Cost of Transport Estimation for Legged Robot Based on Terrain Features Inference from Aerial Scan

Miloš Prágr Petr Čížek Jan Faigl

***Abstract*— The effectiveness of the robot locomotion can be measured using the cost of transport (CoT) which represents the amount of energy that is needed for traversing from one place to another. Terrains excerpt different mechanical properties when crawled by a multi-legged robot, and thus different values of the CoT. It is therefore desirable to estimate the CoT in advance and plan the robot motion accordingly. However, the CoT might not be known prior the robot deployment, e.g., in extraterrestrial missions; hence, a robot has to learn different terrains as it crawls through the environment incrementally. In this work, we focus on estimating the CoT from visual and geometrical data of the crawled terrain. A thorough analysis of different terrain descriptors within the context of incremental learning is presented to select the best performing approach. We report on the achieved results and experimental verification of the selected approaches with a real hexapod robot crawling over six different terrains.**

I. Introduction

Autonomous robots are being deployed in long-term data collection missions in environments with limited or no prior information about the particular terrain the robots are facing to, e.g., in extraterrestrial missions [1]. However, efficient locomotion over a particular terrain greatly influences the mission effectiveness. It is even more prominent with multi-legged robots due to their enhanced traversability capabilities to reside over terrains of different types.

Regarding locomotion of the particular robot, terrains can be distinguished by the traversability cost metric [2] that can be a simple binary division between passable and impassable terrains [3]. Alternatively, there are more elaborating scores such as the Cost of Transport (CoT) [4], [5], which represents a measure of the effectiveness of the robot locomotion. Therefore it is desirable to study the terrain traversability estimation to support mission planning and improve the real-time robot performance in accomplishing its mission goals.

Note, the CoT is inherently a continuous measure influenced by many factors, e.g., terramechanical properties of the terrain, robot morphology, and even seasonal and weather condition changes in long-term missions. Besides, terrains that the robot encounters might not be known in advance, and therefore, a simple classification using a set of pre-learned classes to estimate the CoT is not sufficient for a real-world deployment. Hence, self-learning mechanisms are necessary to estimate the CoT in yet untraversed areas correctly.

Authors are with the Department of Computer Science, Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic. {pragrmi1, petr.cizek, faiglj}@fel.cvut.cz

This work has been supported by the Czech Science Foundation (GAČR) under research project No. 18-18858S. The authors acknowledge the support of the OP VVV funded project CZ.02.1.01/0.0/0.0/16_019/0000765 "Research Center for Informatics".

(a) CoT estimates

(b) Colored height map (c) Multiple artificial terrains

Fig. 1: Multi-Terrain setup and its perceived representation.

In this work, we are concerning the traversability cost estimation for a small hexapod crawling robot using exteroceptive data. A relatively slow speed of crawling robots limits their capability to map a given area fully. Moreover, an easily obstructive close-to-ground viewpoint makes the mapping even more challenging. Nevertheless, a robot capable of traversing an unknown terrain may observe both the terrain appearance and terramechanical properties, and thus it can incrementally build the terrain model describing its CoT.

On the other hand, unmanned aerial vehicles (UAVs) do not suffer from these problems. From a relatively high altitude, a UAV can observe a larger portion of the terrain. Also, a typical UAV often moves faster than most of the ground vehicles of similar size. For these reasons, UAVs can be utilized for mapping unknown terrains, which in turn, may help the robot to improve the mission efficiency.

However, the ground crawling robot and the UAV observe the terrain with a different perspective, and therefore, it is not possible to simply infer the CoT over the terrain observed by the UAV based on the model learned by the ground robots if arbitrary descriptors are used. Hence, construction of such a model together with a selection of suitable feature set for the desired inference is the main goal of this work. In particular, we aim to develop incremental learning of the model to estimate the CoT based on the experience of the hexapod crawling robot with the terrain that can be further utilized for inference of the CoT using aerial terrain view. Such an inference may lead to the annotated aerial terrain cost map as it is visualized in Fig. 1. In this paper, we report on the achieved results towards this challenging goal which is addressed as a thorough experimental analysis of a set of terrain characterization features and learning methods used in the literature for terrain classification from visual and scene geometric data and application of these methods in the problem of the CoT learning and estimation.

