

Non-parametric Learning To Aid Path Planning Over Slopes

Sisir Karumanchi Thomas Allen Tim Bailey
Steve Scheduling
ARC Centre of Excellence For Autonomous Systems (CAS),
Australian Centre For Field Robotics (ACFR),
The University of Sydney,
NSW. 2006, Australia.
Email: s.karumanchi/t.allen/t.bailey/s.scheduling@cas.edu.au

Abstract

This paper addresses the problem of closing the loop from perception to action selection for unmanned ground vehicles, with a focus on navigating slopes. A new non-parametric learning technique is presented to generate a mobility representation where maximum feasible speed is used as a criterion to classify the world. The inputs to the algorithm are terrain gradients derived from an elevation map and past observations of wheel slip. It is argued that such a representation can aid in path planning with improved selection of vehicle heading and velocity in off-road slopes. In addition, an information theoretic test is proposed to validate a chosen proprioceptive representation (such as slip) for mobility map generation. Results of mobility map generation and its benefits to path planning are shown.

1 Introduction

Learning techniques that close the loop from perception to action selection are of particular interest for off-road robotics. This loop closure refers to the need for an intermediate module that processes sensed exteroceptive¹ information such as terrain slopes and colour into a representation which can directly aid in decision making (such as path planning). This task of interpreting exteroceptive data by associating a scalar value of cost or utility is referred to as *Scene Interpretation* in this work. In ground vehicle robotics the focus is usually on identifying hard hazards such as obstacles or classifying predefined environmental states into different degrees of traversability (Ojeda et al. 2006; Jackel et al. 2006; Hadsell et al. 2007; Shneier et al. 2008). Assumptions such as terrain homogeneity, or perpetual existence of a road are often made to simplify the problem. In the absence of such assumptions,

theoretical techniques that use sensed information to aid decision making need to be investigated. This paper addresses the above mentioned problem with a specific focus on negotiating two dimensional slopes given range sensor measurements. However, the learning algorithm presented in this work can be extended to include additional exteroceptive variables such as terrain colour and texture.

Autonomous navigation in unstructured conditions such as non-homogeneous uneven terrain is a challenging problem to solve. In such environments two main issues need to be addressed. First, explicit assumptions about the terrain should be avoided. Second, in addition to hard hazards (such as obstacles), soft hazards (situations where behaviour needs to be adapted) need to be identified and dealt with. For example, terrain slopes are soft hazards and to successfully negotiate them, vehicle behaviour such as velocity, operating gear and vehicle heading needs to be adjusted. In addition to obstacle avoidance, cost representations that aid path planning over slopes require two key properties. i) Orientation sensitivity, as navigating down hill and up hill need to be judged differently ii) The ability to encapsulate platform and controller limitations, as performance on slopes is very sensitive to controller tuning in practice. Tuning a controller to a certain condition (such as flat terrain) can limit its performance in other conditions (such as non-flat terrain)(see Extension 1).

Due to recent developments in Bayesian non-parametric techniques, *learning from experience* architectures offer promise. In such architectures, no assumptions need to be made about the environment. The environment representation is only limited by the available sensor suite and the variables used to define the exteroceptive state. Hence, a non-parametric learning technique is proposed in Section 4 to learn from past observations of behaviour, environment and proprioceptive² feedback. More specifically, the proposed learning

¹Exteroception: perception of external factors that are not under agent control

²Proprioception: perception of internal factors that are affected by envi-

technique uses past observations of slope (environment), velocity (behaviour) and wheel slip (proprioceptive feedback). Better representations in the scene interpretation problem can aid the purpose of bridging the gap between perception and action selection in unstructured environments. One cost/utility representation of the world that is of interest is a mobility map (Wong 2001) (see Figure 1). Here the maximum feasible speed of the vehicle between two points is used as a criterion for continuous classification, as such capturing the net utility of an environment condition with a single value. Such a mobility map explicitly represents traversability of occupied, admissible and unexplored regions and can be used as an objective map for trajectory planning algorithms such as A^* .

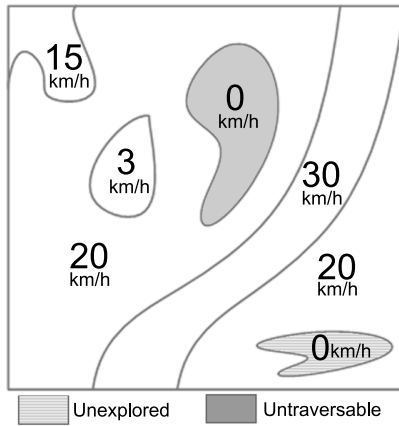


Figure 1: Sample Mobility Map Indicating Maximum Feasible Speed (Shaded areas indicate immobile and unexplored areas)

For navigating slopes, a vectorised representation of a mobility map can offer orientation sensitivity by making the mobility values dependent on the direction of pitch and roll slope definitions. An additional benefit of mobility representation over the traditional heuristic cost spaces is that environment utility is defined in behaviour space or action space³. In behaviour space, scene interpretation can be treated as a learning problem where agents learn about behavioural limits (such as maximum feasible speeds in a mobility map) by physically interacting with the environment. The observed extent of behaviour limitation provides information about environment utility/risk. For example, a lower speed limit in a mobility map indicates higher behaviour limitation due to its restricted ability to interact with the environment. As such, the extent of

environment and one's own behaviour.

³This type of knowledge representation of the world in behaviour space is an embodiment of the concept of 'enaction' which is a current trend in cognitive studies. It proposes a bidirectional link between perception and action such that perception controls action and further action influences perception in a closed loop system (Noë 2005).

limitation is indicative of how costly a particular environmental condition is relative to other conditions. Other examples of behaviour space measures for scene interpretation include kinetic energy limits, and bandwidth of allowable behaviours that indicate the extent of feasible interaction with an environmental condition. In behaviour space interpretations both hard and soft hazards are encapsulated in a continuum as different degrees of behavioural limits. Such an interpretation also captures both platform and controller limitations implicitly.

In summary, it is beneficial to clarify the focus of this paper. The proposed learning approach is a means for gross assessment of an environmental condition given past observations of proprioceptive feedback such as slips. This is the problem of getting a cost/utility/reward function to aid decision making but not a decision making problem in itself. From a decision making point of view, temporal correlations between successive actions and spatial correlations between neighbouring environmental states become important. These correlations are not taken into account in a scene interpretation problem where the goal is get an instantaneous assessment of a given condition. However, given a reward/utility function they can be taken into account in a reinforcement learning or an optimal control framework.

This paper is an extended version of Karumanchi et al. (2009). In this revised version, we extend the learning algorithm to address multimodality and scalability to large datasets. Additionally, an empirical validity test is introduced that addresses the question of an appropriate choice of proprioceptive representation. For example, in the specific problem addressed in this paper the validity test is used to justify the use of wheel slips to interpret terrain slopes.

2 Related Work

In this section, related publications that address scene interpretation as a learning problem are discussed. To our knowledge, the use of learning techniques for a continuous space interpretation⁴ is quite recent and only has a few related publications, which are discussed below.

Existing 'learning from experience' techniques include Reinforcement Learning (Sutton and Barto 1998) and model predictive techniques (Lacroix et al. 2002; Green and Rye 2007). Model predictive techniques aim to quantify the environment from repeated simulation. However, these techniques make assumptions such as the existence of a reward function in Reinforcement Learning, or the existence of accurate models for model predictive techniques. It is difficult to quantify such

⁴In contrast, discrete space interpretation of terrain conditions is more popular in the form of classification problems (Ojeda et al. 2006; Jackel et al. 2006; Hadsell et al. 2007; Shneier et al. 2008).

reward functions or develop accurate models in unstructured environments in a deterministic fashion. This work tackles the problem of scene interpretation which looks at learning a reward/utility function that can be used within reinforcement learning or optimal control frameworks.

Related work in off-road navigation includes work in long range planning using overhead data by Silver et al. (2006) and Sofman et al. (2006). In the former work, the authors use manually engineered traversal costs by hand tuning the cost functions from semantic and geometric data. Sofman et al. (2006) demonstrate a self-supervised approach of predicting terrain cost on far-range or overhead data from noisy local observations of cost. Although their cost definitions are based on heuristics, the contributions (weights) of different feature costs are learned in a probabilistic model. In this work, it is shown that the problem of negotiating soft hazards such as slopes can be better addressed with a mobility representation compared to scalar cost representations used in the above mentioned papers.

Current state of the art in scene interpretation includes Imitation Learning (Silver et al. 2008; Ratliff 2009) and Inverse Reinforcement Learning (Abbeel and Ng 2004). These are relatively new concepts and have been applied to the problem of learning a reward function from example behaviour. However, they rely heavily on expert input. Controller limitations are usually ignored when systems rely purely on expert input. Such limitations can be dealt with implicitly when the vehicle explores its behavioural capabilities on its own terms. This sort of introspective judgement in a self-supervised manner is achievable if one has access to proprioceptive feedback.

A *learning from proprioception* approach is demonstrated in Angelova et al. (2007) for a Mars-Rover platform where the authors represent the environment in proprioception space in terms of expected slip. This approach ignores the influence of velocity on wheel slip. For a Mars-Rover, proprioceptive measures (such as wheel slip) are mainly dependent on environment conditions and behavioral influences (such as velocity) could be ignored because the platform moves slowly. However, this assumption cannot be made for larger platforms where there is a distribution of slip values for a given condition pertaining to all possible behaviours. Also, such an approach is limited to the case when proprioception is a scalar or a weighted average of scalars. The latter usually involves manual tuning of weights which is not an intuitive process. A single scalar cost cannot capture multiple objectives in unstructured conditions. Instead it is beneficial to use a collage of proprioceptive stimuli to judge actions. Traditionally, proprioceptive feedback has been used as a means to instantaneously assess utility or risk (Stavens et al. 2007), much akin to an instantaneous reward formulation in Reinforcement Learning. However, for complex agent-environment interaction problems it is

not the immediate proprioceptive observation but the ensemble of past observations which are informative. In this work, proprioceptive information is dealt with in a statistical framework for the above mentioned reasons.

3 Outline

This paper is organised as follows. Section 4 formally introduces the scene interpretation problem as conditional density estimation and a non-parametric solution using Gaussian Processes is proposed. Section 5 introduces an information theoretic test to validate a chosen proprioceptive representation. Section 6 discusses the test platform and describes collection of training data. In Section 7, results of the validity test for different proprioceptive representations are shown. Additionally, results of scene interpretation are shown on an elevation map derived from laser scans. Mobility maps are derived from the given elevation map by analysing terrain gradients with past observations of vehicle slip collected from a skid-steered vehicle (shown in Figure 2). Section 8 compares path planning over mobility maps with planning over heuristic costs. Section 9 analyses various regressions algorithms for the problem at hand and discusses their limitations. Finally in Section 10, future work is discussed and concluding remarks are presented.



Figure 2: Argo 8x8 Unmanned Ground Vehicle

4 Proprioceptive Scene Interpretation

4.1 Motivation

Current terrain perception modules in unmanned ground vehicles (UGVs) are focused on creating an accurate internal representation of the environment. Exteroceptive parameters such as terrain colour and terrain slope have little value if the vehicle cannot associate them with a value of cost/utility of movement.

