

### 3 Understanding and controlling deviations from required behaviour

#### Practical guidance – automotive

Author: Dr Matthew Gadd, University of Oxford

#### Overview

We consider the required behaviour for autonomous vehicles to be motion which is:

1. Safe and
2. Efficient

and which is also

3. Understandable to occupant and standby alike as well as
4. Guaranteeable to all concerned parties.

We consider behaviour to be both learned from data as well as handcrafted as part of a complex system of rules. In the latter case, we consider that the rulebook being incomplete leads the AV into scenarios not accounted for. In the former case, a distributional shift from training data to the deployed system or incompleteness in the scenarios presented during training may cause deviation from the required behaviour.

More specifically, and in reference to both learned and non-learned systems, we find that the autonomy-enabling mobile robot tasks of motion estimation, localisation, planning, and object detection give rise to these deviations due to

1. Poorly modelled vehicle kinematics,
2. Sensor artefacts,
3. Environmental ambiguity, and
4. Scene distractors

Under each of these scenarios we find that it is important for vehicles to exhibit introspection - both explaining unavoidable deviations in behaviour (pin-pointing it in data), as well as avoiding deviations (if able to identify them ahead of time).

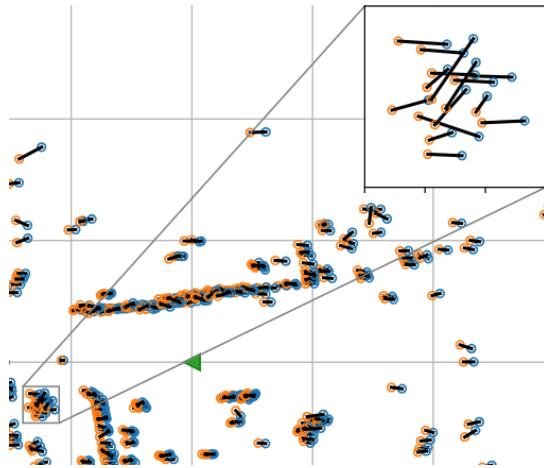
Finally, we also acknowledge that deviations are sometimes unavoidable, and also use them to enrich the learned models for better future behaviour.

The following present some of our findings in these areas.

#### Vehicle modelling

The fidelity of motion estimation is suspect to incorrect data association on a frame-by-frame basis. This is tied to environmental ambiguity and sensor artefacts (see below), but we find in [1] that incorporating a kinematic model of the vehicles motion into the estimation strategy itself is greatly beneficial to the robustness of this crucial task (see Fig. 1). Specifically, we use the strong prior of how a robot is constrained to move smoothly

through its environment allows for easy outlier detection of poor matches that are a dominant source of pose estimate error.



*Fig 1: Deviation in data association for motion estimation can be bolstered by considering realistic platform constraints - our vehicles cannot move sideways.*

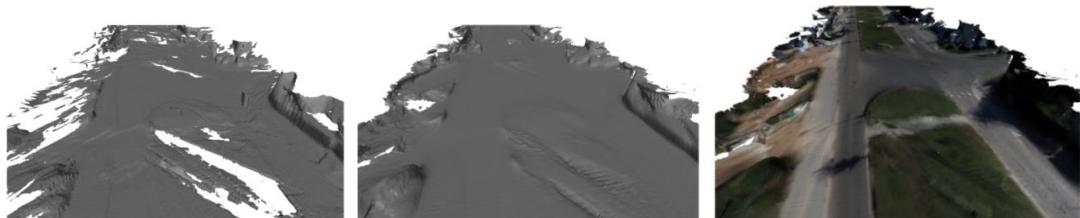
There is also potential for deviation in a robot's localisation capability due to environmental ambiguity - false positives abound in ambiguous scenes, where many recorded measurements might have some similarity to the current scan. Here, again, domain knowledge of vehicle kinematics can be a salve, where in [5] we show that searching for coherent sequences of matching scans from a radar is beneficial to the vehicle's ability to recognise previously visited places.

### Sensor artefacts

Here, sensor failure modes and other noises perturb the vehicle's measurements which are processed by autonomy-enabling systems in order to make predictions and inform behaviour.

This is well documented in the application of vision to mobile autonomy, where lens flare and lens adherents are especially problematic. In radar, the measurement process itself is very complex, leading to multipath reflections, doppler effects, speckle noise, and ringing. While largely not affecting radar, obstructed views particularly impede scene understanding with vision and lasers.

In [9] we learn how to repair artefacts in the construction of detailed 3d maps from stereo vision and laser (see Fig. 2). Here, we acknowledge that these artefacts may be unavoidable in a realistically deployed system - with sensors limited by cost and form factor, etc - but find that we can learn a repairing mechanism which can be deployed alongside the limited systems which can produce in real-time maps which are comparable to super-rich, high quality representations.

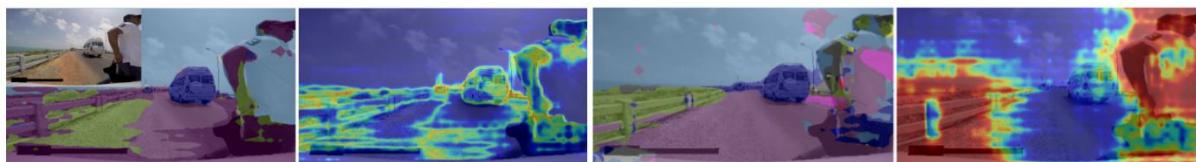


*Fig 2: Sometimes, for practical reasons, we deploy lower quality sensors. We can nevertheless deploy alongside them machine-learned models for correcting their deviations.*

In [3] [4] we exploit the “stochastic” nature of the radar measurement process to train an unsupervised radar place recognition system. Here, we take advantage of these noises and perturbations to expose our network to more realistic data than it otherwise would have been privy to with standard unsupervised representation learning approaches.