The paper is structured as follows. An overview of the

related approaches on the terrain classification and terrain features is in Section II. A description of the proposed inference learning framework, terrain model, and features utilized in the herein reported evaluation is presented in Section III. Section IV details the experimental set-up and reports on the achieved results. Finally, conclusions are drawn in Section V.

II. Related work

Numerous methods on terrain traversability analysis [2], [6] have been presented in recent years that aim to evaluate the terrain properties mostly from geometric data and assign each area a number that characterizes local properties of the terrain geometry. On the other hand, traversability cost metrics [4], [5], like the Cost of Transport (CoT), are more general metrics that incorporate the own robot experience with traversing the terrain. The value of the CoT is inherently continuous. Even though it can be estimated from the results of the terrain classification, the classification relies on a discrete set of pre-learned classes which might not be available or might get irrelevant during the mission, e.g., in extraterrestrial environments or long-term missions. Therefore we are interested in the CoT regression.

The terrain description methods rely on extraction of terrain characterization features that can be roughly categorized into appearance-based visual features, geometric-based features, and methods combining both approaches. Moreover, the approaches can be further categorized based on whether the feature is dependent on the current robot viewpoint, and whether the feature makes use of color information. For example, approaches extracting features from images are inherently viewpoint dependent. In contrast, approaches using the extraction of features from point clouds or aerial scans are position independent. Further, the color information can be valuable for discriminating terrain types; however, it is strongly influenced by the illumination and seasonal changes.

The appearance-based features include approaches to classify terrains based on different colors or textures. The authors of [7] introduce classification based on a simple two-dimensional feature which uses mean color components of superpixels in the Lab color space combined with the SVM classifier. Bayes decision rules with Gaussian mixture models on the RGB color space is used in [8]. Regarding the texture recognition, methods using frequency-based approaches relying on wavelet filters [9], [10], Gabor filters [11], or more recent approach on Steerable Pyramid Masks [12] can be used. A survey on visual terrain classification from a monocular camera is presented in [13]. Further, hybrid approaches utilizing RGB-D camera [14] and stereo camera [15] use a combination of the appearance and geometrical features.

The main advantage of the geometrical terrain features is that they are not affected by illumination changes. However, they can suffer from a low density of the point cloud in a far distance to the robot which is limiting especially for a vast number of approaches that compute statistics based on normals [16]–[19] where point normals are mostly computed by fitting a plane to a local neighborhood of the investigated points. Point-cloud density [20], view-point, and point cloud centroid relations [18], [19] or minimal and maximal curvature of estimates for local neighborhood [21], or histogram based features [22] are being used. Features extracted from LIDAR data are used in [23] with a random forest classifier to differentiate vegetation and estimate the soil plane using the geometrical, reflectance, and color information.

In [24], a set of 13 features is proposed to describe terrain, vegetation, and other objects in an agricultural environment. The set is computed from the local neighborhood of the interesting point and divided into four height features based on the z coordinate, four shape features based on principal component analysis, three orientation features based on normal vectors of the local plane, a distance feature, and a reflectance feature. However, the classifier is trained by SVM from labeled data [24]. Moreover, the approach requires the z-coordinate of the point cloud to be orthogonal to the surface which is made by fitting a global ground plane to the dataset, which may bias the results in a more structured environment.

A self-supervised approach is presented in [25] to teach a terrain classifier from geometric data using the proprioceptive data. However, the approach uses a pre-learned proprioceptive classifier which differentiates only between several terrain classes.

The most similar approach to the herein addressed problem of the assignment of the traversability data to the aerial scan has been presented in [20]. The authors proposed a self-supervised learning approach with a Gaussian mixture model. The traversability cost is estimated from the geometrical clues in the LIDAR data to infer the traversability cost in a map obtained by an aerial recon.

Based on the presented literature survey, we have identified a set of appearance-based and geometric-based features with different properties that are commonly used in the terrain classification, and we adapt these for the CoT regression presented in this paper. Description of the used features is presented in the following section. Besides, we can conclude that the incremental regression and estimation of the traversability value for the observed but yet untraversed areas is still a largely unexplored topic.

