2019 International Conference on Robotics and Automation (ICRA)
Palais des congres de Montreal, Montreal, Canada, May 20-24, 2019

Support Surface Estimation for Legged Robots

Timon Homberger[1], Lorenz Wellhausen[1], Péter Fankhauser[2], Marco Hutter[1]

Abstract— **The high agility of legged systems allows them to operate in rugged outdoor environments. In these situations, knowledge about the terrain geometry is key for foothold planning to enable safe locomotion. However, on penetrable or highly compliant terrain (e.g. grass) the visibility of the supporting ground surface is obstructed, i.e. it cannot directly be perceived by depth sensors. We present a method to estimate the underlying terrain topography by fusing haptic information about foot contact closure locations with exteroceptive sensing. To obtain a dense support surface estimate from sparsely sampled footholds we apply Gaussian process regression. Exteroceptive information is integrated into the support surface estimation procedure by estimating the height of the penetrable surface layer from discrete penetration depth measurements at the footholds. The method is designed such that it provides a continuous support surface estimate even if there is only partial exteroceptive information available due to shadowing effects. Field experiments with the quadrupedal robot ANYmal show how the robot can smoothly and safely navigate in dense vegetation.**

I. Introduction

Legged robots have the ability to maneuver in rugged terrain, which opens a vast range of applications. Their superior agility over wheeled systems is largely because they require only a sparse set of contact points with the ground during locomotion on rough terrain. To select these footholds carefully, an accurate and high-resolution model of the local terrain geometry is necessary. Because geometric environmental perception is usually based on exteroceptive sensors like depth cameras or LiDARs, they can only capture terrain which is in the sensor field-of-view (FOV) which typically does not cover the area immediately underneath the robot. They also suffer from occlusions which leads to problems in environments where the robot penetrates the upper terrain layers and sensor returns do not correspond to rigid geometry. While traversing e.g. tall grass, direct visibility of the supporting ground structure is not given and the actual contact point between the foot tip and the ground differs from the expected terrain height. As a consequence, locomotion planning may yield motion sequences that cause instability or there may be no feasible solution to the foot tip trajectory planning problem at all.

This work was supported in part by the Swiss National Science Foundation (SNF) through the National Centre of Competence in Research Robotics and by the EU H2020 Project THING.

This work has been conducted as part of ANYmal Research, a community to advance legged robotics.

[1]Authors are with the Robotics Systems Lab (RSL), ETH Zürich, Switzerland authors@mavt.ethz.ch

[2]Author is with ANYbotics AG, Zürich, Switzerland

Fig. 1: Support surface estimation allows ANYmal to navigate through densely vegetated environments in which the underlying terrain structure is not perceivable by exteroceptive sensing.

A. Contribution

This work introduces a combined proprioceptive-exteroceptive terrain estimation architecture. It allows estimation of the support surface in environments where it is not perceivable by exteroceptive onboard sensors, while retaining the accuracy of current elevation mapping methods when the support surface is visible. It thereby enables the robot to robustly navigate through rugged environments in which the terrain cannot directly be perceived. While the estimation architecture is applicable to other penetrable surfaces, like water or sand, we hereafter refer to them simply as *vegetation*.

Our method relies on combining data from exteroceptive and proprioceptive sensors. A depth camera provides exteroceptive sensing, which is used to generate a 2.5D elevation map of the perceived geometry [1]. Contact positions between foot tips and the ground obtained from leg joint states and forward kinematics serve as proprioceptive measurements.

The procedure involves estimating the support surface in two different ways. The first method exploits the fact that ground contact positions of the foot tips sample the actual underlying terrain structure sparsely. The second method relies on combined proprioceptive and exteroceptive data. It uses an approximation of the height of the vegetation to reason about the topography of the ground. The two resulting ground estimates are combined using an adaptive weighting scheme to yield the final support surface estimate. The interpolation methods used for approximation of the terrain topography are such that they promote continuity in terrain and vegetation height changes respectively. Similar assumptions on continuity of environmental changes have been made in other works in the field of outdoor navigation [2], [3].

Real world experiments were conducted with the quadrupedal robot ANYmal [4] in environments that feature

uneven terrain and dense vegetation of varying height. Our results show that with the presented method safe and stable navigation through vegetation is feasible even if exteroceptive sensors fail to perceive the underlying terrain structure. Furthermore, we show that the system allows to adapt the support surface estimation mechanism to penetrable as well as rigid terrain. In addition, the method is designed such that it allows to generate a smooth estimate of the support surface even if there are regions that lack exteroceptive data.