In the context of mobility maps a bound on velocities is desired from experience data. For this to occur, exteroceptive states (\tilde{e}) such as terrain slopes must be related with action states (\tilde{a}) such as vehicle velocity. Given no additional information, the two states are independent ($\tilde{e} \perp \tilde{a} | \emptyset$). However, when the appropriate proprioceptive feedback (\tilde{j}) is observed, the two become indirectly related ($\tilde{e} \not\perp \tilde{a} | \tilde{j}$). From the above assumption, the ‘*appropriate proprioceptive feedback*’ is defined as any random variable that is causally dependent on both environmental factors and action states. The causal dependencies are shown as a directed acyclic graph in Figure 3(a). The specific choice of variables for the problem at hand of interpreting terrain slopes in shown in Fig 3(b) for concreteness.

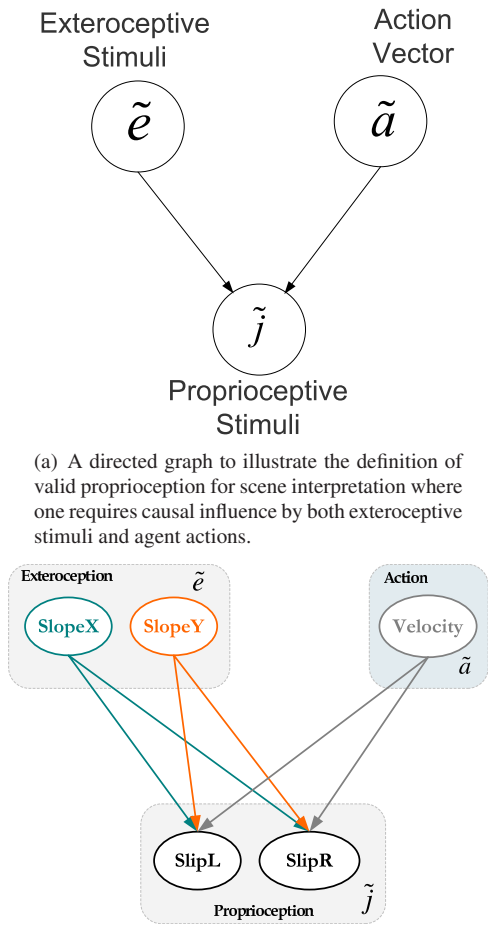


Figure 3: Definition of valid proprioception and the variables used for the specific problem investigated in this paper

Gathering experience corresponds to collecting co-occurrent observations of \tilde{e} , \tilde{a} and \tilde{j} in as many varied conditions as possible⁵. This experience set (\mathbf{H}) serves as a training set for learning. The exploration philosophy for collecting training data is to explore the natural feasibility of vehicle behaviour in as many varied conditions as possible either under manual or autonomous control. The latter has the advantage of exploring controller limitations.

For experience-based scene interpretation, practicality of collecting training data is critical. Exhaustive exploration of behaviour is not practical for large platforms because of safety concerns. Instead of treating the problem as an optimisation over observed proprioceptive stimuli to determine optimal behavioural limits (which requires worst case exploration), the problem can be relaxed to concentrate on suboptimal bounds that capture the extent of behavioural feasibility. In such a scenario, training data only needs to be collected within the thresholds of feasibility. As a result data collection is practical since the vehicle only explores what it can negotiate comfortably. Unexplorable behaviour contributes to scene interpretation by indirectly informing about behaviour limitation.

For the skid-steered vehicle of interest, slip estimates are chosen as proprioceptive feedback. The slip values cannot be measured directly, so they are estimated with an Unscented Kalman Filter (UKF) (Julier et al. 2000) using the two-track process model mentioned in Le et al. (1997) (shown in Appendix B). The reason for using a Kalman Filter is to efficiently deal with sensor noise. The test platform has an on-board Inertial Navigation System (INS) which is used to sense vehicle actions such as velocity with good accuracy. In addition, pitch and roll information from the INS are used to sense the current terrain slope (exteroceptive conditions)⁶.

In the following subsections, the notation is summarised in one place to provide easy reference to all the variables and then the theory is subsequently introduced.

4.2 Nomenclature

\tilde{x} - tilde is used to indicate that a particular variable is a vector.

\tilde{a} - Action vector (velocity)

\tilde{e} - Exteroceptive stimuli (terrain slopes)

\tilde{j} - Proprioceptive stimuli (wheel slips): A vector of measures that indicate dependence of performance on environmen-

⁵Existence of a stationary joint distribution $p(\tilde{e}, \tilde{j}, \tilde{a})$ is assumed. Therefore the experience/training set is a collection of i.i.d. samples from the joint.

⁶The test platform used in this work does not have any suspension, so pitch and roll information from the INS reflects the terrain slope accurately. For a different platform with suspension, it is more appropriate to derive terrain slopes from exteroceptive sensors.

tal conditions and vehicle behaviour.

H - Experience set (training set)

$$\begin{pmatrix} \tilde{e}_1 & \tilde{e}_2 & \cdots & \tilde{e}_N \\ \tilde{j}_1 & \tilde{j}_2 & \cdots & \tilde{j}_N \\ \tilde{a}_1 & \tilde{a}_2 & \cdots & \tilde{a}_N \end{pmatrix}$$

J* - Feasible Proprioception Set $-\{\tilde{j}_1^*, \tilde{j}_2^*, \dots, \tilde{j}_M^*\}$

This is the set of proprioceptive stimuli observed in ideal conditions such as flat terrain in the case of terrain slopes⁷. This set is used as an indication of feasibility. Actions that have a high frequency of proprioceptive response in the region of the **J*** set are considered to be feasible.

E_{test} - Set of test conditions which need to be interpreted (Test set) - $\{\tilde{e}_{test1}, \tilde{e}_{test2}, \dots, \tilde{e}_{testT}\}$.

for example, the set of both horizontal and vertical gradients for each grid cell form the set of test conditions to interpret terrain slopes from an elevation map.

4.3 Problem Definition

There are two main steps to the problem which will be referred to through out this paper.

Step 1: Identify a representation of \tilde{j} which satisfies the desired causal dependencies.

Step 2: Given a choice of \tilde{j} , perform proprioceptive scene interpretation (discussed below).

Step 1 is covered in Section 5. The following section covers Step 2 assuming that a correct representation of \tilde{j} has been identified and estimates or observations of those states are available.

Before velocity limits can be derived from experience data, an intermediate goal is to infer the *feasible behaviour distribution* for any test condition given the set of all past observations (**H**) and the comfortable proprioception set (**J***) which is chosen by the user to be observations in ideal/nominal conditions. For interpreting slopes, observations from flat terrain conditions are labelled as ideal and used as a reference. This process can be intuitively understood as training the robot what to look for (in proprioception) when exploring feasibility of actions in unknown conditions.

⁷The choice of what corresponds to ideal conditions is a design variable chosen by the user.

Once feasible behaviour distribution is inferred, an upper bound using the cumulative density function (CDF) can determine velocity limits for use in mobility maps. The process of deriving a mobility map given a set of test conditions is outlined in Algorithm 1.

Input: Elevation Map (A set of elevation values)

- Apply the Sobel operator (Gonzalez and Woods 2008) to determine gradient maps in pitch and roll directions (Test set $-\mathbf{E}_{test}$).

foreach \tilde{e}_{test} **in** \mathbf{E}_{test} **do**

 - Infer feasible behaviour distribution from past experience (**H** & **J***)
 - Determine behavioural limit (Maximum Feasible Speed)

end

Output: Mobility Map (Set of all associated mobility values ordered according to their respective test condition in \mathbf{E}_{test})

Algorithm 1: Scene Interpretation Process For An Elevation Map

Determining the feasible behaviour distribution is a conditional density estimation problem. The feasible behaviour distribution for a selected environment condition is $p(\tilde{a}|\tilde{e}_{test}, i = 1)$ ⁸ where i is an indicator variable to represent the feasibility constraint $\tilde{j} \in \mathbf{J}^*$.

$$i = \begin{cases} 1 & \tilde{j} \in \mathbf{J}^* \\ 0 & \tilde{j} \notin \mathbf{J}^* \end{cases} \quad (1)$$

$p(\tilde{a}|\tilde{e}_{test}, i = 1)$ ⁹ is a *measure of confidence* in taking an action \tilde{a} given past experience (**H**). Confidence for an action is based on how often proprioception observed under that action was within the set of proprioceptive stimuli observed in ideal conditions (**J***) i.e. actions that generated stimuli in the region of feasible proprioception **J*** are preferred. $p(\tilde{a}|\tilde{e}_{test}, i = 1)$ can be comprehended as the probability of any action satisfying the constraint of $\tilde{j} \in \mathbf{J}^*$ in any given environmental condition.

⁸Equivalent to $p(\tilde{a}|\tilde{e}_{test}, i = 1, \mathbf{H})$ - dependence on the training set **H** is not shown for conciseness.

⁹It is preferable to model in action space ($p(\tilde{a}|\tilde{e}_{test}, i = 1)$) instead of proprioception space ($p(i = 1|\tilde{e}_{test}, \tilde{a})$) as in practice the training data samples are overwhelmingly populated in the region of feasibility ($i = 1$) and any attempt to model $p(i = 0|\tilde{e}_{test}, \tilde{a})$ would be inaccurate which is required to probabilistically model in the i space. It is also undesirable to explore in infeasible regions for safety concerns on large platforms. This relates to the argument about worst case exploration made in the previous section.

4.4 A Non-Parametric Approximation

In this section a hierarchical non-parametric¹⁰ approach is presented to approximate the global conditional density $p(\tilde{a}|\tilde{e}_{test}, i = 1)$. Initially, an inner module approximates the function $\tilde{a} = f(\tilde{e}, \tilde{j})$ within a Bayesian non-parametric framework using Gaussian Processes. While the regression module infers local conditional distributions, the global conditional distribution is treated as a kernel density estimation problem where the number of kernels grow as the number of elements in the \mathbf{J}^* set grows. Together, the density $p(\tilde{a}|\tilde{e}_{test}, i = 1)$ can be adapted online as the sets \mathbf{J}^* and \mathbf{H} grow.

The whole process is captured in the following equation where the desired distribution is derived by marginalising $p(\tilde{a}, \tilde{j}|\tilde{e}, i)$ over \tilde{j} .

$$p(\tilde{a}|\tilde{e}, i = 1) = \int p(\tilde{a}, \tilde{j}|\tilde{e}, i = 1)d\tilde{j} \quad (2)$$

$$= \int p(\tilde{a}|\tilde{e}, \tilde{j})p(\tilde{j}|i = 1)d\tilde{j} \quad (3)$$

$p(\tilde{j}|i = 1)$ corresponds to the distribution of desired proprioception. In this application, since one has access to M samples of \tilde{j} in the \mathbf{J}^* set (i.i.d. samples from $p(\tilde{j}|i = 1)$), the above equation can be approximated as a weighted sum of conditional distributions at the observed \tilde{j} locations.

$$p(\tilde{a}|\tilde{e}, i = 1) \approx \sum_{\tilde{j}_i \in \mathbf{J}^*} \pi_{\tilde{j}_i} p(\tilde{a}|\tilde{e}, \tilde{j} = \tilde{j}_i) \quad (4)$$

where $\pi_{\tilde{j}_i}$ are mixing components; $\sum \pi_{\tilde{j}_i} = 1$

If all samples in the \mathbf{J}^* set are given equal importance.

$$p(\tilde{a}|\tilde{e}, i = 1) \approx \frac{1}{M} \sum_{\tilde{j}_i \in \mathbf{J}^*} p(\tilde{a}|\tilde{e}, \tilde{j} = \tilde{j}_i) \quad (5)$$

Equation 5 is in the form of kernel density estimation, but with variable kernels, as $p(\tilde{a}|\tilde{e} = \tilde{e}_{test}, \tilde{j} = \tilde{j}_i)$ is inferred from data. This can be viewed as an infinite mixture of conditional densities as the number of components grows when the \mathbf{J}^* set is allowed to grow. If the local conditionals are approximated to be Gaussian then the global approximation turns out to be a Gaussian mixture.