III. Inference Learning Framework

The main goal of this paper is to report on a thorough analysis of terrain characterization features used in the inference of the CoT perceived by the hexapod crawling robot to the aerial scan. This section describes the used framework for the extraction of the terrain features and CoT learning, i.e., the utilized terrain features, the learning procedure, and the sampling strategy for the inference learning. The individual building blocks are described in the following sections.

A. Terrain Characterization Features

Based on the literature survey and preliminary results, we consider the following features for benchmarking. We select the features that are computationally inexpensive, so that can be utilized on various mobile robotic platforms. Moreover,

we only use descriptors that are viewpoint robust under the herein described conditions.

1) Appearance-based features: Two point cloud based color features and a texture recognition using wavelets [9] have been selected. For the color features, both the RGB and Lab color spaces have been considered with either channel values of the sampled point (denoted as Point in the reported results) or a channel mean of the points in a $r = 0.2$ m spherical neighborhood (denoted as Mean in the reported results). As the feature is purely based on color, it is robust to viewpoint changes; however, less to illumination changes. During the preliminary evaluation, the wavelet features exhibit low performance presumably due to a large viewpoint change between the robot and the aerial scan, and they have been left out of the comparison.

2) Geometric-based features: We use a modified version of the terrain feature sets presented in [24]. In particular, we have used 11 out of 13 features, namely the shape feature, height feature, orientation feature and all of them combined in a full feature, leaving out the reflectance and distance features as those do not suit our experimental setup. The ground plane and normal are estimated by fitting a plane to the $k = 5$ nearest neighbors of the sampled point. We consider the utilized geometric features to be viewpoint robust. As the coordinate frame is based on the global ground plane estimate, the height feature and shape features are robust to viewpoint changes under the assumption that the aerial scan captures the area with sufficient precision. The orientation feature robustness depends on the quality of the aerial scan, although different descriptor values are assigned to the terrains sloped in different directions. A spherical region with $r = 0.3$ m radius is considered when querying the neighborhood of the sampled point.

B. Learning algorithms

We have considered four approaches on top of the utilized terrain features that are capable of regression from which two of them support incremental online learning.

1) Support Vector Regression: (SVR) [26] is a maximum-margin regression algorithm, here utilized with the radial basis function kernel.

2) Regression Tree: which uses recursive partitioning with the depth $d = 5$.

3) Incremental Gaussian Mixture Network Model: (IGMN) [27], [28] is an online incremental learning approach, which creates and updates the Gaussian mixture model based on streamed data points. The IGMN allows a full prediction of the data point based on an incomplete input of any kind. We used our implementation of the Fast-IGMN that is an improvement of the IGMN presented in [28]. The Fast-IGMN improves the IGMN time complexity to $\mathcal{O}(NKD^2)$, where N is the number of data points, K is the number of components, and D is the data point dimensionality. Experimentally, we parametrized the IGMN with the $k = 10$ components, grace period $v_{\text{min}} = 100$, minimal accumulated posterior $sp_{\text{min}} = 3$, and scaling factor $\delta = 1$.

4) Hoeffding Tree: or Very Fast Decision Tree Learner (VFDT) [29] is an online incremental decision tree learning algorithm that utilizes Hoeffding bound to create the output asymptotically identical to that of the conventional learner. We used a slightly modified VFDT implementation of [30]. However, unlike the other utilized approaches, Hoeffding tree is used with a discrete number of classes, i.e., $k = 10$.

C. Sampling and Learning

This section explains how the above-selected terrain features and learning algorithms (forming building blocks) are combined in an inference learning framework that estimates the CoT in the aerial scan of the terrain. Our model has two major life stages: (i) the learning phase, when the robot learns the model based on the RGB-D input and pairs it with the CoT; and (ii) the inference phase, where the learned model is used to evaluate the terrain observed from the aerial scan.

