*Simultaneous Localization And Mapping: Present Future and the Robust-Perception Age.*" arXiv preprint arXiv: 1606.05830. 2016.

[7] Extracted from `youtube.com/watch?v=BVRMh9NO9Cs`

Expand the geometric perception

- Ages of mobile robotics[1]:
  - Robust-perception age (2015-now):
    - **Understand** the **concepts** of parts of the **sensor readings** (Semantic information)
    - **Associate them** to the **map** (Semantic mapping)
    - Enhance robot's autonomy and robustness, **facilitate** more **complex tasks**

(A) The sire of the fire truck.



(B) The car door.

Figure: Self-Driving System of an autonomous driving car.[7]

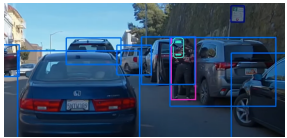
[1] Cesar, Cadena, et al. "*Simultaneous Localization And Mapping: Present Future and the Robust-Perception Age.*" arXiv preprint arXiv: 1606.05830. 2016.

[7] Extracted from `youtube.com/watch?v=BVRMh9NO9Cs`

Expand the geometric perception

- Ages of mobile robotics[1]:
  - Robust-perception age (2015-now):
    - **Understand** the **concepts** of parts of the **sensor readings** (Semantic information)
    - **Associate them** to the **map** (Semantic mapping)
    - Enhance robot's autonomy and robustness, **facilitate** more **complex tasks**
    - Essential for **high-level reasoning** and **human-robot interaction**

(A) The sire of the fire truck.



(B) The car door.

Figure: Self-Driving System of an autonomous driving car.[7]

[1] Cesar, Cadena, et al. "*Simultaneous Localization And Mapping: Present Future and the Robust-Perception Age.*" arXiv preprint arXiv: 1606.05830. 2016.

[7] Extracted from `youtube.com/watch?v=BVRMh9NO9Cs`

Semantic in the robotics field

Amount of publications returned by **Google Scholar** for the keywords "**Semantic**, **robotics**":

- until 2013: 38.600
- 2013 - 2015: 15.800
- 2015 - 2017: 17.300
- 2017 - 2019: 21.100
- 2019 - now: 25.500

Semantic in the robotics field

Amount of publications returned by **Google Scholar** for the keywords "**Semantic**, **robotics**":

- until 2013: 38.600
- 2013 - 2015: 15.800
- 2015 - 2017: 17.300
- 2017 - 2019: 21.100
- 2019 - now: 25.500

An **evidence** of the **increasing use of semantics in robotics** by the research community

Hypothesis

Semantic information *associated with the* **spatial** *and* **temporal organization** *of the environment help mobile robotics to* **overcome** *the* **limitations** *to deal with* **high-level tasks**

Semantic information within mobile robotics

- Three questions:

Semantic information within mobile robotics

- Three questions:
  - **Which** type of **semantic information** is **relevant** to the **task**?

Semantic information within mobile robotics

- Three questions:
  - **Which** type of **semantic information** is **relevant** to the **task**?
  - **How** to perform the **inference/estimation** of the semantic information?

Semantic information within mobile robotics

- Three questions:
  - **Which** type of **semantic information** is **relevant** to the **task**?
  - **How** to perform the **inference/estimation** of the semantic information?
  - **How** to **use** the **semantic information** to improve the **robot's performance**?

Semantic information within mobile robotics

- Three questions:
  - **Which** type of **semantic information** is **relevant** to the **task**?
  - **How** to perform the **inference/estimation** of the semantic information?
  - **How** to **use** the **semantic information** to improve the **robot's performance**?
- We **investigate** these questions in the context of a **high-level task**: **object search** (OS)

Object search problem

- **Challenging** and unsolved **task for robots**

Object search problem

- **Challenging** and unsolved **task for robots**
- **Goal:** the robot **searches** for a **target object** in an **unknown environment**

Object search problem

- **Challenging** and unsolved **task for robots**
- **Goal:** the robot **searches** for a **target object** in an **unknown environment**
- Computation of the **optimal solution** *vs* Estimation that the current goal **is not promising**