¹⁰Non-parametric techniques are preferred for learning from experience (memory-based learning) problems as they make the least assumptions about the global form of the distribution.

4.4.1 Gaussian Process Regression

Inferring the local conditional distribution $p(\tilde{a}|\tilde{e}, \tilde{j} = \tilde{j}_i)$ from observed data at each \tilde{j}_i location can be treated as a Bayesian Regression problem ($f: \{\tilde{e}, \tilde{j}\} \rightarrow a$) (Rasmussen and Williams 2006; Bishop 2006), where \tilde{e}, \tilde{j} are augmented together to form the input vector \tilde{x} and a is the output y .

$$a_i = f(\tilde{e}_i, \tilde{j}_i) + \varepsilon \quad (6)$$

where

$$\varepsilon - \text{Noise} \sim N(0, \beta^{-1})$$

$$\beta - \text{Noise precision}$$

A Gaussian Process (GP) is completely specified by its covariance function $K(x, x')$ and its choice defines the space of functions (latent variables - f) that can be generated (Rasmussen and Williams 2006). Further, the output is assumed to be zero mean. This is reflected in the prior over the latent variables $p(f)$. Because of this zero mean assumption in GPs, predictions are biased towards null behaviour region (zero) if no data is observed in the test conditions. In the scene interpretation problem, this translates to being cautious in unexplored or underexplored environments which is desired.

$$P(f) = N(0, K) - \text{Prior On Functions} \quad (7)$$

where

f – latent variables

$K(x, x')$ – covariance function

$x = \{\tilde{e}, \tilde{j}\}$ – input values

The predictive distribution is a Gaussian with the following form:

$$P(a|\tilde{e}_{test}, \tilde{j}_{test}, \mathbf{H}) = N(\mu(x), \text{var}(x)) \quad (8)$$

where

$$\mu(x) = K(x_{test}, X)[K(X, X) + \beta^{-1}I_N]^{-1}\mathbf{y}$$

$$\text{var}(x) = K_{test, test} + \beta^{-1} - K_{test}^T[K + \beta^{-1}I_N]^{-1}K_{test}$$

$\mathbf{H} - \{a_{train}, \tilde{e}_{train}, \tilde{j}_{train}\}_{1..N}$ – Training data

$\mathbf{y} - \{a_{train}\}_{1..N}$ – Training outputs

$X - \{\tilde{e}_{train}, \tilde{j}_{train}\}_{1..N}$ – Training inputs

$x_{test} - \{\tilde{e}_{test}, \tilde{j}_{test}\}$ – Test input

$$K_{test} = K(X, x_{test})$$

$$K_{test, test} = K(x_{test}, x_{test})$$

For the scene interpretation problem, the commonly used squared exponential covariance function is chosen. This

choice has the stationarity property of associating observations within a local neighbourhood which is desired.

$$K(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2l^2}(x - x')^2\right) \quad (9)$$

The hyper-parameters of the GP $\theta = \{l, \sigma_f, \beta^{-1}\}$ are learned by maximising the log likelihood of the training data (\mathbf{H}) using a numerical optimisation technique.

GP regression is a discriminative approach, additional exteroceptive or proprioceptive states could be augmented into the input vector. This allows for incorporating additional sensors or proprioceptive measures into the scene interpretation process. However, the main limitation is the inversion of an $N \times N$ matrix ($[K(X, X) + \beta^{-1}I_N]^{-1}$) which is an $O(N^3)$ operation (where N is the size of the dataset). In the next section, extensions to GP regression are discussed to address multimodality and scalability to large datasets.

4.4.2 Gaussian Process Extensions To Address Multimodality and Sparsity

The mapping from $\{\tilde{e}, \tilde{j}\}$ to \tilde{a} can be multimodal as it is essentially an inverse modelling problem. There can be multiple actions that give rise to the same proprioceptive response.

In GP literature, multimodality has been addressed as infinite mixture of Gaussian Process experts (iMoGPe) in Rasmussen and Ghahramani (2002) and Meeds and Osindero (2006). The essential idea is to probabilistically cluster the data points in the joint space of inputs and outputs before stochastically assigning an individual GP model to each cluster. In the version of Meeds and Osindero (2006) an infinite Gaussian mixture model (Rasmussen 2000) is used as a gating function to stochastically weight the contribution of each expert to any given input point for prediction. An appealing property of this model is that the number of clusters is inferred from the data with a Dirichlet process prior. Inference is performed with MCMC techniques such as Gibbs sampling for indicator variables and model parameters in conjunction with Hybrid Monte Carlo sampling of the GP hyperparameters. Example clustering results of sample multimodal datasets in 1D and 2D are shown in Figures 4(a) and 11(b).

Since the main aim of this paper is to infer an upper bound on velocities, it is sufficient to capture multimodality as noise in the outputs. The variance estimate of the original GP model (Equation 8) is a function of inputs only, the variance of the outputs is assumed to be homogeneous and is estimated as a hyperparameter. Heteroscedastic¹¹ Gaussian Process models

¹¹Heteroscedasticity is term used in statistic and GP literature to indicate non-uniform noise levels. In contrast, homoscedasticity implies the uniform noise assumption.

(Kersting et al. 2007; Goldberg et al. 1998) extend the basic GP model to deal with input dependent noise. They use a second GP to model the log variance of the outputs from the mean to predict local noise for any given test input. The dual GP models of target and noise processes are optimised iteratively in an Expectation-Maximisation (EM) like fashion. This captures the variance in the outputs at any location as input dependent noise. An example result on a 1D dataset with homoscedastic and heteroscedastic models is shown in Figure 4(b). For this dataset it can be seen that the homoscedastic noise assumption is undesirable as it causes the noise to be inflated in some regions. This is corrected with the heteroscedastic model, which accounts for the locality in the noise estimates.

The other limitation of the standard GP model is the cubic computational complexity of matrix inversion. Sparse approximations in Gaussian Processes is an active research area (Quiñero Candela and Rasmussen 2005). This section only presents some approximations that scale well to the heteroscedastic problem on large datasets. An obvious approach to address cubic complexity is to perform matrix inversion on a smaller subset of the original data. A better approach is to use the Projected Process (PP) approximation (Rasmussen and Williams 2006) which in addition to retaining a subset ($m < n$), projects the non-included data points into the basis space spanned by the included points. This is achieved through a likelihood approximation ($q(y|f_m) \approx p(y|f)$) derived from the class of likelihood distributions that only depend on the latent function values of the included subset (f_m). This approximation can be derived through minimisation of the KL-divergence between the approximated posterior and the true posterior ($KL(q(f|y)||p(f|y))$) (Csató 2002). The key aspects of the posterior approximation are shown below in Equation 10. The key factor in the approximation is the expectation of latent values for the non-included subset ($E(f_{n-m}|f_m)$) given the approximated posterior over the included latent function values ($q(f_m|y)$).

$$p(f/y) \propto p(f)p(y|f) \approx q(f|y) \propto p(f_m)q(y|f_m) \quad (10)$$

where

$$p(y|f) = N(f, \beta^{-1}I)$$

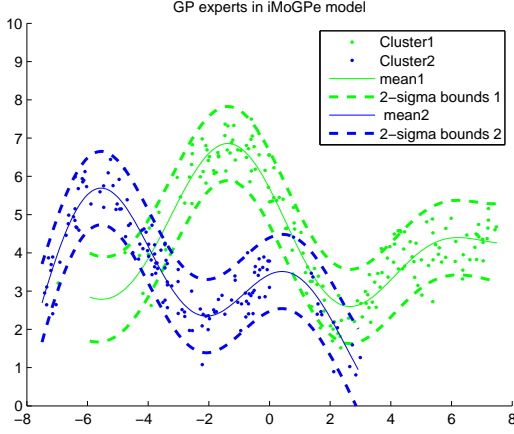
$$q(y|f_m) = N\left(\begin{bmatrix} f_m \\ E(f_{n-m}|f_m) \end{bmatrix}, \beta^{-1}I\right)$$

n – dataset size

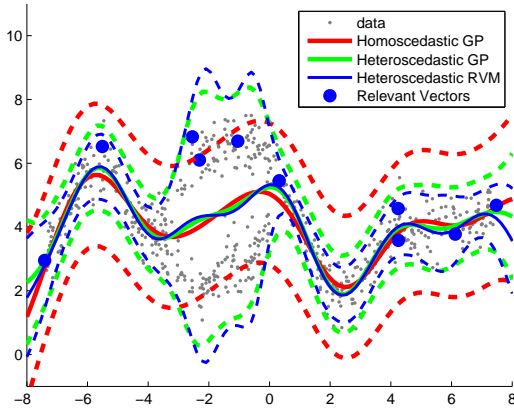
m – subset size

β – Noise precision

The PP approximation can be extended to the heteroscedastic model of Kersting et al. (2007) and its prediction equations are given below. An efficient Cholesky form implementation of the PP approximation is presented in Seeger et al. (2003).



(a) Infinite Mixture of Gaussian Process Experts: Addressing Multimodality



(b) Heteroscedasticity: Addressing Input Dependent Noise (dashed line: 2-sigma bound; solid line: mean estimate)

Figure 4: GP Extensions on a 1D example to addressing multimodality and input dependent noise.

$$\begin{aligned}\mu(x) &= k_m(x_{test})^T (K_{mm} + K_{mn}R^{-1}K_{nm})^{-1}K_{mn}R^{-1}\mathbf{y} \\ \text{var}(x) &= K_{test,test} + R^* - k_m(x_{test})^T K_{mm}^{-1}k_m(x_{test}) + \dots \\ & k_m(x_{test})^T (K_{mm} + K_{mn}R^{-1}K_{nm})^{-1}k_m(x_{test})\end{aligned}$$

where

$$K_{mm} = K(X_{1:m}, X_{1:m})$$

$$K_{mn} = K(X_{1:m}, X_{1:n})$$

$$k_m(x_{test}) = K(X_{1:m}, x_{test})$$

$$K_{test,test} = K(x_{test}, x_{test})$$

R/R^* – training/test noise matrices (Heteroscedasticity)

$$R = R^* = \beta^{-1}I \text{ (Homoscedasticity)}$$

It is quite common to choose the points in the subset via a greedy selection based on heuristics that are indicative of the novelty of a given input location (Csató 2002). However, for multimodal distributions it is important to consider the joint space of inputs and outputs instead of just input space novelty as it is important to preserve the relative proportions of the output values. For the problem at hand, it is better to use a random selection instead of a greedy selection of the subset.

