Data Association Aware Semantic Mapping and Localization via a Viewpoint-Dependent Classifier Model

V. Tchuiev, Y. Feldman and V. Indelman

Technion - Israel Institute of Technology

vovatch@campus.technion.ac.il
yurif@cs.technion.ac.il
vadim.indelman@technion.ac.il

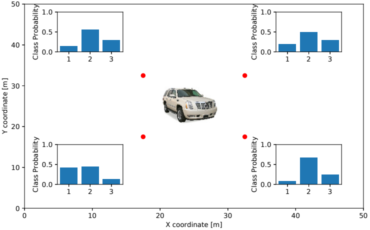
November 7, 2019

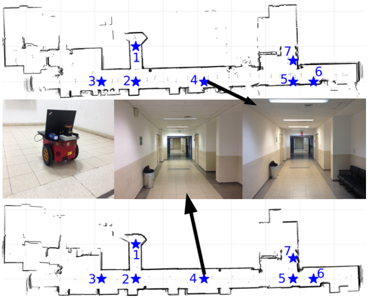
ANPL Autonomous Navigation and Perception Lab

TECHNION Israel Institute of Technology

Introduction

- We propose an **object based SLAM**approach.
- **Classification** is a key component for object based SLAM.
- Key challenge: operation in **perceptually aliased** environments.
    - Classification aliasing
    - Data Association (DA) aliasing





[1]

[1] Right image from: Pathak, Shashank, Antony Thomas, and Vadim Indelman. "A unified framework for data association aware robust belief space planning and perception." The International Journal of Robotics Research 37, no. 2-3 (2018): 287-315.

Related Work

Existing works:

- Consider *most likely class* semantic measurements.
  (Mu et al. 2016, Bowman et al. 2017)
- Utilize a viewpoint dependent classifier with DA solved:
  (Velez et al. 2012, Teacy et al. 2015, Feldman and Indelman 2018)
- Consider hypotheses for DA only:
  (Pathak et al. 2018)

Our work:

- Considers semantic measurements of class probability vectors.
- Uses a viewpoint dependent classifier model for DA-aware semantic SLAM.
- Considers joint DA and classification hypotheses.

Contribution

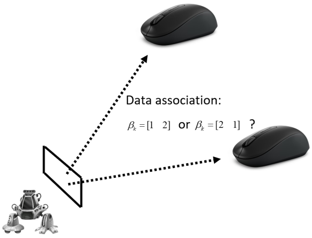
We utilize the *coupling* between **classifier outputs** and **relative viewpoint** between object and camera to:

1. Assist in **data association** (DA) disambiguation.
2. Improve **accuracy** and reduce **uncertainty** of pose inference for the robot and the objects in the scene.

- **We aim to maintain the hybrid belief:**

$$\mathbb{P}(\mathcal{X}_k, C, \beta_{1:k} | \mathcal{H}_k)$$

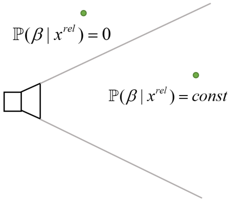
- $\mathcal{X}_k$: all robot and object poses.
- x_k: robot pose at time k.
- C: class hypothesis of all objects.
- β_k: data association realization.
- z_k^{geo}: geometric measurement of an object.
- z_k^{sem}: class probability vector of an object.
  - $Z_k^{geo} \doteq \{z_k^{geo}\}$, $Z_k^{sem} \doteq \{z_k^{sem}\}$: geometric and semantic measurements.
  - $\mathcal{H}_k \doteq \{Z_{1:k}^{geo}, Z_{1:k}^{sem}, a_{0:k-1}\}$: measurement history.