B. Related Work

Incorporating terrain perception into the control architecture of legged robots allows for foot tip trajectory planning and thus choosing good motion sequences to reach a given navigation goal [5]. The data of a depth sensor, combined with the state estimate of the system are used to generate a model of the environment. In our work, a robot-centric elevation mapping approach is used, i.e. the terrain structure is expressed in a coordinate frame that is attached to the robot. Advantages compared to world-centric mapping are lower impact of state estimator drift on the elevation map [1], [6]. Furthermore world-centric mapping approaches require accurate global localization to allow for the generation of a consistent map of the environment [7].

A method for ground detection in an agricultural context is described in [8] which is based on an iterative probabilistic inference algorithm to estimate ground height. It uses a histogram of perceived heights in vegetation and assumes visibility into the vegetation to various depths or even partly to the ground. However, our work addresses general environments containing areas of completely obscured terrain and does not rely on partial visibility of the ground.

For quadruped locomotion in rugged terrain Kolter et al. [9] introduce an algorithm to fill in missing parts of the terrain map. Using a stereo camera, many locations cannot be perceived in rough terrain, as objects can significantly obstruct the line of sight. Missing map regions are filled-in using an approach that is inspired by texture synthesis methods in computer graphics. Our support surface estimation scheme is designed such that it can deal with such occlusions of exteroceptive elevation data.

The work of [10] describes a method to infer vegetation height from statistical elevation data, such as entropy and variance for large scale terrain mapping from satellite imagery. The hexapod presented by [11] features a vegetation detection method that relies on comparison of proprioceptive and exteroceptive terrain variance estimates and switches between two locomotion modes.

The authors of [12] present a system for ground estimation concerning locomotion on terrain which is totally obscured by vegetation. They use Markov random fields to model the assumptions of continuous changes of terrain and vegetation height. Using a 3D voxel map and a terrain classification scheme that relies on a multitude of exteroceptive sensors, voxel cells are assigned to the categories "vegetation", "ground", "obstacle" and "free space". Furthermore they use a hidden semi Markov model for the class distributions in vertical cell columns. Other research has been conducted on robust terrain classification for mobile robots [13], [14], [15] which could be leveraged to predict vegetation with discrete height values. However, this work presents a support surface estimation system that does not rely on classification but rather regresses the height estimate directly from sensor measurements.

Our method applies a Gaussian process regression (GPR) scheme which has been applied in a similar fashion for large scale terrain modelling before. The authors of [16] use GPR to generate large scale multiresolution models inferred from partly uncertain and incomplete data. In the work of [17] GPR is used for terrain regression with a depth camera on a quadruped. To manage computational cost they introduce a data separation method, referred to as model tiling. This concept is applied and described in this work (see section II-F.2).

II. Method

Our method involves generating estimates of the support surface in two different ways and then fusing them. The first is derived from the contact positions of the foot tips with the ground and the second via estimating the vegetation height. All estimates, each of which is represented as a 2.5D height map, are referred to as layers of the elevation map [18].

A. Overview and Definitions

In Figure 2 the different layers that are involved in the support surface estimation procedure are depicted. The *visible topography* layer represents the topography that is directly perceivable by the depth camera. The *support surface* is the layer which the robot is walking on. Fitting a smooth function into the sparse set of footholds serves as a purely proprioceptive estimate of the support surface which we refer to as the *foothold map* layer. The height difference between a foothold and the *visible topography* layer at the footholds location is defined as a *vegetation height measurement*. Fitting a smooth function through a discrete set of *vegetation height measurements* yields an estimate of the *vegetation height*. We refer to the support surface estimate obtained by subtracting the vegetation height estimate from the *visible topography* layer as the *vegetation subtraction* layer. *Vegetation subtraction* layer and *foothold map* layer are weighted based on empirical environment characteristics, i.e. foothold variance and vegetation height variance, and fused to obtain our final estimate of the support surface.

B. Gaussian Process Regression

For the computation of the *foothold map* layer (see section II-C) and as a part of the occlusion adaptive smoothing (see section II-F.2), a GPR scheme is used [19], which is a non-parametric regression method. A Gaussian process defines a distribution of functions in the form of a mean and a covariance function.

A finite set of samples[1] $\{z_1, ..., z_n\}$ is distributed as

[1]Here assumed to be scalars, but can be higher dimensional in other application scenarios.