A related GP approximation technique that achieves the same prediction equation of the Projected Process mean ($\mu(x)$) but a different variance form ($\text{var}(x)$) is known as the Subset of Regressors (SR) approximation (Rasmussen and Williams 2006). This approximation can be understood as a generalised linear regression that models the mean process ($\mu(x)$) as a weighted average of basis vectors with a Gaussian prior on the weights (shown below).

$$\begin{aligned}\mu(x) &= \sum_{i=1}^m \alpha_i K(x_{test}, x_i) \\ \alpha_m &\sim N(0, K_{mm}^{-1})\end{aligned}$$

The Relevance Vector Machine (RVM) model of Tipping (2001) can be derived from the SR approximation by assuming a diagonal prior on the weights ($\alpha_m \sim N(0, \text{diag}(K_{mm})^{-1})$). Therefore, the RVM model can be thought of as a special case of the Gaussian Process model with a degenerate covariance function (Rasmussen and Williams 2006). The significant advantage of the RVM model is that the likelihood factorises into an incremental form which leads to an accelerated training algorithm presented in Tipping and Faul (2003). The model is grown sequentially by a greedy selection of points that maximise the likelihood. Unlike the PP model, greedy selection is not based on heuristics, instead the true likelihood is directly maximised with sequential updates. In practice, the RVM achieves excellent sparsity of orders of magnitude ($m \ll n$) compared to other techniques. However, RVM training is only fast for fixed covariance hyperparameters. When hyperparameter updates are required the training speed suffers because of a need to recalculate the BASIS (K) matrix at each parameter update. However, training complexity still scales quadratically as opposed to the cubic complexity of the standard GP. Regardless of training complexity, the prediction is very fast because of the strong sparsity achieved. The heteroscedastic model of Kersting et al. (2007) can be performed with RVM models instead of the GP models. Example results of prediction and the chosen relevant vectors from the RVM model are shown in Figure 4(b).

The computation complexity of the different techniques presented in this section is outlined in Table 1. For the scene interpretation problem it is acceptable to have slow training time as this can be done offline. However, predictions need to be fast as they need to be performed online. Inference in the

Table 1: Computational Complexity Of Different Extensions (it : iterations; n : data size; m : subset size; k : number of clusters)

	Homoscedastic GP	Heteroscedastic GP - PP	Heteroscedastic RVM	iMoGPe
<i>Training</i>	$O(n^3)$	$O(it \times 2 \times n \times m^2)$	$O(it \times 2 \times m^3)$	$\approx O(it \times (n^3/k))$
<i>Prediction</i>	$O(n^2)$	$O(n \times m^2)$	$O(m^2)$	$\approx O(k \times (n/k)^2)$
<i>Storage</i>	$O(n^2)$	$O(n \times m)$	$O(n \times m)$	$\approx O(k \times (n/k)^2)$

iMoGPe model is slow and does not have a good termination criterion. The clustering results and convergence are sensitive to the quality of the initial guess for the GP hyperparameters. In practice, the MCMC sampling is performed for a sufficiently long number of iterations to ensure convergence. The cubic complexity of the matrix inversion is relaxed by the clustering process. Even for small number of clusters ($k \approx 10$) this results in a notable speed up. However, the number of iterations required for convergence is still very high (around 1000 iterations in our experiments) to scale to large datasets. In comparison, the number of iterations for the heteroscedastic models was small (less than 10 in our experiments). For the above reasons, the heteroscedastic architecture is sufficient for the scene interpretation problem to determine an upper bound. However it does not offer an alternative to the iMoGPe model if the goal is to sample from the conditional distribution.

A common criticism of the RVM and SR models is that the degenerate covariance function leads to variance and mean predictions at unseen locations collapsing to zero. In contrast, variance predictions in unknown locations expand for the standard GP model. The interpretation of uncertainty estimates is application dependent. For the scene interpretation problem, the former is more desirable as the zero bias translates towards having a preference for null behaviour in unseen locations. Also in practice, since training is only performed on available data, estimates of hyperparameters may cause extrapolation artifacts near the dataset boundaries regardless of the technique. For the above reasons, the proprioceptive scene interpretation algorithm is only performed on input locations presented in the training data. The end result of upper bound values and training locations are treated as a new training dataset for a secondary RVM model. The advantage of this approach is two fold. First, the zero mean bias of the RVM model in unseen locations is exploited. Second, the resulting RVM model is extremely sparse which makes it suitable for fast online predictions.

In summary, for the problem at hand of generating mobility maps the heteroscedastic GP model with the projected process approximation is used for training and then a second sparse RVM model is trained to be used for fast online predictions. This choice of algorithms is evaluated in Section 9 and its shown to be the best tradeoff in terms of modelling accuracy

and training complexity.

5 Information Theoretic Proprioception Validity Test

Validity of a proprioceptive representation (\tilde{j}) is application dependent. In practice, the initial choice of representation is based on engineering insight into the physics of agent-environment interactions. However, given the statistical nature of agent-environment interactions and a data driven approach to proprioception it is important to ascertain if sufficient information is available in the training data for the task at hand. This section proposes an information theoretic test to address the above by looking at dependencies observed in the experience data. More specifically for the problem addressed in this paper, this test serves to justify the use of behavioural limits derived from a given choice of proprioceptive measures for scene interpretation.

The need for proprioceptive feedback was previously introduced in Section 4.1 as a means to relate vehicle behaviours with environmental conditions. Having access to valid proprioception provides a mechanism to analyse behaviour on different environmental conditions. For this reason, any random variable that is causally dependent on both environment and vehicle behaviour (shown in Figure 3(a)) is of initial interest.

The aim of a proprioceptive scene interpretation module is to capture the relative difficulty between different environmental conditions. Difficulty of a condition is judged on the effort required in adapting behaviour to negate environmental disturbances. To validate utility of a proprioceptive representation (\tilde{j}) for scene interpretation, one needs to show that adapting behaviour minimises environmental effects. Given experience data, the sensitivity of \tilde{j} to environmental factors (\tilde{e}) can be plotted with respect to actions (\tilde{a}) using information theoretic measures of statistical dependence. This trend can serve as a visual tool to demonstrate that adapting behaviour regulates environmental influence on proprioception.

More specifically the argument to validate a chosen proprioceptive representation is posed as follows. First, a proprioceptive space is chosen according to domain knowledge such as wheel slip or traction coefficients. Second, the dependence trend between exteroceptive and proprioceptive measurements

is plotted against actions (velocity) using information theoretic measures. If a monotonic trend is observed it indicates that a bound on velocity¹² also corresponds to a bound on the environment influence on proprioception. The above discussion is summarised in the following proposition.

Proposition: Behaviour limits can be used to regulate the influence of environmental effects on proprioception if and only if a monotonic relationship is observed between $dependence(\tilde{e}, \tilde{j} | \tilde{a} = \tilde{a}_i)$ and \tilde{a} (see Figure 5).

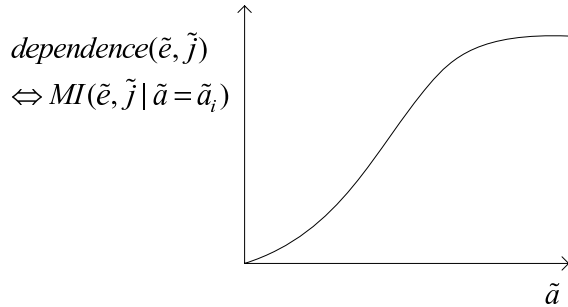


Figure 5: Validity requires a monotonic relationship between environment-proprioreception ($\tilde{e} - \tilde{j}$) dependence and the action variable (\tilde{a}). $MI(\cdot, \cdot)$ is mutual information which is used as a measure of dependence.

For the problem addressed in this paper, a monotonic trend implies that for low velocities the proprioceptive response is relatively similar in different conditions. However, at higher velocities the response varies widely on different conditions indicating a high dependence on environment. Thus by limiting velocities with an upper bound the volatility in proprioceptive response due to environmental effects is also bounded.

If the ($\tilde{e} - \tilde{j}$) dependence trace is flat, it indicates that behavioural effects are negligible. Therefore the link between actions and proprioception in Figure 3(a) could be ignored (as in Angelova et al. (2006, 2007)). If the ($\tilde{e} - \tilde{j}$) dependence trace is arbitrary, then there is little to be gained with velocity limits or a mobility based interpretation strategy either due to a bad choice of proprioceptive representation or just insufficient data collection in the given experience (training) data.

A common metric used to evaluate dependencies between two random variables is mutual information. Similarly conditional mutual information is used as a measure of conditional dependence. In the above test, specific conditional mutual information $[MI(\tilde{e}, \tilde{j} | \tilde{a} = \tilde{a}_i)]$ ¹³ is used as a test of dependence between \tilde{e} and \tilde{j} for a given $\tilde{a} = \tilde{a}_i$. In this paper, a normalised

¹²Kinetic energy bounds could be used instead of velocities for higher dimensional actions.

¹³calculation of $MI(\tilde{e}, \tilde{j} | \tilde{a} = \tilde{a}_i)$ requires the joint $p(\tilde{e}, \tilde{j} | \tilde{a} = \tilde{a}_i)$. Given the experience data it can be approximated with a discrete representation such as a normalised histogram.

variant of the mutual information is used to compare between alternative proprioceptive representations. This normalisation is achieved by dividing mutual information by the log of the joint space cardinality: $NMI(X, Y) = MI(X, Y) / \log(|X| * |Y|)$.

6 Test Platform, Testing Environment and Data Collection

The testing platform is a skid-steered vehicle (Figure 2). The platform is equipped with sensors to measure wheel speed, engine RPM, gearbox RPM and brake pressures. It also has an onboard Inertial Navigation System (INS) with access to raw accelerometer and gyro readings from the onboard IMU. The testing environment has access to DGPS (Differential GPS) corrections for the navigation module. The INS system along with GPS/DGPS observations delivers very good localisation (5cm accuracy) and vehicle actions such as velocity are available with good accuracy. Pitch and roll information from the INS are used to sense terrain slope (exteroceptive conditions) so that an elevation map can be interpreted from terrain gradients.

Training data was collected while executing 30 second exploration maneuvers in various terrain conditions. The exploration maneuvers included an acceleration phase, a coasting phase, a turning phase (both left and right turning) and a braking phase to ensure sufficient proprioceptive excitation. The different terrain conditions include flat terrain, uphill, downhill, positive and negative side slope conditions on grass, and a few runs over flat tarmac and a flat gravel road. The exploration runs were repeated for three distinct behaviours (*slow*: $< 1\text{m/s}$, *normal*: $1 - 2\text{m/s}$ and *fast*: $2 - 3\text{m/s}$) on each of the terrain conditions, so as to achieve sufficient exploration in behaviour space. In total, 20 minutes of data was collected at 20Hz.

7 Scene Interpretation Results

As mentioned in Section 4.3, there are two steps to the problem. Before the scene interpretation process can be implemented, it is important to select a proprioceptive representation that is conducive to the proposed scene interpretation strategy. This can be achieved with the help of the validity test mentioned in Section 5.

7.1 Step 1: Choice of Proprioceptive Representation

Results of the validity test are shown in Figure 6 for three different proprioceptive representations shown below.

- $\tilde{J}_{SlipsOnly} : \{Slip_L, Slip_R\}$
- $\tilde{J}_{\mu'sOnly} : \{\mu_L, \mu_R\}$ – Traction Coefficients
- $\tilde{J}_{All} : \{Slip_L, Slip_R, \mu_L, \mu_R\}$

The process models for estimating slip and traction estimates in an UKF framework are given in Appendix B and C.

To compare the traces of the validity test in terms of monotonicity, the slope of a linear functional fit on the bin values is used (see Table 2), where the higher the slope the better the monotonicity. The linear fit is shown in red in Figure 6. The mutual information measure is normalised by the log of the cardinality of the joint space to compare across different proprioceptive representations.