The framework operates on individual datasets consisting of georeferenced RGB-D, i.e., the color RGB and depth images from the robot and georeferenced RGB-D aerial scan of the whole environment. Besides, the ground robot collects the power readings used for estimation of the CoT [5] as

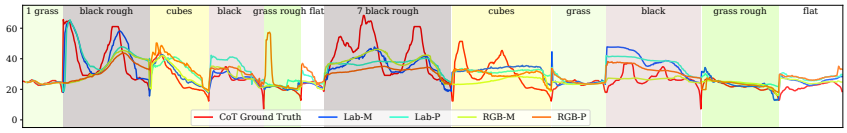
$$\text{CoT} = \frac{P}{m\, g\, v}, \tag{1}$$

where P is the instantaneous power consumption, m is the weight of the robot, $g = 9.81$ ms^{-2} is the gravitational acceleration, and v is the robot speed. In the regression task, we understand the CoT to be a function of the robot type, the robot gait, and the local terrain property. However, the experimental platform and locomotion gait are fixed; hence, the CoT is estimated only from the local terrain property. Note that the evaluated aerial scan is independent of the robot trajectory, i.e., the robot learned model can be applied to a different location; however, for the herein presented benchmarking, it is necessary that the trajectory of the robot is contained within the aerial scan.

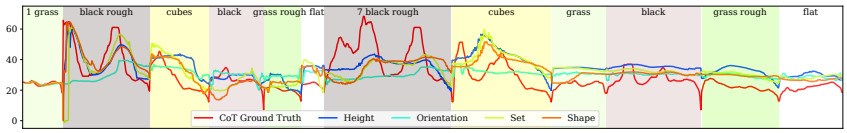
It is essential to address the fact that the robot knows the CoT only after it successfully traverses the terrain and estimates its velocity from the georeferenced data. Therefore, we introduce feature storage that maintains a dictionary of the georeferenced features extracted from the robot field of view to deal with this delay in the acquisition of the RGB-D–CoT pairs that are necessary for the learning phase. Thus, whenever an arbitrary location is reached by the robot, the feature storage is queried, and all features located sufficiently close, i.e., in a spherical region with $r = 0.2$ m, are passed to the learning framework together with the measured CoT of the current location. The temporary feature storage is of limited size, and it is randomly pruned when its capacity overflows. Hence, the robot is not creating a persistent feature map of the environment that would grow over time, and it incrementally learns the forthcoming terrain.

During the learning phase, points in front of the robot are sampled according to the scheme described above. The obtained features are then used for learning of the CoT model. In the case of the incremental learning, the features are first used to query the model for the value of the CoT.

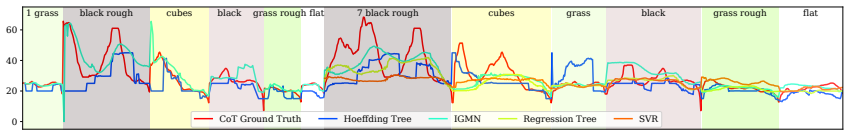
(a) The rough terrains.



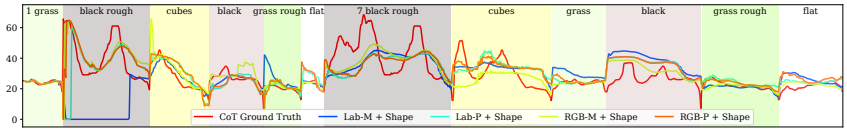
(b) CoT estimation for different variants of appearence-based features using the IGMN.



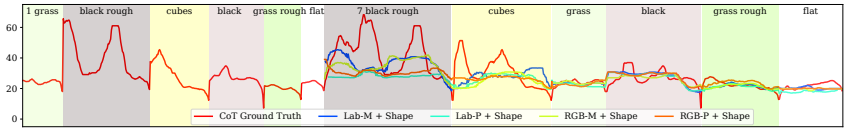
(c) CoT estimation for different variants of geometric-based features using the IGMN.



(d) Comparison of CoT estimation using the RGB-M-Shape feature and different learning algorithms.

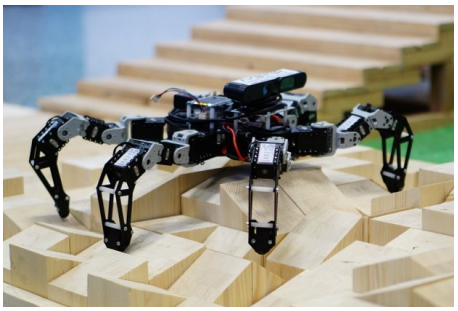


(e) Comparison of CoT estimation for different variants of combined appearence and geometric features using the IGMN.