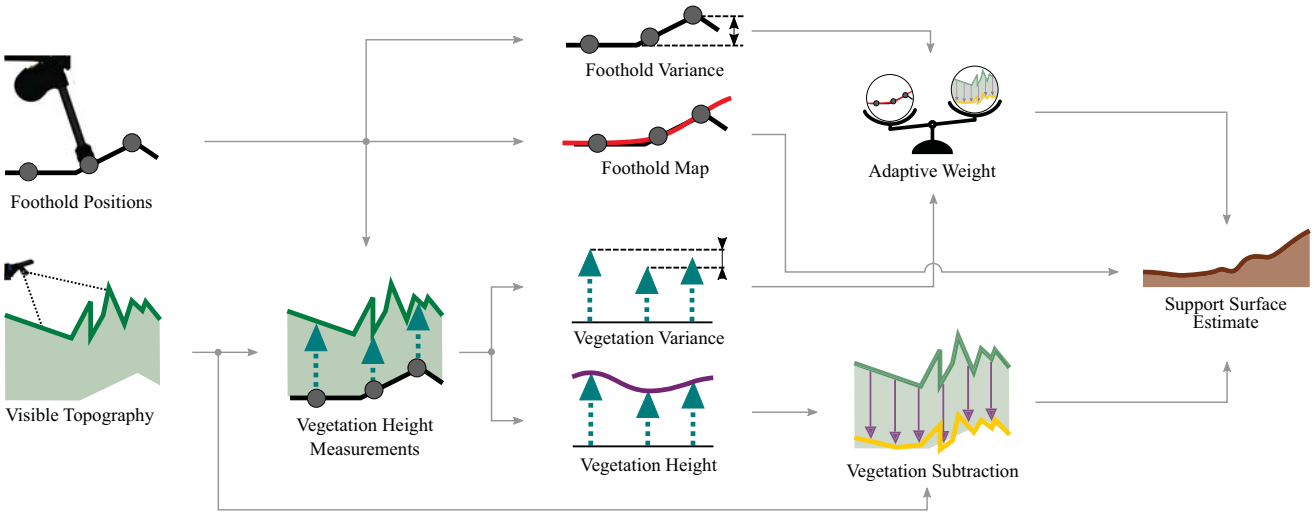


Fig. 2: Overview of the Support Surface Estimation Procedure. (Visualized as cross sections of the respective layers of the elevation map.) Foothold positions are used to obtain a purely proprioceptive height estimate. Additionally, the footholds are compared to the visible topography to estimate the vegetation height, which is used to compute the *vegetation subtraction* layer. *Foothold map* layer and *vegetation subtraction* layer are fused with a weighting scheme which is based on their variance. This yields the support surface estimate.

$$P(z_1, ...z_n|\mathbf{x}_1, ...\mathbf{x}_n) \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{K}) \tag{1}$$

with $\mu \in \mathbb{R}^n$ the mean and the covariance matrix $\mathbf{K}$. $\mathbf{K}$ is specified as $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j) + \sigma_n \delta_{ij}$ where $k(\mathbf{x}_i, \mathbf{x}_j)$ is a covariance kernel and σ_n is a global noise parameter. Given a set of training data $D = \{\mathbf{x}_i, z_i\}_{i=1}^n$ the target value z^* for a given test input[2] $\mathbf{x}^*$ is distributed as $\mathcal{N}(\mu^*, v^*)$ where

$$\mu^* = \mathbf{k}^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{z} \tag{2}$$

$$v^* = k^* + \sigma_n^2 - \mathbf{k}^T(\mathbf{K} + \sigma_n^2\mathbf{I})^{-1}\mathbf{k} \tag{3}$$

with $\mathbf{z}$ the vector of training targets and $\mathbf{K} \in \mathbb{R}^{n\times n}$, $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, $\mathbf{k} \in \mathbb{R}^n$, $k_j = k(\mathbf{x}^*, \mathbf{x}_j)$, $k^* = k(\mathbf{x}^*, \mathbf{x}^*)$ [16], [17].

For this application in the field of terrain mapping the input data $\{\mathbf{x}_i\}_{i=1}^n$ are the horizontal locations, i.e. 2D positions of elevation map cells. The target values $\{z_i\}_{i=1}^n$ are the corresponding elevation values of the cells. The covariance kernel $k(\mathbf{x}_i, \mathbf{x}_j)$ is therefore a function of the spatial relation between two cell positions.

C. Foothold Map Layer

As outlined in section II-A the *foothold map* layer is a proprioceptive estimate of the support surface. It is obtained by fitting a smooth curve into a discrete set of footholds, using a GPR method as described in the previous section. The training input data points are the horizontal components of the foothold positions. The training targets are the heights of the footholds, while the resulting test targets are the elevation values of the *foothold map* layer.

Figure 3 demonstrates that the type of covariance kernel used in GPR has an influence on the shape of the *foothold map* layer. The squared exponential[3] kernel, one of the most commonly used covariance kernels, which is defined as

$$k_{\text{SQE}}(x_i, x_j) = \sigma_f^2 \exp\left(\frac{(x_i - x_j)^2}{2l^2}\right), \tag{4}$$

yields an extrapolation of the terrain slope and has a relatively strong smoothing effect on discontinuities.

The Ornstein-Uhlenbeck Kernel

$$k_{\text{OU}}(x_i, x_j) = \sigma_f^2 \exp\left(\frac{\|x_i - x_j\|_1}{l}\right) \tag{5}$$

immediately slopes towards zero outside of the range of the footholds. However, the kernel yields little interpolation error at discontinuities in the input data.