From Figure 6 and Table 2, the slips-only representation is seen to be the best performer in terms of monotonicity. On the other hand, the traction coefficients are influenced more by the environmental conditions (slopes) than behaviour as the trace is more or less flat¹⁴. This is expected, as the traction coefficients are mainly affected by the gravitational components acting on the vehicle when it is on slopes. This is also due to the fact that the vehicle was operated at low velocities in low gear for safety reasons ($< 3m/s$ or $10km/h$). Based on the monotonicity scores the slips-only proprioceptive representation $\{Slip_L, Slip_R\}$ is chosen (see Figure 3(b)).

Table 2: Validity Test Monotonicity Scores For Different Proprioceptive Representations

	$\tilde{J}_{SlipsOnly}$	$\tilde{J}_{\mu'sOnly}$	\tilde{J}_{All}
Fine Bin Size	0.0225	0.0132	0.0185
Coarse Bin Size	0.0203	0.0087	0.0161

7.2 Step 2: Mobility Map Generation

After a valid representation is chosen, the scene interpretation process can be applied. Given training data, and the set of test conditions, the extent of observed movement limitation for each of the test condition needs to be derived. In this section, laser data collected over 100x100m off-road terrain is used to derive an elevation map shown in Figure 7 (top). The Sobel operator (Gonzalez and Woods 2008) was applied to the elevation map image to derive pitch and roll gradients, that together form the set of test conditions (\mathbf{E}_{test}). Each grid cell has its corresponding exteroceptive state ($\tilde{e}_{test} : \{slope_{Pitch}, slope_{Roll}\}$) value which needs to be associated with a corresponding velocity limit.

¹⁴Note that the y-axis in Figure 6 is a dependency measure between \tilde{e} and \tilde{j} conditioned on actions. A flat non-zero trace informs that there is a constant dependence irrespective of actions.

$$\mathbf{E}_{test} : \begin{pmatrix} slope_{Pitch1} & slope_{Pitch2} & \cdots & slope_{PitchT} \\ slope_{Roll1} & slope_{Roll2} & \cdots & slope_{RollT} \end{pmatrix}$$

The slip estimate is two dimensional as observations from the UKF using the two-track process model (see Appendix B) consist of slip observations for both left and right tracks ($\tilde{j} : \{slip_{Left}, slip_{Right}\}$).

The set of all slip observations obtained from flat terrain conditions are chosen to be the nominal proprioception set \mathbf{J}^* . Given the desired proprioception set and the experience data from the training runs, a Gaussian Process with a squared exponential covariance function was optimised and the proprioceptive scene interpretation process described in Section 4.3 was implemented for the set of slope queries (\mathbf{E}_{test}) derived from the gradient maps (see Figure 7).

Each slope condition query results in a Gaussian mixture (see Equation 5). By selecting an upper bound on each conditional distribution the maximum feasible speed is determined. The upper bound can be determined from the cumulative density function. Also, a caching data structure is used to prevent interpretation of the same condition twice. This significantly improves the speed of the interpretation process.

The end result of such queries on an elevation map is a mobility map shown in Figure 7. The mobility map interprets the obstacles in the scene (trees) as untraversable with a velocity limit of zero, and the rest of the traversable regions on a continuous scale between 0-7kmph. This is visualised with pixel intensity, the brighter the pixel intensity, the easier it is to traverse.

Mobility is defined in vehicle frame, the direction of movement affects pitch and roll slopes which in turn affects mobility values. In this paper, A^* path planning is performed on a grid based representation, hence eight possible directions for slope are considered. Mobility for a given grid cell is a vector of values pertaining to eight possible orientations. Only one such mobility map is shown in Figure 7. All eight mobility maps are shown in Figure 8. Particularly of interest are maps in Subfigures 8(e) and 8(a), the values for going downhill (8(e)) are significantly smaller than going uphill (8(a)), indicating the need for increased caution.

8 Path Planning For Slopes

In Figure 9, path planning over a vector of mobility maps is compared with planning over a scalar cost map. This scalar cost is the maximum gradient of all eight orientations, and the corresponding ‘traversability’ map is shown in Figure 8(i).

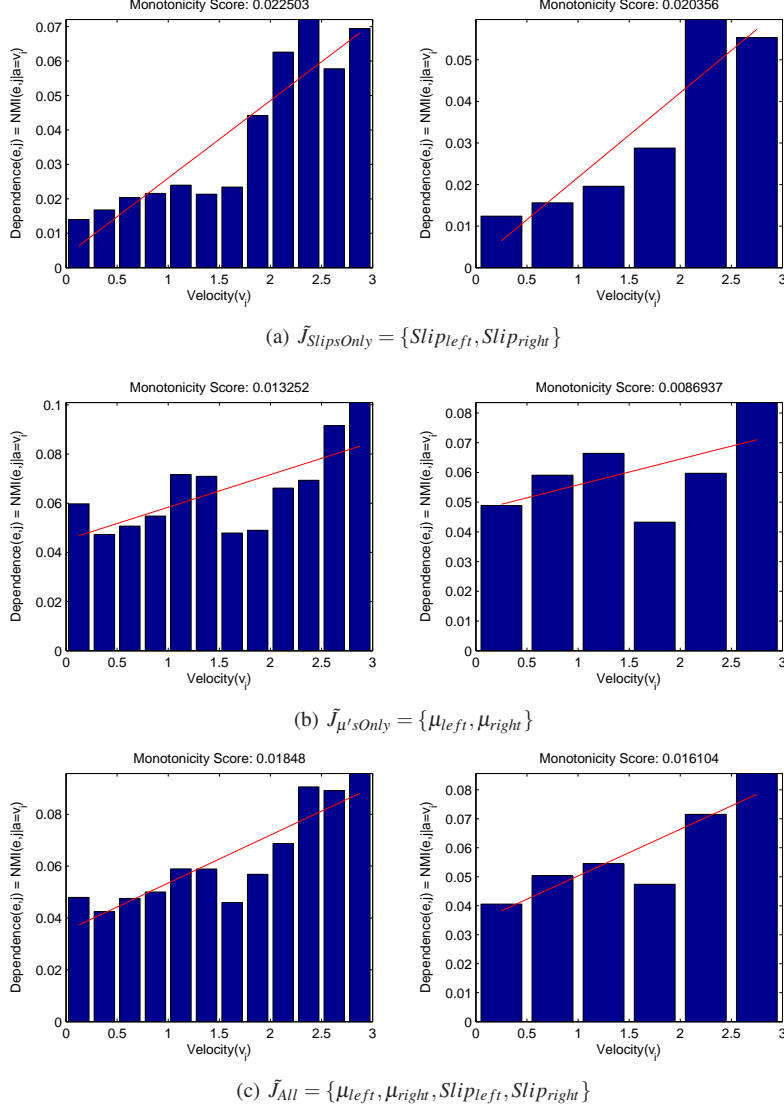


Figure 6: Validity tests with two different bin sizes: X-axis = velocity(v -m/s) & Y-axis = $NMI(\tilde{e}, \tilde{j}|a = v_i)$

The key benefit of the mobility maps with respect to planning is that the cost is orientation sensitive. To leverage this benefit in the A^* algorithm, the arc cost of a connection between two nodes was given as a function of the particular mobility map associated with the direction of this arc. Figure 9(b) demonstrates the desired sensitivity to platform configuration, whereby the path \vec{AB} , and the reverse path \vec{BA} take different routes, since the path taken to go downhill is treated differently from going uphill. In the scalar cost map case, shown in Figure 9(a), the paths \vec{AB} and \vec{BA} are the same.

The A^* paths in Figure 9 only offer heading commands and no information about velocity, so the heuristic cost path in Fig-

ure 9(a) are usually operated with a constant speed preselected for cautious navigation (usually about 1m/s or 3.6 km/hr for the platform in question). For the second case, information from the eight mobility maps can be used to regulate or bound velocities. Average ‘maximum feasible speeds’ for the paths in Figure 9(b) are 6.1121 km/hr in the forward path (white) and 5.9079 km/hr in the backward path (green). These values are an improvement from that of the cautious case as the mobility values adjust to situations of caution by slowing down and situations of confidence by speeding up.

In a separate experiment, controller performance is compared for paths executed on the experimental platform on a

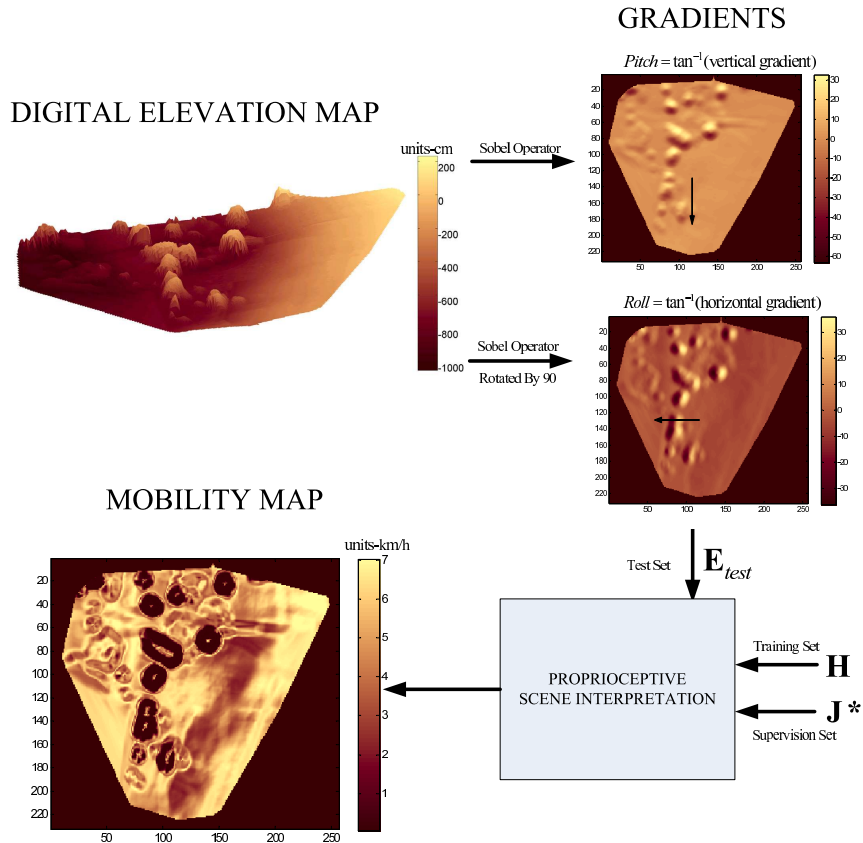


Figure 7: Scene interpretation of an elevation map ($\sim 100 \times 120 \text{m}$ at 0.5m grid resolution) derived from laser data (units of mobility = km/hr). The mobility map shown assumes that the vehicle intends to travel in the downwards direction shown in the pitch gradient map.

5° hill (see Figure 10(a)) with and without mobility information. For a given set of four waypoints, A^* was used to plan across the heuristic scalar cost representation and across mobility maps. The path given by the scalar framework is executed for constant values of velocity starting from 1m/s to 3m/s ($3.6 - 10.8 \text{km/h}$). In contrast the second path planned over mobility maps uses values from the mobility maps to regulate velocity. Figures 10(b) and 10(c) show the input waypoints, the planned paths and the executed paths.