Object search problem

- **Challenging** and unsolved **task for robots**
- **Goal:** the robot **searches** for a **target object** in an **unknown environment**
- Computation of the **optimal solution** *vs* Estimation that the current goal **is not promising**
- The robot's **perception** is **crucial**

Object search problem

- **Challenging** and unsolved **task for robots**
- **Goal:** the robot **searches** for a **target object** in an **unknown environment**
- Computation of the **optimal solution** *vs* Estimation that the current goal **is not promising**
- The robot's **perception** is **crucial**
- We **complement** the robot's **perception** by **inferring semantic information** from:

Object search problem

- **Challenging** and unsolved **task for robots**
- **Goal:** the robot **searches** for a **target object** in an **unknown environment**
- Computation of the **optimal solution** *vs* Estimation that the current goal **is not promising**
- The robot's **perception** is **crucial**
- We **complement** the robot's **perception** by **inferring semantic information** from:
  - Texts

Object search problem

- **Challenging** and unsolved **task for robots**
- **Goal:** the robot **searches** for a **target object** in an **unknown environment**
- Computation of the **optimal solution** *vs* Estimation that the current goal **is not promising**
- The robot's **perception** is **crucial**
- We **complement** the robot's **perception** by **inferring semantic information** from:
  - Texts
  - Semi-dynamic objects

Object search problem

- **Challenging** and unsolved **task for robots**
- **Goal:** the robot **searches** for a **target object** in an **unknown environment**
- Computation of the **optimal solution** *vs* Estimation that the current goal **is not promising**
- The robot's **perception** is **crucial**
- We **complement** the robot's **perception** by **inferring semantic information** from:
  - Texts
  - Semi-dynamic objects
- **Human-like** behavior

Exploiting semantic information in indoor environments

Mathias Fassini Mantelli

Federal University of Rio Grande do Sul
Institute of Informatics
Postgraduate Program in Computing

Advisor: Prof[a]. Mariana Luderitz Kolberg Fernandes
Co-advisor: Prof. Renan Maffei

May 18, 2022

Equations - Semantic OS System based on text

- Map segmentation: $\Psi(c_k) = \sum_c^{\mathbf{T}} Q(c)K(\|c - c_k\|)$

$$Q(c) = \begin{cases} 1 & \text{, if } c \text{ is a free cell} \\ 0 & \text{, otherwise.} \end{cases}$$

$$K(d) = \begin{cases} a & \text{, if } d \leq r \\ 0 & \text{, otherwise,} \end{cases}$$

$$\Upsilon(c_k) = \lfloor \Psi(c_k)/\delta \rfloor$$

Equations - Semantic OS System based on text

- Growing Direction factor: $\zeta\big(S(c)\big) = \dfrac{\Big(L^{<}\big(S(c)\big) - L^{>}\big(S(c)\big)\Big)}{\max\Big(L^{<}\big(S(c)\big) + L^{>}\big(S(c)\big), w_g\Big)},$

$$\gamma(\theta_f(c)) = 1.0 + \left|\frac{\theta_f(c) - \theta_i\big(S(c)\big)}{\pi}\right| * -2.0$$

$$\varphi_g(c) = \frac{\zeta\big(S(c)\big) * \gamma\big(\theta_f(c)\big) + 1.0}{2.0}$$

Equations - Semantic OS System based on text

- Parity factor: $\varphi_p(c) = 0.5 + \frac{L^{=}\left(S(c)\right) - L^{\neq}\left(S(c)\right)}{\max\left(L^{=}\left(S(c)\right) + L^{\neq}\left(S(c)\right), w_p\right)} * 0.5$