For both kernels σ_f is a signal variance parameter and l the lengthscale. The latter can intuitively be seen as a scaling of the spatial distance within which two output datapoints correlate to a certain degree [16].

Using the kernel addition rule, we define a new kernel

$$k_{\text{COMB}} = \sigma_f^2\big(0.94\exp\Big(\frac{(x_i - x_j)^2}{2l^2}\Big) + 0.06\exp\Big(\frac{\ell 1(x_i - x_j)}{l}\Big)\big). \tag{6}$$

Figure 3 shows that the combined kernel yields extrapolation of the terrain slope, as well as good interpolation at discontinuities.

D. Vegetation Subtraction Layer

We obtain a *vegetation height* estimate by applying a distance based averaging function to a discrete set of *vegetation height measurements*.

The elevation of a cell of the *vegetation height* layer $z_{\text{VHE},ij}$ is computed from a set of N *vegetation height measurements* $z_{\text{VHM},m}$ as

$$z_{\text{VHE},ij} = \frac{\sum_{m=1}^{N} \frac{z_{\text{VHM},m}}{(d_{m,ij})^5}}{\sum_{m=1}^{N} \frac{1}{(d_{m,ij})^5}} \tag{7}$$

[2]In this work z^* is referred to as the test target and $\mathbf{x}^*$ as the test input.

[3]In literature also referred to as Gaussian kernel or the radial basis function.

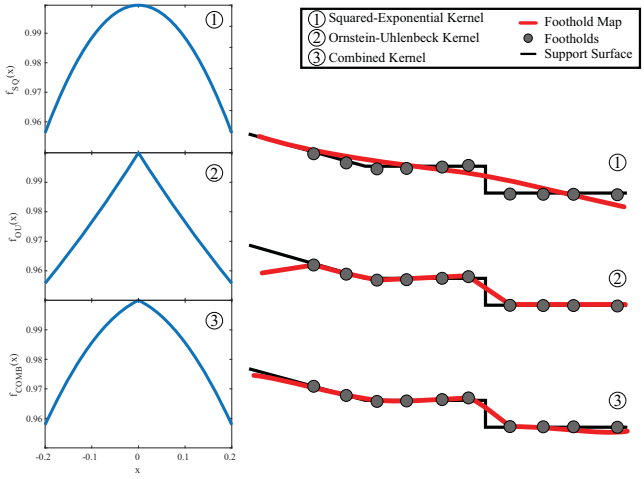


Fig. 3: Left: Functional Form of Kernels, Right: Orthonormal sideways projection of the *foothold map* layer for different covariance kernels, equal kernel parameters and equal support surface structure.

where $d_{m,ij}$ is the horizontal distance between the location of the considered cell and the location of a *vegetation height measurement* $z_{\text{VHM},m}$.

Note that in areas where no foot tip contact data is available, i.e. outside of the range of the footholds, this method yields a highly continuous vegetation height estimate.[4]

For the calculation of the *vegetation subtraction* layer our estimate of the *vegetation height* is subtracted from the *visible topography* layer.

E. Proprioceptive Terrain Characteristics

The weights for combining the *foothold map* layer and the *vegetation height* layer to obtain our final support surface estimate are derived from estimated surface characteristics. We choose them to be the empirical foothold variance as a representation of the uncertainty of the *foothold map* layer and the empirical *vegetation height* variance as the uncertainty of the *vegetation height*.

1) Foothold Variance Estimation: The foothold variance V_F is calculated as

$$\text{V}_\text{F} = \frac{\sum_{m=2}^{N} \left(z_{F,m} - z_{F,m-1}\right)^2}{N-2} - \left(\frac{\sum_{m=2}^{N} \left(z_{F,m} - z_{F,m-1}\right)}{N-2}\right)^2 \tag{8}$$

where $z_{FH,m}$ is the height of the foothold. V_F is the variance of the difference in elevation between consecutive footholds.[5] It serves as a measure for the discontinuity of the support surface and is calculated with respect to the last N footholds.

2) Vegetation Height Variance Estimation: The *vegetation height* variance V_VH is the variance of consecutive *vegetation height measurements* $h_{\text{VH},m}$ and is computed as

$$\text{V}_\text{VH} = \frac{\sum_{m=2}^{N} \left(h_{\text{VH},m}\right)^2}{N-1} - \left(\frac{\sum_{m=2}^{N} h_{\text{VH},m}}{N-1}\right)^2. \tag{9}$$

It serves as a measure for the discontinuity of the *vegetation height*.