For this experiment, training data were collected under autonomous control where the tuned controller is part of the system. The control system was tuned on flat grass conditions, hence it does not perform well in non-flat conditions (especially downhill where the system becomes significantly underdamped for higher velocities)(see Extension 1). By performing the same scene interpretation process as before with this new training data, the controller limitations are captured into the mobility representation. In Table 3 the paths are compared

for average speed and tracking performance. Tracking performance is judged by standard control theoretic metrics such as root mean squared error (RMSE), L_∞ norm (max error value) and the L_2 norm. It can be seen that the path over mobility maps shows improved controller performance for similar average and maximum speeds as the path with the highest constant velocity of 3m/s . The improvement in tracking performance is more pronounced (shown in Table 4) for the case of traveling downhill where controller limitations are prominent. This improvement is due to the vehicle slowing down in downhill conditions and speeding up on flat terrain.

While the results in Figure 9 illustrate the desired orientation sensitivity in paths, results in Table 3 show that controller limitations are captured in the mobility representation and help in regulating velocity to improve controllability. These results illustrate that accounting for directional mobility are of benefit for UGV planning. Use of mobility information leads to a safer choice of paths with a reduced risk of excessive slippage,

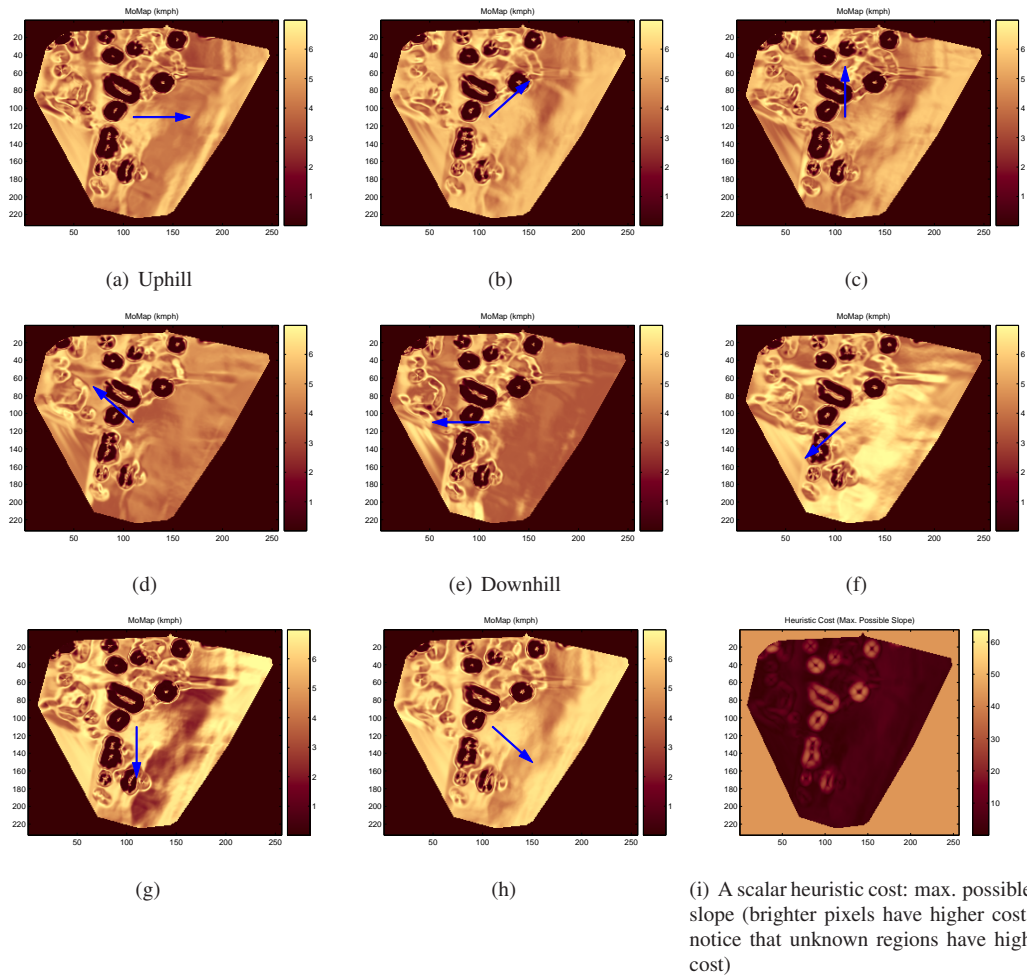
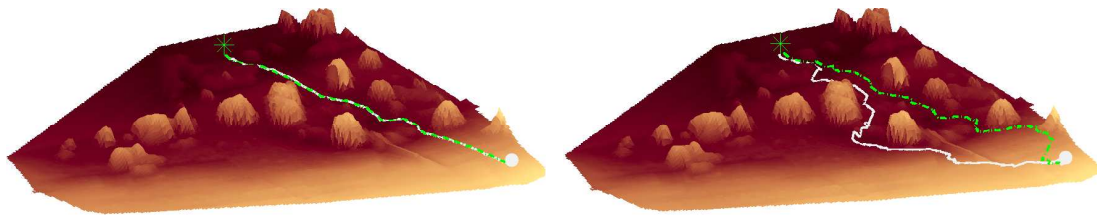


Figure 8: Mobility Maps for eight possible vehicle headings (a-h) and an alternative scalar cost representation (i)



(a) A* path planning on heuristic cost map (path distance: 101m) (b) A* path planning on directional mobility maps (path distance: 121m forward and 109m backward)

Figure 9: Path planning on heuristic cost maps vs. directional mobility maps. [start waypoint - ●, goal waypoint - *, forward path - solid line, backward path - dotted line]

and improved controllability as the system automatically decides on the feasible velocity of operation.

9 Choice of Regression Model

In Section 4.4.2, extensions to Gaussian Process regression were presented to relax the homoscedastic unimodal assumption. In this section, scene interpretation results for different GP extensions are compared and their limitations are discussed.

For the task of deriving an upper bound, the accuracy of GP extensions can be analysed on a toy multimodal dataset with known upper bound. Such a dataset is shown in Figure 11(a). Figure 11(c) shows the clustering results of the iMoGPe model discussed in Section 4.4.2. The iMoGPe model separates the different surfaces and fits a separate GP model for each cluster. The plot in Figure 11(c) shows normalised mean squared error (nMSE) of the predictions for different values of percentage thresholds that are used to determine the upper bound¹⁵. The predictions were performed for ten different random seeds and the error bars for the nMSE are also shown. It can be seen that iMoGPe has the best performance and the heteroscedastic models perform better than the homoscedastic version. The iMoGPe model gets more accurate for high values peaking at 80 percent, the inaccuracy and high variance at the low values is due to the gap in probabilistic mass between various clusters. Both heteroscedastic RVM and heteroscedastic GP models achieve similar accuracy. For a good choice of UB threshold, heteroscedastic models can match the accuracy of the iMoGPe model. The 70-75 percent threshold has the best performance. This is expected because of the Chebyshev’s inequality which states that approximately 70 percent of the probability mass of a unimodal distribution lies between one standard deviation from the mean.

In Figure 12, pair wise differences in the mobility maps are shown for different models. Mobility maps under the homoscedastic assumption lead to conservative estimates of velocity bounds. The upper bound estimates are biased towards lower values as they are visited more often than higher values in data collection which leads to tight noise estimates in training under the homoscedastic model. Figures 12(a), 12(b) and 12(c) show the change in mobility maps between the homoscedastic GP and other models. At worst, the velocity bounds in the homoscedastic model are lower than other models by 1m/s (3.6 kmph). This bias is better accounted when input-dependent noise is considered in the het-

eroscedastic models. In comparison the difference maps in Figures 12(d), 12(e) and 12(f) mostly have minimal difference values ($< 0.2m/s$ (1kmph)) except at dataset boundaries where different models lead to different extrapolation artifacts. For a path planning application, it is the relative difference of mobility between neighbouring states rather than the absolute value that is of significance. The conservative estimates of homoscedastic case can be justified for path planning if the dataset is homogeneously multimodal. However, the assumption breaks down if multimodality is non-stationary and heteroscedastic models are essential to deal with it.

In implementation, training under the heteroscedastic RVM model has more tuning parameters than the heteroscedastic GP model. Greedy selection of relevant vectors in the RVM model is sensitive to random seeds when training is performed on different subsets of the data. Modeling ambiguity is also a problem in the iMoGPe model, where different initial guesses for the hyperparameters caused the algorithm to give different clustering solutions. In contrast, the Heteroscedastic GP training algorithm has fewer tuning parameters and is less sensitive to subset selections because of the projected process approximation. However, it was sensitive to the choice of subset size. In our experiments, the subset size was chosen according to available computing resources.

In an attempt to quantify model uncertainty due to sparse exploration and subset training, different heteroscedastic models were trained 10 times on different subsets. Figure 13 shows the resulting velocity bounds and their errorbars for pure pitch and roll conditions. The worst case ambiguity in the upper bounds can be seen to be around $\pm 0.5m/s$. The regions of high variance are due to lack of training data in that region and different training subsets lead to different interpolation artifacts. This modelling ambiguity is unavoidable in practice and is a shortcoming of a data driven approach.

The training times for different models on a 2GHz PC in a Matlab implementation are shown in Table 5. The iMoGPe is the least feasible on large datasets. It took about 9 hours to train on a subset size of 2000 points for 1000 iterations. In the heteroscedastic case, the projected process approximation gives the opportunity to control complexity with the choice of subset size. The fusion of iMoGPe and projected process approximation is complicated as it involves choosing subset points in conjunction with determining the cluster label¹⁶. The heteroscedastic RVM is faster to train than the GP version as the subset size is grown incrementally. In our experiments the subset size was around 100 datapoints and is learned intrinsi-

¹⁵The upper bound is determined from the cumulative density function by searching for an input that produces an output equal to the selected threshold. Theoretically, this is a query from the quantile function (inverse of the cumulative density function).

¹⁶Wood et al. (2008) present an incremental approach to iMoGPe which can incorporate the projected process approximation in the form mentioned in Csató (2002). However, the approach described in Wood et al. (2008) does not consider hyperparameter updates. The full inference problem of selecting a sparse subset, cluster label and hyperparameters is still an open problem

Table 3: Trajectory following results: Average speed and tracking performance

	Speed performance			Cross track error performance		
	time(s)	Mean Vel(m/s)	Max. Vel(m/s)	RMSE	L_∞ norm	L_2 norm
$A^* + 1m/s$	130	0.95	1.70	1.15	2.34	158.65
$A^* + 2m/s$	73	1.71	2.62	1.61	3.66	254.56
$A^* + 3m/s$	65	1.99	3.41	2.14	5.69	536.04
A^* +Mobility Maps	64	2.07	3.29	1.98	4.54	360.03

Table 4: Trajectory following results for downhill only: Average speed and tracking performance

	Speed performance			Cross track error performance		
	time(s)	Mean Vel(m/s)	Max. Vel(m/s)	RMSE	L_∞ norm	L_2 norm
$A^* + 1m/s$	19	0.89	1.36	1.98	2.34	68.38
$A^* + 2m/s$	10	2.02	2.48	2.35	3.66	106.51
$A^* + 3m/s$	8	2.67	3.41	3.13	5.69	350.63
A^* +Mobility Maps	13	1.58	2.21	2.23	3.58	97.80

cally from the RVM model. As mentioned previously in Section 4.4.2, for the purpose of scene interpretation it is acceptable to have slow training time but fast online predictions are important. Hence, a dual prediction strategy was suggested where the velocity bounds are inferred at training locations and a second RVM model is fitted to the inferred answers. This RVM predictor model is shown for the heteroscedastic GP model in Figure 14. The relevant vectors selected in the model are also shown as black dots and their size is proportional to their inferred weight. This model is downloaded onto the vehicle for fast online predictions which is made possible because of the excellent sparsity achieved (41 relevant vectors).