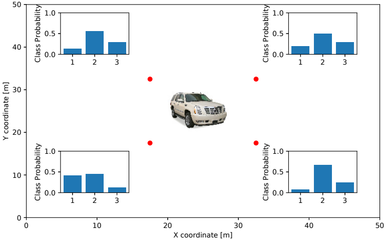
Assumptions

- A single robot within a static environment.
- A known number of objects.
- Models: motion $\mathbb{P}(x_k|x_{k-1}, a_{k-1})$, geometric $\mathbb{P}(Z_k^{geo}|\mathcal{X}_k, \beta_k)$, and classifier $\mathbb{P}(Z_k^{sem}|\mathcal{X}_k, C, \beta_k)$ are Gaussian.

- The object observation model $\mathbb{P}(\beta_k|x^{rel})$ determines if DA realization is feasible given relative pose.



- $z_k^{sem} \in \mathbb{R}^M$ is viewpoint dependent.



- The model is assumed Gaussian $\mathbb{P}(z_k^{sem}|c, x^{rel}) = \mathcal{N}(h_c, \Sigma_c)$, where h_c and Σ_c **depend on object class** c **and relative pose** x^{rel}.

- Split the hybrid belief to *continuous* and *discrete* parts:

$$\mathbb{P}(\mathcal{X}_k, C, \beta_{1:k}|\mathcal{H}_k) = \underbrace{\mathbb{P}(\mathcal{X}_k|C, \beta_{1:k}, \mathcal{H}_k)}_{b^C_{\beta_{1:k}}[\mathcal{X}_k]} \underbrace{\mathbb{P}(C, \beta_{1:k}|\mathcal{H}_k)}_{w^C_{\beta_{1:k}}}$$

- $b^C_{\beta_{1:k}}[\mathcal{X}_k]$ is the **continuous** belief given a class and DA realization.
- $w^C_{\beta_{1:k}}$ is the **weight** of $b^C_{\beta_{1:k}}[\mathcal{X}_k]$, computed separately for each C and $\beta_{1:k}$.
- We keep all continuous beliefs with large enough weights.

Continuous Belief and Weights

- **Continuous belief** update:

$$b^{C}_{\beta_{1:k}}[\mathcal{X}_k] \propto b^{C}_{\beta_{1:k-1}}[\mathcal{X}_{k-1}]\mathbb{P}(x_k|x_{k-1}, a_{k-1})\mathbb{P}(\mathcal{Z}_k|\mathcal{X}_k, C, \beta_k)$$

- **Weight** update:

$$w^{C}_{\beta_{1:k}} \propto w^{C}_{\beta_{1:k-1}} \int_{\mathcal{X}_k} \mathbb{P}(\beta_k|\mathcal{X}_k) b^{C}_{\beta_{1:k}}[\mathcal{X}_k] d\mathcal{X}_k$$

- **Small weights are pruned** to keep the number of realizations small.
- **Viewpoint dependent classifier model** in $\mathbb{P}(\mathcal{Z}_k|\mathcal{X}_k, C, \beta_k)$ assists in *inference* DA, and *reduces number of realizations* when pruned.

- **Comparison** between without and with classifier model.
- **Highly aliased** scenario with 6 identical objects with different orientations.
- **Uninformative prior** on initial robot pose, causing multiple probable hypotheses.