F. Support Surface Update

After each stance phase the support surface estimate is updated in a circular area which is centered at the newest foothold position. For this update the *foothold map* layer, the *vegetation subtraction* layer, the foothold variance and the *vegetation height* variance are used as input for an adaptive weighting procedure. We apply GPR-based smoothing with a squared-exponential kernel. The training target for cell position x_{ij} is

$$z_{\text{train},ij} = w_{\text{FML},ij} z_{\text{FML},ij} + w_{\text{VSL},ij} z_{\text{VSL}} \tag{10}$$

where $z_{\text{FML},ij}$ and $z_{\text{VSL},ij}$ are the elevation of cell c_{ij} of the *foothold map* layer and the *vegetation subtraction* layer respectively. $w_{\text{FML},ij}$ and $w_{\text{VSL},ij}$ are their corresponding weights (see section II-F.1).

Because the computational complexity of GPR scales cubic with the number of input data points, a model tiling method is used to reduce the computational expense. The set of input data points is divided into overlapping circular subsets, called tiles, and GPR is performed on each tile. In overlapping areas the results are combined by averaging. By resizing these tiles, time efficiency and accuracy can be balanced [17].

1) Terrain Adaptive Weighting Procedure: The weight $w_{\text{FML},ij}$ of the *foothold map* layer of cell c_{ij} is determined by

$$w_{\text{FML},ij} = \exp\left(-d_{ij} \frac{a \text{V}_\text{F}}{\text{V}_\text{VH}^b}\right) \tag{11}$$

where d_ij is the horizontal distance between the newest foothold and the cell c_{ij}.[6] V_F is the empirical foothold variance, V_VH the empirical vegetation height variance and a, b are two parameters which are used to control the relative impact of V_F and V_VH on the weight. The weight of the *vegetation subtraction* layer is $w_{\text{VSL},ij} = 1 - w_{\text{FML},ij}$.

To illustrate the effect of Equation (11) we will now give a qualitative description of the emerging behavior in different scenarios, pictured in Figure 4. If $\text{V}_\text{F} \ll \text{V}_\text{VH}$, the support surface sampled by the footholds is very even while there are big fluctuations in the vegetation height. Consequently, we would like our support surface estimate to rely mostly on proprioceptive measurements. In contrast, if V_VH is small, we can assume that the vegetation height is almost constant and the *vegetation subtraction* allows us to make accurate predictions about the support surface from exteroceptive sensors. This also captures the special case where the support surface is completely unobstructed and

[4]It is a close to constant function, as the impact of a specific vegetation height measurement $z_{\text{VHM},m}$ decays in proportion to $(d_{m,ij})^5$.

[5]Note that by considering the differences between consecutive foothold elevations, uniform slopes are assigned zero variance.

[6]Therefore in close proximity to the newest foothold there is more weight on the purely proprioceptive source of data, i.e. the *foothold map* layer.

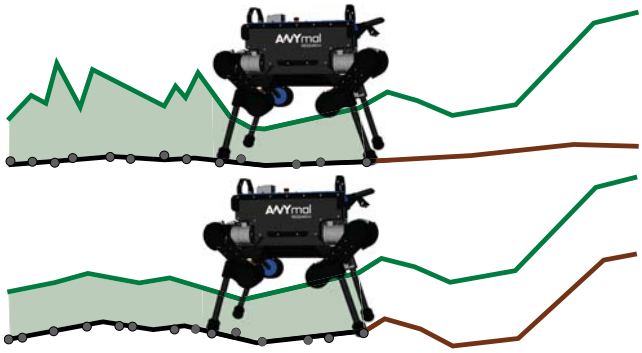


Fig. 4: The methods effect on support surface prediction outside of the range of the footholds. (Visualized as cross sections of the respective layers of the elevation map.) (Top) Terrain is relatively continuous, therefore relatively constant terrain evolution is predicted by the support surface estimate. (Bottom) Vegetation height is relatively constant, support surface estimate is therefore mostly based on the *visible topography*.

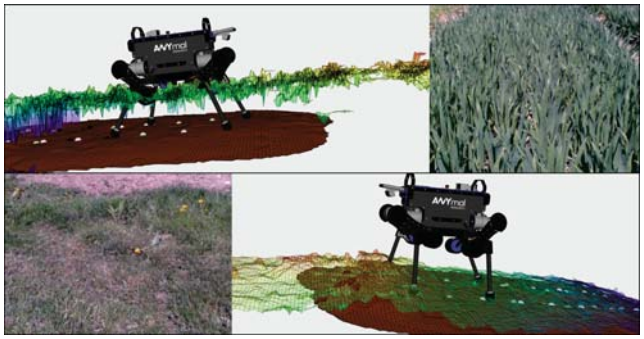


Fig. 5: The *visible topography* (colored) and the support surface estimation (brown) at two different instants during the first experiment run with corresponding camera images. Upper: Dense continuous vegetation, lower: Short grass.

vegetation height as well as its variance are close to zero. Because the *vegetation height* is close to zero, the *vegetation subtraction* layer is almost identical to the *visible topography* layer and high weight is placed on it due to the low V_{VH}, thereby letting the method closely mimic the behavior of the purely exteroceptive height mapping system.