10 Conclusion and Discussion

This paper presents a non-parametric learning technique to interpret terrain slopes from wheel-slip observations, and describes the benefits of defining costs in behaviour space (as behavioural limits). First, this enables experience-based learning and encapsulates platform and controller limitations in a common representation. Second, behaviour dependent costs can be created for aiding decision making, such as orientation sensitive costs for UGV path planning. Finally, proprioceptive feedback such as wheel slip can be incorporated and used to learn effectively in complex environments.

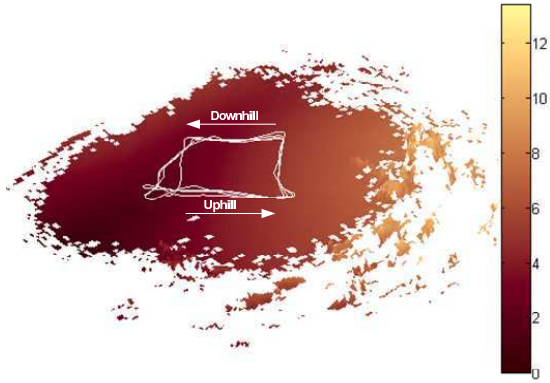
An additional benefit for using proprioceptive feedback is that it minimises input required from an expert. Instead of providing low-level supervision on what decisions to make in each condition, expert supervision is now concentrated on

higher level design choices of where to explore and what to label as nominal/feasible operation. By looking at dependence trends in the observations, questions about the validity of proprioceptive feedback measures can be answered. Therefore, an information theoretic test was proposed to validate the use of vehicle slip observations to interpret terrain conditions. Also, sparse approximations to non-parametric regression were discussed in order to make this work feasible for training on large datasets and generating fast online predictions.

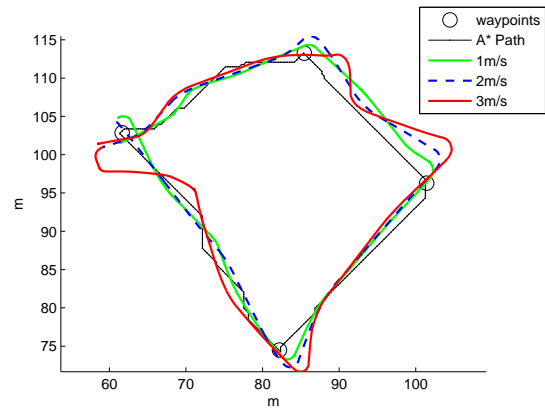
Although the experiments in this paper were concentrated on interpreting terrain slopes from range sensor measurements, additional terrain states such as colour and texture can be incorporated without increase in complexity as GP regression is a discriminative modelling technique.

Non-parametric techniques and experience based learning are a sensible approach in off-road unstructured conditions where simplifying assumptions about the environment cannot be made. They are practical for off-line analysis of sensor information to be used as prior data for path planning. More importantly, they provide a theoretical approach to the process of defining and generating costs.

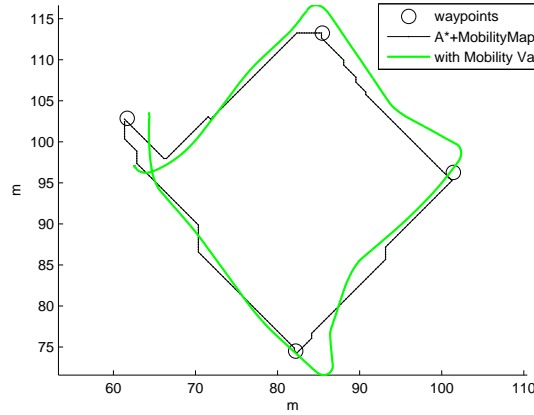
In this current work, a simple upper bound on velocity was used as a scene interpretative metric for combinatorial planning with A^* . Temporal correlations between successive velocities was not taken in account for decision making; for that to be achieved mobility values can be used as instantaneous rewards into a Reinforcement Learning formulation. Also, the validity test proposed in this work was confined to a 1D case. For more complicated platforms, the velocity limits could be replaced with kinetic energy limits as an alternative approach to capture control effort in agent-environment inter-



(a) Trajectory test environment - A 5° Hill



(b) A* Path on scalar heuristic cost and the executed trajectories with different constant velocities

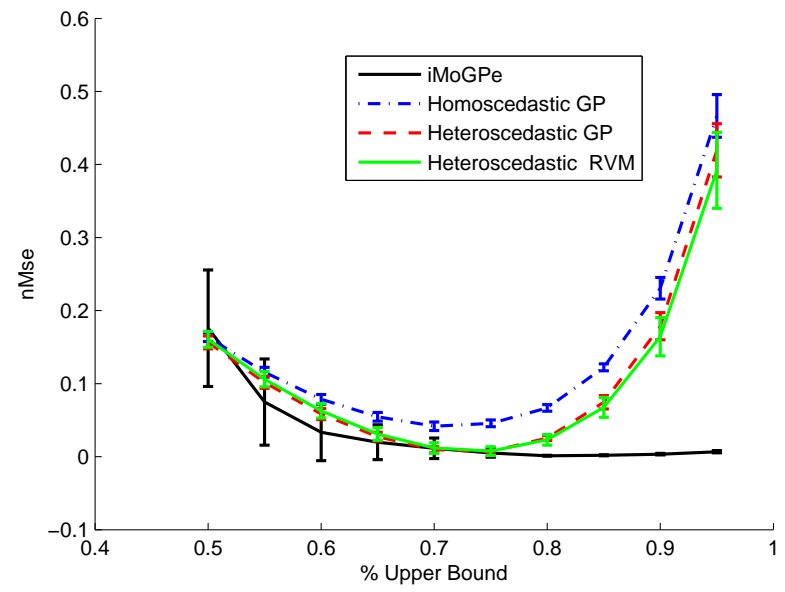
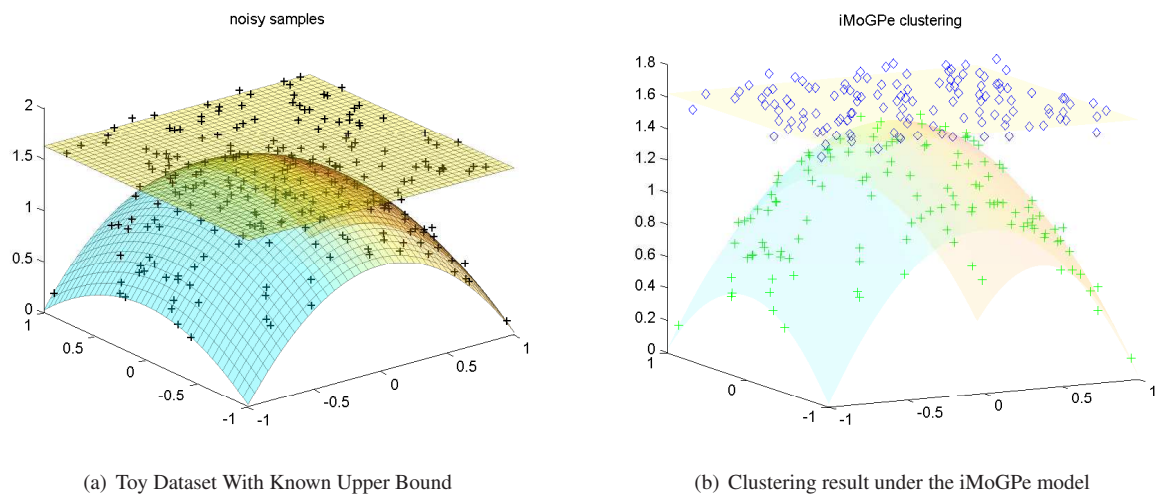


(c) A* Path on mobility maps and the executed trajectory with mobility values

Figure 10: Trajectory following experiment

Table 5: Training Times (M : subset size, N : dataset size)

	Homoscedastic GP - PP	Heteroscedastic GP - PP	Heteroscedastic RVM	iMoGPe
Data Size:	≈ 15 minutes ($M = 2000; N = 10000$)	≈ 1 hour 20minutes ($M = 2000; N = 10000$)	≈ 30 minutes ($M \approx 100; N = 10000$)	≈ 9 hours ($N = 2000$)
Notes:	Projection Process Approx.	Projection Process Approx.	Subset(M) is learned	



(c) nMSE vs. Upper Bound Threshold For Different Techniques

Figure 11: Choice of Regression Models for a Multimodal Dataset

actions. However in the future, it would be valuable to look at higher dimensional problems, in conjunction with stochastic trajectory planning algorithms that sample from behaviour-centric conditional distributions such as the one proposed in this work.

Acknowledgment

This work is supported by the ARC Centre of Excellence programme, funded by the Australian Research Council (ARC)

and the New South Wales State Government. The authors would also like to thank James Underwood and Thierry Peynot for their support during the field experiments.

References

Abbeel, P. and Ng, A. Y. (2004). Apprenticeship learning via inverse reinforcement learning. In 21st International Conference on Machine Learning. Banff, Canada.

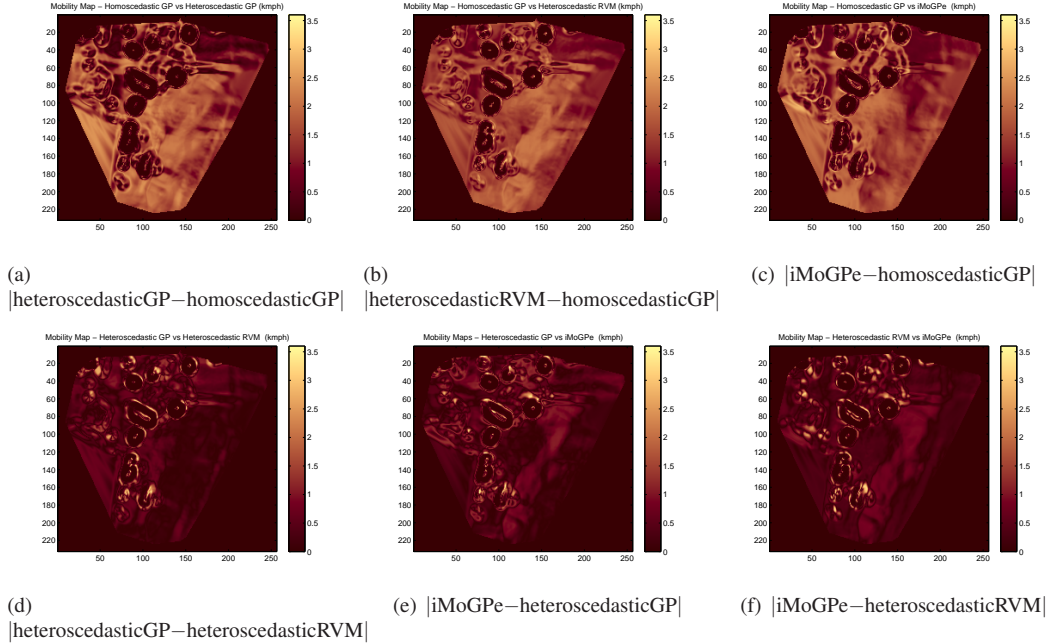


Figure 12: Mobility map differences for different techniques

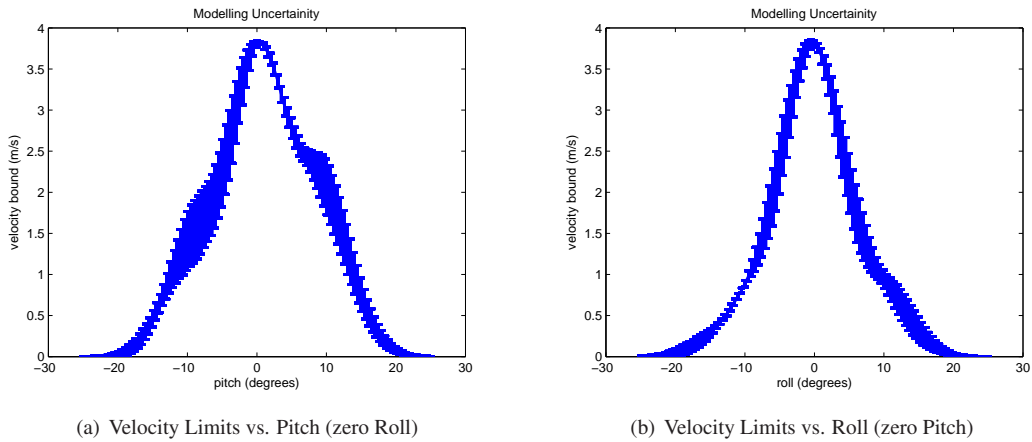


Figure 13: Modelling Uncertainty

Angelova, A., Matthies, L., Helmick, D., and Perona, P. (2006). Slip prediction using visual information. *Robotics: Science and Systems (RSS)*.