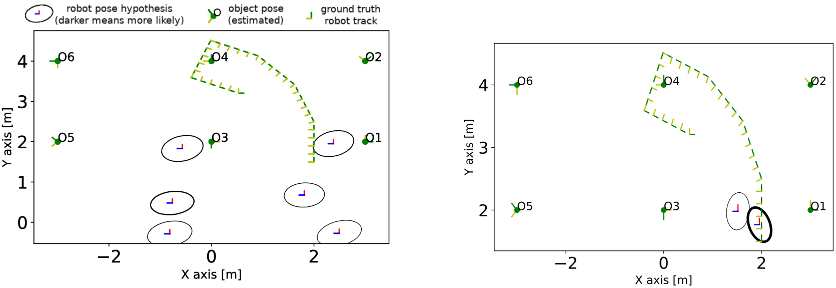


Figure: Time $k = 1$, without (left) and with (right) classifier model

Experiment

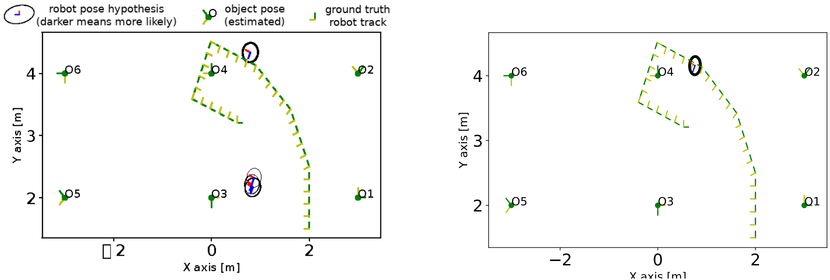


Figure: Time $k = 15$, without (left) and with (right) classifier model

Experiment

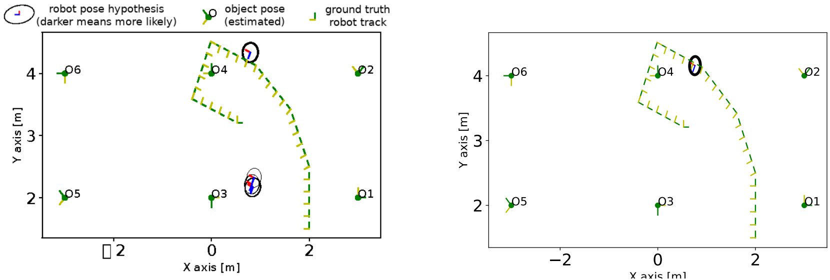


Figure: Time $k = 15$, without (left) and with (right) classifier model

$\Longrightarrow$ With classifier:

- fewer belief components
- More accurate localization

Experiment

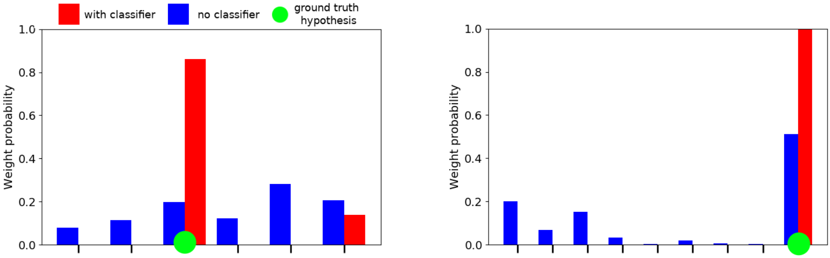


Figure: Weight comparison, times $k = 1$ (left) and $k = 15$ (right)

$\Longrightarrow$ With classifier:

- fewer belief components

Experiment

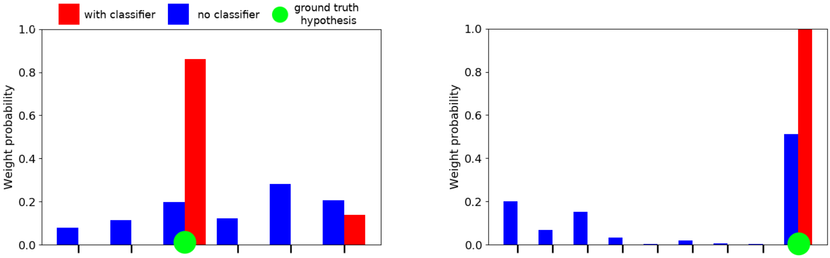


Figure: Weight comparison, times $k = 1$ (left) and $k = 15$ (right)

$\Longrightarrow$ With classifier:

- fewer belief components
- stronger disambiguation

Conclusion and Summary

- We propose an approach **addressing DA and classification ambiguity** that maintains a **hybrid belief.**
- We utilized the **coupling** between *object class* and *relative viewpoint* via a viewpoint dependent classifier model.
- Performance improvement in a highly aliased scenario was demonstrated for disambiguation and localization.

Thank you for listening!