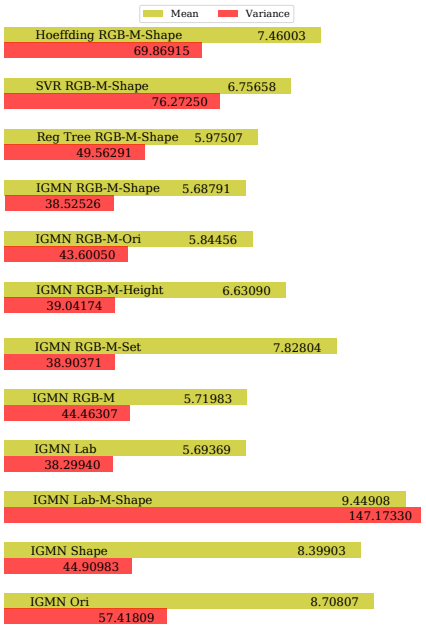


(f) Comparison of CoT estimation for different variants of combined appearence and geometric features using Regression Tree.

(b-f) Learned and predicted values of the CoT in time for different combination of features and learning algorithms. The first six terrains (parts) represent the learning phase. The following six parts represent the inference phase.



(g) The used hexapod robot.



(h) Mean absolute error and its variance of selected features. Variance bars not to scale.

Fig. 2: Experimental Hardware, Setup, and Results.

For the inference phase, the robot is virtually walked along the same known trajectory while sampling points from the aerial scan, which is necessary for the evaluation purposes. The robot stores the geolocated features in the same manner as in the learning phase. Similarly, when an arbitrary location is reached, the feature storage is queried, and all features located sufficiently close are considered. These features are then used to query the CoT model for the particular value of the CoT. Additionally, for the incremental learning enabled approaches, the model is further taught by these features combined with the measured CoT.

IV. Experimental Evaluation

In this section, we report on the experimental results and verification of the proposed inference learning framework with terrain features benchmarking using real hexapod crawling robot. The reported results are organized as follows. First, the robotic platform and the experimental set-up are introduced. After that, the results themselves are presented and discussed.

A. Hexapod Crawling Robot

The used robot is an electrically actuated low-cost hexapod crawling robot depicted in Fig. 2g. It has six legs, each with three joints attached to the trunk which hosts the electronics and sensory equipment. The RGB-D ASUS Xtion Pro Live camera has been utilized for the terrain perception and the Hall-effect-based current sensor for estimation of the robot instantaneous power consumption. The camera provides the data with 30 Hz frequency, and the power consumption data are provided with 62 Hz frequency. The locomotion over the rough terrain is performed by the adaptive motion gait [31], which uses the estimation of the ground-reaction forces based on the position data provided by the joint actuators.

B. Experimental Setup and Terrains

The experimental data have been obtained on the laboratory test-track consisting of three-meter length path over different surfaces. Six experiments, each with different terrain and three trials, have been performed. The terrains with