2) Occlusion Adaptive Smoothing: In some rugged environments, e.g. high grass, objects close to the camera block the line of sight of the depth camera, causing a number of elevation map cells not to have a height value in the *visible topography* layer. As a consequence, for these cells the elevation of the *vegetation subtraction* layer cannot be calculated and rather than following the weighting procedure described in section II-F.1, $w_{FML} = 1$ is set.

Some areas of the *visible topography* may contain more blank spots than others. Therefore, an occlusion-adaptive GPR method is used, i.e. the lengthscale of the GPR applied in a model tile is adapted to the number of blank cells it contains:

$$l = \max\left(l_0 \cdot \frac{o_c}{o_{c,0}}, l_{min}\right), \tag{12}$$

where l is the adapted lengthscale, o_c is the number of blank cells and l_0, l_{min} as well as $o_{c,0}$ are tuning parameters.

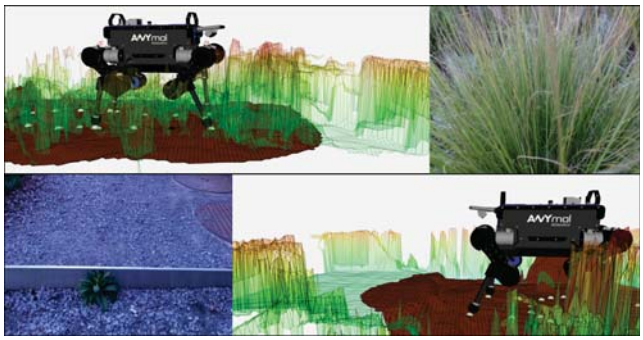


Fig. 6: The *visible topography* (colored) and the support surface estimation (brown) at two different instants during the the second experiment run with corresponding camera images. Upper: Dense vegetation, lower: Rigid ground with step, *visible topography* and support surface estimate are essentially identical (explanation in Section II-F.1).

III. Experiments

We verified our approach on ANYmal quadruped[4] by comparing the foothold height prediction performance of our support surface estimation method to a purely exteroceptive method and using the flat ground assumption as a baseline.

A. Experimental Setup

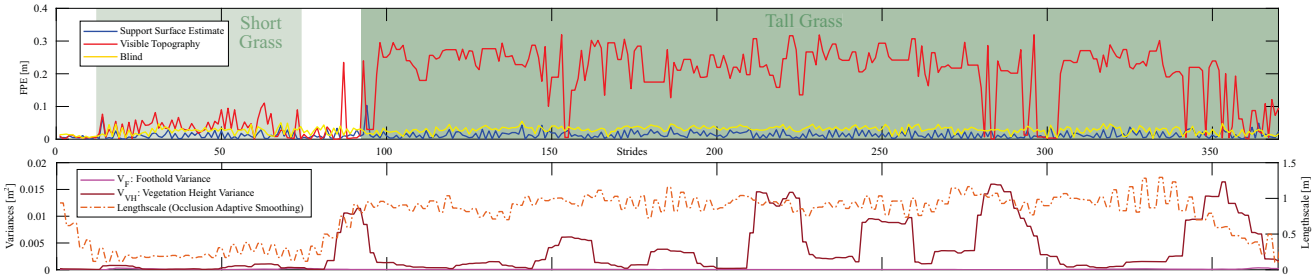
The quadruped was equipped with a Intel Realsense D435 depth camera. For the model tiling method of GPR (see section II-F) the tile diameter was set to $0.23\,\mathrm{m}$, the tile resolution was set to $0.08\,\mathrm{m}$. Experiments were performed in two locations, an agricultural field and an outdoor garden, which where not altered for the experiments. The vegetation consisted of bushes and grass, ranging from 0.1 to $0.8\,\mathrm{m}$ in height. In our first experiment, ANYmal took its first few steps on flat asphalt. It then transitioned to slightly inclined terrain, featuring well-kept, short grass, followed by dirt, before entering a field of young wheat plants, shown in Figure 1 (left). In this experiment, the *vegetation height* was fairly constant, with occasional brief clearings created by sectioning of the field.

Our second location was an unkept garden, featuring vegetation of highly varying height, shown in Figure 1 (right). Here, clearings are frequent and some plants grow taller than the robot itself. The robot was initialized standing inside the vegetation, walking for some time before exiting it and ascending a single, visible step, pictured in Figure 5. The path continued through vegetation and, after exiting, descended the same step again.