## Environmental ambiguity

Specific to the way in which the sensor technology works, different parts of the environment may have very similar “appearance” which can be confusing for downstream tasks, leading to errors in scene understanding. For example, the visual appearance of vegetation dominating the camera view may lead to deviation in localisation while a straight road without many intersections, perceived as two straight lines of powerful returns, is not that distinctive as measured by radar.



*Fig 4: Vision systems should understand what parts of the scene are causing deviation due to being “out-of-distribution” with respect to the way they were trained.*

We address the latter in [7], presenting a system for both estimating confidence in and correcting the solution of a motion estimation system. At run-time, we use only the sensor data itself to detect failures based on a weakly supervised classification procedure which is trained offline by data fusion. The belief in how likely a current failure is passed to a correction system that performs an update based on the reliability of the estimate.

On the other hand, the environment may be somewhat surprising or unknown. In this sort of scenario, lacking representation in the training data we of course do not expect our machines to understand the category of objects they are perceiving, but we do expect that the learned system will at least be able to know what parts of the perceived scene it does not understand (i.e. be introspective). An example is given in Fig. 4, taken from our work in [10] in which we develop this facility in scene segmentation with cameras.

## Scene distractors

Perturbation of the sensor stream and therefore deviation from ideal performance in autonomy enabling tasks and autonomous behaviour itself is also possible due to the presence of unaccounted for and uncontrollable objects and actors in the scene.

For example, in [6] we present a system for improving the performance of laser lateral localisation systems in the face of missing kerb-side data due to other active road participants as well as idle vehicles which obscure observation of the kerb itself (see Fig. 5).



*Fig 5: Stationary and dynamic scene distractors of kerb detection for lateral localisation in an urban environment.*

Often, scene distractors are heuristically associated with their effect. In contrast, in [8] we train a filter for distractors by directly supervising against the task at hand. This allows the robot to give implicit meaning to what a scene distractor actually is, rather than basing this off of a human interpretation. This is especially important for sensors such as radar, where human interpretability of the data is tenuous. We let the algorithm decide what is apposite and what is irrelevant.

In general, we advocate that it is important for an intelligent system to understand when its measurements are markedly different from the observations it was privy to when acquiring the behaviour that it is expected to exhibit. In [2], then, we present a system for out-of-distribution detection and selective segmentation of camera images which both knows how to avoid making predictions when faced with distracting scene content which it cannot explain as well as boosts scene understanding performance in canonical situations.

## Summary

In summary, we present four sources of deviation from required behaviour for mobile robots. These are in reference to the way in which our machines should understand their perceptions as well as shortcomings in their programming in order to move robustly and safely through a dynamic world. Throughout, we advocate for a combination of data-driven approaches which are nevertheless structured by strong human priors.

## References

- [1] R. Aldera, M. Gadd, D. De Martini, and P. Newman, "Leveraging a Constant-curvature Motion Constraint in Radar Odometry," in IEEE International Conference on Robotics and Automation (ICRA), (Philadelphia, PA, USA), September 2022.
- [2] D. Williams, M. Gadd, D. De Martini, and P. Newman. "Fool Me Once: Robust Selective Segmentation via Out-of-Distribution Detection with Contrastive Learning," in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA) 2021.
- [3] M. Gadd, D. De Martini, P. Newman, "Unsupervised Place Recognition with Deep Embedding Learning over Radar Videos", Workshop on Radar Perception for All-Weather Autonomy at the IEEE International Conference on Robotics and Automation (ICRA), 2021
- [4] M. Gadd, D. De Martini, and P. Newman, "Contrastive Learning for Unsupervised Radar Place Recognition," in IEEE International Conference on Advanced Robotics (ICAR), (Ljubljana, Slovenia), December 2021
- [5] M. Gadd, D. De Martini, and P. Newman, "Look Around You: Sequence-based Radar Place Recognition with Learned Rotational Invariance," in IEEE/ION Position, Location and Navigation Symposium (PLANS), Portland, OR, USA, 2020.
- [6] T. Suleymanov, M. Gadd, L. Kunze, and P. Newman, "LiDAR Lateral Localisation Despite Challenging Occlusion from Traffic," in IEEE/ION Position, Location and Navigation Symposium (PLANS), Portland, OR, USA, 2020.

[7] R. Aldera, D. De Martini, M. Gadd, and P. Newman, “What Could Go Wrong? Introspective Radar Odometry in Challenging Environments,” in IEEE Intelligent Transportation Systems (ITSC) Conference, Auckland, New Zealand, 2019.

[8] R. Weston, M. Gadd, D. De Martini, P. Newman, and I. Posner, “Fast-MbyM: Leveraging Translational Invariance of the Fourier Transform for Efficient and Accurate Radar Odometry,” Under Review

[9] S. Saftescu, M. Gadd, and P. Newman, “Look Here: Learning Geometrically Consistent Refinement of Inverse-Depth Images for 3D Reconstruction,” International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), 2021.

[10] D. Williams, M. Gadd, D. De Martini, and P. Newman, “Fool Me Once: Robust Selective Segmentation via Out-of-Distribution Detection with Contrastive Learning”, IEEE International Conference on Robotics and Automation (ICRA), 2021