Angelova, A., Matthies, L., Helmick, D., and Perona, P. (2007). Learning and prediction of slip using visual information. *Journal of Field Robotics*.

Bishop, C. M. (2006). *Pattern Recognition And Machine Learning*. Springer.

Csató, L. (2002). *Gaussian Processes - Iterative Sparse Approximations*. Phd thesis.

Goldberg, P., Williams, C., and Bishop, C. (1998). Regression with input-dependent noise: A gaussian process treatment. In *Advances in Neural Information Processing Systems (NIPS)*. MIT Press.

Gonzalez, R. C. and Woods, R. E. (2008). *Digital Image Processing*. Prentice Hall, 3rd edition edition.

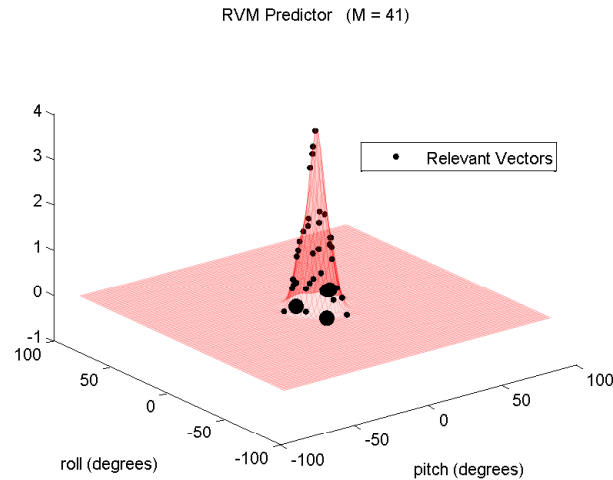


Figure 14: RVM fit to velocity bounds inferred from Heteroschedastic GP Model

- Green, A. R. and Rye, D. (2007). Sensible planning for vehicles operating over difficult unstructured terrains. In *IEEE Aerospace Conference*, pages 1–8.
- Hadsell, R., Sermanet, P., Erkan, A. N., Ben, J., Han, J., Flepp, B., Muller, U., and LeCun, Y. (2007). Online learning for offroad robots: Using spatial label propagation to learn by long-range traversability. *Proceedings of Robotics: Science and Systems*.
- Jackel, L. D., Krotkov, E., Perschbacher, M., Pippine, J., and Sullivan, C. (2006). The darpa lagr program: Goals, challenges, methodology, and phase i results. *Journal of Field Robotics*, 23(11-12):945–973.
- Julier, S. J. (1996). *Process Models for the Navigation of High-Speed Land Vehicles*. Ph.d thesis, The University of Oxford.
- Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F. (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators. In *IEEE Transactions on Automatic Control*, volume 45.
- Karumanchi, S., Allen, T., Bailey, T., and Scheduling, S. (2009). Non-parametric learning to aid path planning over slopes. In *Proceedings of Robotics: Science and Systems*. Seattle, USA.
- Kersting, K., Plagemann, C., Pfaff, P., and Burgard, W. (2007). Most likely heteroschedastic gaussian process regression. In *Proceedings of 24th International Conference on Machine Learning*. Corvallis, OR.
- Lacroix, S., Mallet, A., Bonnafous, D., Bauzil, G., Fleury, S., Herrb, M., and Chatila, R. (2002). Autonomous rover navigation on unknown terrains: functions and integration. *International Journal of Robotics Research*, 21(10-11):917–942.
- Le, A. T., Rye, D., and Durrant-Whyte, H. (1997). Estimation of track-soil interactions for autonomous tracked vehicles. *IEEE International Conference on Robotics and Automation*, 2:1388–1393.
- Meeds, E. and Osindero, S. (2006). An alternative infinite mixture of gaussian process experts. In *Advances in Neural Information Processing Systems 18 (NIPS)*, pages 883–890.
- Noë, A. (2005). *Action in Perception*. MIT PRESS.
- Ojeda, L., Borenstein, J., Witus, G., and Karlsen, R. (2006). Terrain characterization and classification with a mobile robot. *Journal Of Field Robotics*.
- Quiñonero Candela, J. and Rasmussen, C. E. (2005). A unifying view of sparse approximate gaussian process regression. *Journal of Machine Learning Research*, 6(12):1935–1959.
- Rasmussen, C. E. (2000). The infinite gaussian mixture model. In *Advances in neural information processing systems*.
- Rasmussen, C. E. and Ghahramani, Z. (2002). Infinite mixtures of gaussian process experts. In *Advances in Neural Information Processing Systems 14*, pages 881–888. MIT Press.
- Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press.

Ratliff, N. (2009). Learning to Search: Structured Prediction Techniques for Imitation Learning. Phd thesis.

Seeger, M., William, C., and Lawrence, N. (2003). Fast forward selection to speed up sparse gaussian process regression. In Workshop on Artificial Intelligence and Statistics 9.

Shneier, M., Chang, T., Hong, T., Shackleford, W., Bostelman, R., and Albus, J. S. (2008). Learning traversability models for autonomous mobile vehicles. *Autonomous Robots*, 24(1).

Silver, D., Bagnell, J., and Stentz, A. (2008). High performance outdoor navigation from overhead data using imitation learning. In *Robotics: Science and Systems IV*. Zurich, Switzerland.

Silver, D., Sofman, B., Vandapel, N., Bagnell, J., and Stentz, A. (2006). Experimental analysis of overhead data processing to support long range navigation. In *IEEE International Conference on Intelligent Robots and Systems (IROS)*, pages 2443 – 2450.

Sofman, B., Ratliff, E. L., Bagnell, J., Cole, J., Vandapel, N., and Stentz, A. (2006). Improving robot navigation through self-supervised online learning. *Journal of Field Robotics*, 23(12).

Stavens, D., Hoffmann, G., and Thrun, S. (2007). Online speed adaptation using supervised learning for high-speed, off-road autonomous driving. In *Proc. International Joint Conference on Artificial Intelligence (IJCAI)*.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. The MIT Press.

Tipping, M. E. (2001). Sparse bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1:211244.

Tipping, M. E. and Faul, A. C. (2003). Fast marginal likelihood maximisation for sparse bayesian models. In *Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics*. Key West, FL.

Wong, J. Y. (2001). *Theory Of Ground Vehicles*. New York: Wiley, 3rd edition edition.

Wood, F., Grollman, D. H., Heller, K. A., Jenkins, O. C., and Black, M. J. (2008). Incremental nonparametric bayesian regression. Technical Report Technical Report CS-08-07, Brown University Dept. Computer Science.

Appendix A: Index to Multimedia Extensions

Extension	Media Type	Description
1	Video	Controller performance in downhill conditions

Appendix B: Process Model For Slip Estimation

Proprioceptive measures such as slip and traction coefficients cannot be measured directly, so they are estimated with an UKF. Previous work used such process models to try and improve localisation for navigation (Julier 1996). In contrast, for proprioception, good positioning solution is assumed and estimators/observers are designed to monitor the selected proprioceptive states at all times. Wong (2001) defines longitudinal slip as the rotation of the tyre without equivalent translatory progress when a wheel torque is applied:

$$slip(i) = \frac{V - R\omega}{\max(R\omega, V)}$$

where

V = Longitudinal Velocity

ω = Wheel Rate

R = Tyre Radius

In order to get slip estimates from an UKF, we use a slightly modified version of the two-track kinematic model presented in Le et al. (1997) aimed at tracked vehicles. This modified version fixes the problem of poor observability when the vehicle is stationary or close to stationary in the original model. The UKF estimates slip given accurate GPS observations of position. Further details can be found in Le et al. (1997).

$$\dot{X} = \frac{R}{2} (\omega_L + d\omega_L + \omega_R + d\omega_R) (\cos \phi - \sin \phi \tan \alpha);$$

$$\dot{Y} = \frac{R}{2} (\omega_L + d\omega_L + \omega_R + d\omega_R) (\sin \phi + \cos \phi \tan \alpha);$$

$$\dot{\phi} = \frac{R}{WB} ((\omega_R + d\omega_R) - (\omega_L + d\omega_L));$$

$$d\dot{\omega}_L = v_1(t);$$

$$d\dot{\omega}_R = v_2(t);$$

$$\dot{\alpha} = v_3(t);$$

where

$v_1(t), v_2(t), v_3(t)$ – noise inputs;

X – Northing

Y – Easting

ϕ – Heading

ω_L, ω_R – Wheel Rates

R – Tyre Radius

WB – Wheel Base

$d\omega_L, d\omega_R$ – Wheel Rate Increments

$$i_L = \frac{-d\omega_L}{\omega_L}$$

$$i_R = \frac{-d\omega_R}{\omega_R}$$

i_L, i_R – Longitudinal Slip

α – Slip Angle

m – Vehicle Mass

where

V – Longitudinal Velocity

ψ – Turnrate

V_L, V_R – Track Velocities

F_{ZL}, F_{ZR} – Normal loads

TW – Track Width

I – Moment of Inertia

β – Pitch

μ_L, μ_R – Traction Coefficients

$v_1(t), v_2(t)$ – noise inputs;

Appendix C: Process Model For Traction Estimation

The extent of traction generated is usually described in terms of adhesion/traction coefficients which are defined as the ratio of resultant tractive forces with the normal loads exerted on the ground by the vehicle:

$$\mu = \frac{F}{F_z} = \frac{\text{Traction}}{\text{Normal Load}}$$

Similar to the kinematic two-track model, a dynamic version can be used to estimate traction coefficients given accurate measurements of velocity and turnrate from the navigation system.

$$\dot{V} = \frac{1}{m}(\mu_L F_{ZL} + \mu_R F_{ZR}) - g \sin(\beta);$$

$$\dot{\psi} = \frac{TW}{2I}(\mu_R F_{ZR} - \mu_L F_{ZL}) - b_I \psi;$$

$$\dot{\mu}_L = v_1(t);$$

$$\dot{\mu}_R = v_2(t);$$