Equations - Semantic OS System based on text

Goal Doors	Value of α				Indices (m)
	Greedy	Semantic			
	0,00%	80,00%	90,00%	100,00%	
54	117,28	226,48	199,68	132,75	Median
	121,94	226,66	200,31	134,77	Average
	41,59	2,80	2,22	30,64	Std. Dev.
	82,19	221,42	197,88	100,40	Min
	230,01	231,11	204,13	186,80	Max
55	132,09	72,44	64,20	114,96	Median
	132,68	72,61	63,88	136,47	Average
	46,71	0,28	1,06	36,03	Std. Dev.
	74,45	72,34	60,89	113,42	Min
	243,17	73,18	64,44	225,75	Max
111	171,78	236,78	137,85	60,86	Median
	170,69	243,37	135,02	67,34	Average
	70,15	71,52	4,39	16,99	Std. Dev.
	65,99	136,81	128,22	46,73	Min
	275,52	383,46	138,75	108,38	Max
124	196,31	61,88	59,75	67,86	Median
	148,50	61,86	63,31	67,05	Average
	70,09	0,10	11,24	17,08	Std. Dev.
	49,56	61,68	59,64	49,58	Min
	200,36	62,01	95,30	86,16	Max

Color scale

Smallest value of the row — Biggest value of the row

Figure: Normal map.

Equations - Semantic OS System based on text

Goal Doors	Value of α				Indices (m)
	Greedy	Semantic			
	0,00%	80,00%	90,00%	100,00%	
54	82,03	29,97	25,22	131,57	Median
	98,88	29,68	24,98	135,18	Average
	41,52	0,89	0,90	15,13	Std. Dev.
	72,61	27,16	22,44	118,57	Min
	207,20	30,03	25,40	166,40	Max
55	152,67	69,26	41,35	128,48	Median
	154,29	62,34	46,28	109,80	Average
	54,78	10,38	8,13	41,57	Std. Dev.
	63,60	49,44	41,11	35,00	Min
	232,54	71,31	61,34	138,67	Max
111	273,64	250,07	219,97	214,02	Median
	200,16	230,27	205,45	209,96	Average
	102,57	54,63	45,38	14,65	Std. Dev.
	61,86	153,58	124,08	188,42	Min
	278,57	292,28	267,79	231,98	Max
124	205,26	177,16	160,49	173,34	Median
	165,73	145,59	137,05	156,20	Average
	82,65	42,73	38,69	60,25	Std. Dev.
	58,15	92,99	80,53	35,59	Min
	291,36	181,82	162,30	200,52	Max

Color scale

Smallest value of the row — Biggest value of the row

Figure: Inverse map.

Equations - Semantic OS System based on text

Goal Doors	Value of α				Indices (m)
	Greedy	Semantic			
	0,00%	80,00%	90,00%	100,00%	
76	341,49	190,95	168,68	114,34	Median
	314,90	184,36	172,17	125,41	Average
	131,74	21,80	11,11	25,96	Std. Dev.
	143,97	123,07	167,93	96,47	Min
	491,69	199,53	203,74	171,91	Max
135	552,44	285,57	283,22	318,24	Median
	433,75	282,87	304,28	326,52	Average
	194,93	9,18	39,92	27,01	Std. Dev.
	93,17	256,95	280,69	292,30	Min
	555,11	287,15	380,32	374,25	Max
148	554,20	151,65	368,56	405,21	Median
	525,30	151,64	369,44	395,57	Average
	167,12	0,75	1,72	30,34	Std. Dev.
	114,65	149,97	368,09	327,40	Min
	671,81	153,01	372,60	434,30	Max
185	132,18	130,84	130,94	92,01	Median
	167,09	130,87	131,00	102,94	Average
	166,77	0,22	0,29	57,33	Std. Dev.
	51,79	130,57	130,71	51,72	Min
	552,68	131,22	131,55	193,82	Max

Color scale

Smallest value of the row — Biggest value of the row

Figure: Hotel map.

Equations - Semantic OS System based on text

Goal Doors	Value of α				Indices (m)
	Greedy	Semantic			
	0,00%	80,00%	90,00%	100,00%	
765	143,30	26,52	26,01	25,12	Median
	155,03	33,41	28,88	24,52	Average
	25,82	10,35	4,60	1,98	Std. Dev.
	140,63	24,86	24,81	18,88	Min
	221,49	48,96	34,29	25,33	Max

Color scale

Smallest value of the row — Biggest value of the row

Figure: KTH map.