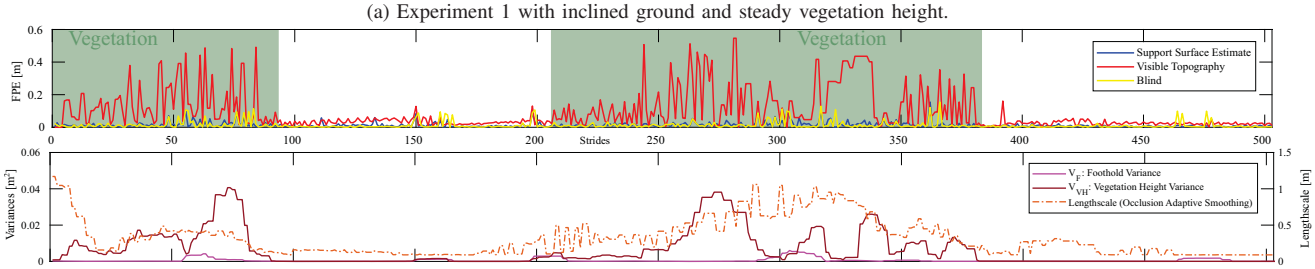
1) Results: As our error metric we define the foothold prediction error (FPE) as the height difference between the predicted foothold and the actual foothold. The predicted foothold height is the elevation of the respective layer at the position of the upcoming foothold.[7] It is compared to the FPE of 'blind' locomotion planning, i.e. assuming that consecutive footholds have the same height and thus not relying on the terrain model at all.

In experiment 1, for which result plots are shown in Figure 7a, all methods perform equally well during the first

[7]For the support surface estimation, FPE is calculated before the support surface estimation is updated, which occurs after the foot tip has entered stance phase.



(a) Experiment 1 with inclined ground and steady vegetation height.



(b) Experiment 2 with strongly varying vegetation and steps on observable ground (around strides 170 and 470).

Fig. 7: Foothold prediction error (FPE) plots. In the segments with green background the environment consists of vegetation, while in the segments with white background the ground is directly observable. V_F, V_{VH} are as assessed by the front right foot tip and are similar for the other foot tips.

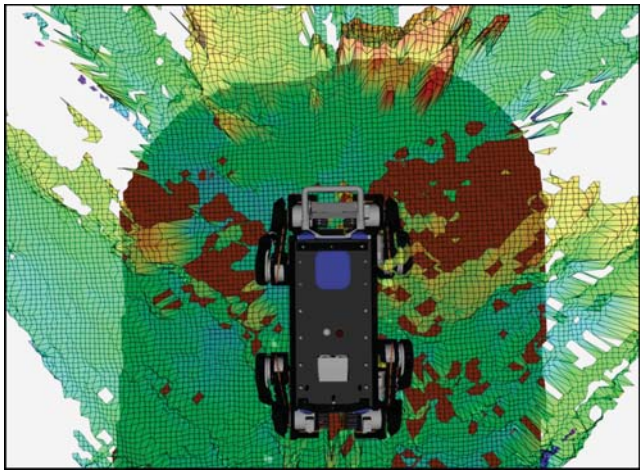


Fig. 8: The *visible topography* (colored) and the support surface estimation (brown) in vegetation during the second experiment. Top down view showing missing data in the *visible topography*, caused by shadowing.

15 steps on flat asphalt. On the inclined terrain, our method outperforms both other methods. The blind locomotion planner does not account for the inclination and therefore has a higher mean FPE. However, as the inclination is not very steep the FPE is of comparable order of magnitude. One notable exception is the spike of our method around stride 90. This occurs because, before entering the tall vegetation, the *vegetation height* is constant and small such that a step in the support surface is predicted. After taking one step into the vegetation our method notices the change in *vegetation height*, i.e. increased V_{VH}, and updates the estimate accordingly. As expected, the *visible topography* has constantly high FPE inside the vegetation.

Figure 7b shows the FPE for the second experiment run through strongly varying terrain, shown in Figure 6. As expected, while in vegetation, the FPE of our support surface estimate is significantly lower than the FPE of the *visible topography*. Where no vegetation is present, the FPE is similar for *visible topography* and support surface estimate with the latter being slightly lower. This is mostly owed to the fact that the proprioceptive measurements can somewhat correct for the accumulating state-estimator drift by updating the estimate underneath the robot. Performance of the blind locomotion planner is generally similar to our method because the experimental environment has mostly flat terrain. One important exception are the steps around stride 170 and 470 where our method as well as the *visible topography* predict correctly while the blind planner, by design, does not. Figure 8 shows a top down view of ANYmals perception in dense vegetation during the second experiment run. Although the *visible topography* is partially missing elevation data a continuous support surface estimate is provided. Generally in vegetation, relatively high *Lengthscale* is observed which yields increased gain of the occlusion adaptive smoothing method in these sections.