TABLE I: Error Rates for Individual Terrain Characterization Features

Learning	Feature		Error		Learning	Feature		Error		Learning	Feature		Error	
	Color	Geom.	Mean	Var		Color	Geom.	Mean	Var		Color	Geom.	Mean	Var
Hoeffding	RGB-M	Shape	7.4	69.8	IGMN	None	Ori	8.7	57.4	Reg Tree	Lab-P	None	6.4	37.9
Hoeffding	None	Height	7.4	56.7	IGMN	RGB-M	None	5.7	44.5	Reg Tree	Lab-P	Shape	7.1	70.9
Hoeffding	Lab-M	None	8.2	86.1	IGMN	RGB-M	Height	6.6	39.0	Reg Tree	None	Ori	8.1	60.3
Hoeffding	Lab-M	Shape	8.6	77.0	IGMN	RGB-M	Ori	5.8	43.6	Reg Tree	None	Set	7.9	59.0
Hoeffding	Lab-P	None	8.9	102.2	IGMN	RGB-M	Set	7.8	38.9	Reg Tree	None	Shape	6.9	77.2
Hoeffding	Lab-P	Shape	7.7	57.3	IGMN	RGB-M	Shape	**5.7**	**38.4**	SVR	RGB-M	Shape	6.8	76.3
Hoeffding	None	Ori	9.0	112.2	IGMN	RGB-P	None	7.3	59.4	SVR	None	Height	7.8	73.9
Hoeffding	RGB-M	None	8.1	70.1	IGMN	RGB-P	Shape	6.7	40.4	SVR	Lab-M	None	7.1	81.5
Hoeffding	RGB-M	Height	7.4	59.8	IGMN	None	Set	9.0	58.5	SVR	Lab-M	Shape	7.9	64.9
Hoeffding	RGB-M	Ori	7.9	66.0	IGMN	None	Shape	8.4	44.9	SVR	Lab-P	None	6.7	63.8
Hoeffding	RGB-M	Set	9.3	87.9	Reg Tree	RGB-M	None	6.3	51.5	SVR	Lab-P	Shape	7.8	68.7
Hoeffding	RGB-P	None	7.7	63.5	Reg Tree	RGB-M	Height	6.4	48.4	SVR	None	Ori	7.6	79.9
Hoeffding	RGB-P	Shape	8.3	72.5	Reg Tree	RGB-M	Ori	5.8	47.5	SVR	RGB-M	None	7.2	74.8
Hoeffding	None	Set	8.6	76.2	Reg Tree	RGB-M	Set	6.8	63.3	SVR	RGB-M	Height	7.7	74.0
Hoeffding	None	Shape	9.1	82.5	Reg Tree	RGB-M	Shape	5.9	49.6	SVR	RGB-M	Ori	7.4	77.2
IGMN	None	Height	9.0	37.9	Reg Tree	None	Height	7.3	58.6	SVR	RGB-M	Set	7.7	78.1
IGMN	Lab-M	None	**5.7**	**38.3**	Reg Tree	RGB-P	None	6.2	41.5	SVR	RGB-P	None	7.0	75.5
IGMN	Lab-M	Shape	9.4	147.2	Reg Tree	Lab-M	None	6.9	66.6	SVR	RGB-P	Shape	6.9	70.7
IGMN	Lab-P	None	7.5	39.6	Reg Tree	RGB-P	Shape	6.6	58.7	SVR	None	Set	7.7	78.3
IGMN	Lab-P	Shape	7.2	67.6	Reg Tree	Lab-M	Shape	6.1	46.0	SVR	None	Shape	7.3	73.7

the increasing difficulty of traversing are: PVC flooring (flat), turf-like carpet (grass), and semi-transparent soft black fabric (black) represent different flat terrains. Then, wooden blocks covered with the turf-like carpet (grass rough), wooden blocks covered with the black fabric (black rough), and bare wooden blocks (blocks) are considered as the rough terrain scenarios. The wooden blocks are 10×10 cm large with variable height and slope. The three rough terrain setups are shown in Fig. 2a. The same turf-like carpet and black-fabric have been used for the flat and the rough terrain setups.

The robot has been remotely guided over the course of the test-track while collecting visual and power consumption data. The visual data have been then processed using the incremental localization technique [32] to extract localization information that has been further used to estimate the robot velocity and for calculation of the CoT according to (1).

The collected data represent an unbiased belief of the robot about the traversability of the selected terrains. Besides, an aerial scan has been captured for each terrain type from the elevated camera to allow dense reconstruction of the whole track course simulating an aerial scan.

C. Results and Discussion

First, the individual trials over different terrains have been merged into a single pass dataset with six terrains each repeated two times as it is visualized in Figs. 2b-2f. Then, the proposed framework has been used according to the description presented in Section III-C. The algorithm has learned on the first six terrains, whereas the CoT value for the following terrains has been inferred from the aerial scan. The incremental learning approaches learn from, but also return the evaluation, for all the terrains, i.e., the terrains used in both the learning and inference phases.

The quantitative measures of the mean error between the predicted and ground-truth CoT, and its variance are reported for all tested combinations in Table I. Additionally, some of the results are visualized in Fig. 2h. Note that the measure is computed from all the returned values and for the incremental learning approaches, the metric includes results returned on the first six learning terrains.

The preliminary analysis has shown a low quality of the results provided by the Hoeffding trees and SVR learning algorithms. Therefore, the qualitative evaluation is focused on the best performing terrain features using the IGMN and Regression Trees. From the quantitative comparison, the best performing features are the sole appearance-based the LAB Mean feature and RGB Mean feature together with the geometric-based Shape feature.

A good performance of the sole Lab Mean feature (see Fig. 2b) is not surprising partially because of the experimental setup where the four well distinguishable colors appear on the terrains. Besides, only a little difference between the CoT values for three out of six terrains, namely flat, grass and grass rough, has been observed, which is correct behavior as we are not interested in the terrain classification, but rather in the CoT estimation. The CoT over the wooden blocks is less uniform, with low-cost areas being similar to the flat or grass datasets and high-cost peaks. Finally, the black and rough black datasets are the most costly with the high-cost peaks. Presumably, it is caused by the inability of the robot to find a proper grip on the fabric covered terrain. The made observations comply with the CoT map presented in Fig. 1, where the blocks have assigned a range of different costs, whereas the black fabric is assigned the high costs only.

From the further qualitative analysis, we can see that the standalone geometric features do not perform well (see Fig. 2c). However, the combination of the color and shape feature provides, in our opinion, the best results as the combination is able to better cope with the high peaks and low values of the CoT. In Table II, we present the aerial-ground scan correlation of the RGB Mean Shape feature. Although the individual feature dimensions do not exhibit a high correlation for all the terrains, for each terrain, there is at least one dimension with considerable correlation. The comparison of different models favors the IGMN setup in both quantitative and qualitative measures (see Fig. 2d,

TABLE II: Walk- and Environment-scan Correlation Rates for the RGB Mean Shape feature. For each of the examined terrain type pairs the metric is computed from 1000 randomly selected points. Each point is represented by the RGB-Means Shape descriptor computed from the environemnt scan, and the same descriptor computed from the walk scan. Then each dimension of the descriptor is reported separately, i.e., the correlation for one dimension in one terrain type is computed from 2 1000-length vectors. A median of the individual dimension correlations is also reported.

Terrain Type	Individual Feature Building Blocks							Full Feature
	Shape				RGB Means			Median
Grass Flat	0.62	0.40	0.34	0.66	0.66	0.47	0.64	0.62
Grass Rough	0.31	0.33	0.29	0.19	0.78	0.75	0.58	0.33
Black Flat	0.56	0.46	0.42	0.53	0.86	0.85	0.84	0.56
Black Rough	0.36	0.48	0.38	0.36	0.74	0.75	0.75	0.48
Flat	0.63	0.51	0.43	0.78	0.09	0.03	0.08	0.43
Cubes	0.55	0.34	0.31	0.47	0.79	0.80	0.78	0.55

Fig. 2e, and Fig. 2f). It is most likely due to its incremental learning property that allows the model to adapt quickly to CoT changes.

A rather interesting property of our datasets is that at the far end of each examined terrain, there is usually a section of a flat ground that has been traversed by the robot. Such a border represents a change of the terrain type. When investigating the recovered data, it is possible to observe that the most of the good performing setups are capable of reacting on such a terrain change and presume a lower value of the CoT in that region.

V. Conclusion

In this paper, we present a framework for model learning the CoT in a two-viewpoint setups, where the model is firstly learned by a small ground hexapod crawling robot, which can observe not only the exteroceptive terrain properties but also its associated CoT, and then the model is used for CoT inference from an aerial scan to yet untraversed areas. From a set of several feature setups, we chose the best performing combination of the RGB Mean and Shape feature which forms the descriptor of only seven dimensions. Several learning setups have been evaluated and based on the achieved results, the incremental Gaussian mixture and post-estimation regression trees suit best the selected features. In future, we aim to utilize this system in planning tasks and explore the transferability of the terrain traversability evaluation between different robotic platforms.

References

[1] D. Brown and G. Webster, "Now a stationary research platform, nasa's mars rover spirit starts a new chapter in red planet scientific studies," *NASA, press release*, 2010.

[2] P. Papadakis, "Terrain traversability analysis methods for unmanned ground vehicles: A survey," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 4, pp. 1373–1385, 2013.

[3] A. Stelzer, H. Hirschmüller, and M. Görner, "Stereo-vision-based navigation of a six-legged walking robot in unknown rough terrain," *The International Journal of Robotics Research*, vol. 31, no. 4, pp. 381–402, 2012.

[4] J. Nishii, "An analytical estimation of the energy cost for legged locomotion," *Journal of theoretical biology*, vol. 238, no. 3, pp. 636–645, 2006.

[5] N. Kottege, C. Parkinson, P. Moghadam, A. Elfes, and S. P. Singh, "Energetics-informed hexapod gait transitions across terrains," in *ICRA*, 2015, pp. 5140–5147.