IV. Conclusion

In this work we presented a method to estimate the height of the support surface for robot locomotion with a combination of proprioceptive and exteroceptive measurements. It enables perceptive locomotion when the ground cannot be sensed with depth sensors directly, while still making accurate terrain geometry predictions for obstacle negotiation in unobstructed terrain. Even in the extreme case of a camera failure the system provides support surface estimates using proprioception only. In future work, proprioceptive sensing and hence terrain prediction could be further improved by tactile sensors in the feet to identify deformable terrain. Tighter integration with the locomotion system to perform active probing during sharp terrain transitions could increase robustness in these scenarios. Robustness in transitions could be furthermore improved by predicting the vegetation height from visual appearance and using it as a prior when stepping into vegetation.

REFERENCES

[1] P. Fankhauser, M. Bloesch, and M. Hutter, "Probabilistic terrain mapping for mobile robots with uncertain localization," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3019–3026, Oct 2018.

[2] F. Rogers-Marcovitz, N. Seegmiller, and A. Kelly, "Continuous vehicle slip model identification on changing terrains," in *RSS 2012 Workshop on Long-term Operation of Autonomous Robotic Systems in Changing Environments*, July 2012.

[3] D. Williamson, N. Kottege, and P. Moghadam, "Terrain characterisation and gait adaptation by a hexapod robot," in *Australasian conference on Robotics and Automation (ACRA)*, 2016.

[4] M. Hutter, C. Gehring, D. Jud, A. Lauber, C. D. Bellicoso, V. Tsounis, J. Hwangbo, K. Bodie, P. Fankhauser, M. Bloesch, R. Diethelm, S. Bachmann, A. Melzer, and M. Hoepflinger, "Anymal - a highly mobile and dynamic quadrupedal robot," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2016, pp. 38–44.

[5] P. Fankhauser, "Perceptive locomotion for legged robots in rough terrain," Ph.D. dissertation, ETH Zurich, 2018.

[6] Y. Junho, D. Ashwin, C. Soon-Jo, and H. Seth, "Vision-based localization and robot-centric mapping in riverine environments," *Journal of Field Robotics*, vol. 34, no. 3, pp. 429–450. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21606

[7] S. Thrun, "Robotic mapping: A survey," in *Exploring Artificial Intelligence in the New Millenium*. Morgan Kaufmann, 2002.

[8] P. Ross, A. English, D. Ball, B. Upcroft, and P. Corke, "Finding the ground hidden in the grass: Traversability estimation in vegetation."

[9] J. Z. Kolter, Y. Kim, and A. Y. Ng, "Stereo vision and terrain modeling for quadruped robots," in *2009 IEEE International Conference on Robotics and Automation*, May 2009, pp. 1557–1564.

[10] Z. Petrou, I. Manakos, T. Stathaki, C. Mcher, and M. Adamo, "Discrimination of vegetation height categories with passive satellite sensor imagery using texture analysis," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 4, pp. 1442–1455, April 2015.

[11] M. Bjelonic, N. Kottege, T. Homberger, P. Borges, P. Beckerle, and M. Chli, "Weaver: Hexapod robot for autonomous navigation on unstructured terrain," *Journal of Field Robotics*, vol. 0, no. 0. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.21795

[12] C. Wellington, A. Courville, and A. T. Stentz, "A generative model of terrain for autonomous navigation in vegetation," *The International Journal of Robotics Research*, vol. 25, no. 12, pp. 1287 – 1304, December 2006.

[13] J. Kim, D. Kim, J. Lee, J. Lee, H. Joo, and I. S. Kweon, "Non-contact terrain classification for autonomous mobile robot," in *2009 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Dec 2009, pp. 824–829.

[14] E. M. DuPont, C. A. Moore, and R. G. Roberts, "Terrain classification for mobile robots traveling at various speeds: An eigenspace manifold approach," in *2008 IEEE International Conference on Robotics and Automation*, May 2008, pp. 3284–3289.

[15] F. Schilling, X. Chen, J. Folkesson, and P. Jensfelt, "Geometric and visual terrain classification for autonomous mobile navigation," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sept 2017, pp. 2678–2684.

[16] S. Vasudevan, F. Ramos, E. Nettleton, H. Durrant-Whyte, and A. Blair, "Gaussian process modeling of large scale terrain," in *2009 IEEE International Conference on Robotics and Automation*, May 2009, pp. 1047–1053.

[17] P. Christian, M. Sebastian, P. Sam, K. Kristian, R. Nicholas, and B. Wolfram, "A bayesian regression approach to terrain mapping and an application to legged robot locomotion," *Journal of Field Robotics*, vol. 26, no. 10, pp. 789–811. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/rob.20308

[18] P. Fankhauser and M. Hutter, "A Universal Grid Map Library: Implementation and Use Case for Rough Terrain Navigation," in *Robot Operating System (ROS) The Complete Reference (Volume 1)*, A. Koubaa, Ed. Springer, 2016, ch. 5. [Online]. Available: http://www.springer.com/de/book/9783319260525

[19] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.