[6] G. Reina, M. Bellone, L. Spedicato, and N. I. Giannoccaro, "3D traversability awareness for rough terrain mobile robots," *Sensor Review*, vol. 34, no. 2, pp. 220–232, 2014.

[7] K. Otsu, M. Ono, T. J. Fuchs, I. Baldwin, and T. Kubota, "Autonomous Terrain Classification with Co- and Self-Training Approach," in *IEEE Robotics and Automation Letters*, 2016, pp. 1–6.

[8] W. Mou and A. Kleiner, "Online Learning Terrain Classification for Adaptive Velocity Control," in *SSRR*, 2010, pp. 1–7.

[9] M. H. Bharati, J. J. Liu, and J. F. MacGregor, "Image texture analysis: Methods and comparisons," *Chemometrics and Intelligent Laboratory Systems*, vol. 72, no. 1, pp. 57–71, 2004.

[10] X. Li and Z. Tian, "Wavelet Energy Signature: Comparison and Analysis," pp. 474–480, 2006.

[11] S. Grigorescu, N. Petkov, and P. Kruizinga, "Comparison of texture features based on Gabor filters," *IEEE Transactions on Image Processing*, vol. 11, no. 10, pp. 1160–1167, 2002.

[12] S. Dash and U. R. Jena, "Texture classification using Steerable Pyramid based Laws' Masks," *Journal of Electrical Systems and Information Technology*, pp. 1–13, 2016.

[13] Y. Gao, C. Spiteri, M.-T. Pham, and S. Al-Milli, "A survey on recent object detection techniques useful for monocular vision-based planetary terrain classification," *Robotics and Autonomous Systems*, vol. 62, no. 2, pp. 151–167, 2014.

[14] S. Bartoszyk, P. Kasprzak, and D. Belter, "Terrain-aware motion planning for a walking robot," in *RoMoCo*, 2017, pp. 29–34.

[15] T. Homberger, M. Bjelonic, N. Kottege, and P. V. K. Borges, "Terrain-dependant Control of Hexapod Robots using Vision," in *International Symposium on Experimental Robotics*, 2016, pp. 92–102.

[16] R. B. Rusu, Z. C. Marton, N. Blodow, and M. Beetz, "Learning informative point classes for the acquisition of object model maps," *ICARCV*, pp. 643–650, 2008.

[17] R. B. Rusu, N. Blodow, and M. Beetz, "Fast Point Feature Histograms (FPFH) for 3D registration," *ICRA*, pp. 3212–3217, 2009.

[18] R. B. Rusu, G. Bradski, R. Thibaux, and J. Hsu, "Fast 3D recognition and pose using the viewpoint feature histogram," *IROS*, pp. 2155–2162, 2010.

[19] A. Aldoma, M. Vincze, N. Blodow, D. Gossow, S. Gedikli, R. B. Rusu, and G. Bradski, "CAD-model recognition and 6DOF pose estimation using 3D cues," in *ICCV*, 2011, pp. 585–592.

[20] S. Boris, E. Lin, J. A. Bagnell, J. Cole, N. Vandapel, and A. Stentz, "Improving Robot Navigation Through Self-Supervised Online Learning," *Journal of Field Robotics*, vol. 23, no. 11-12, pp. 1059–1075, 2006.

[21] Z. C. Marton, D. Pangercic, N. Blodow, J. Kleinehellefort, and M. Beetz, "General 3D modelling of novel objects from a single view," *IROS*, pp. 3700–3705, 2010.

[22] S. Lazebnik, C. Schmid, J. Ponce, S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine region," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1265–1278, 2010.

[23] D. M. Bradley, J. K. Chang, D. Silver, M. Powers, H. Herman, P. Rander, and A. Stentz, "Scene Understanding for a High-mobility Walking Robot," *IROS*, vol. 2015-Decem, pp. 1144–1151, 2015.

[24] M. Kragh, R. N. Jørgensen, and H. Pedersen, "Object Detection and Terrain Classification in Agricultural Fields Using 3D Lidar Data," in *ICVS*, vol. 9163, Copenhagen, 2015, pp. 188–197.

[25] C. A. Brooks and K. Iagnemma, "Self-supervised terrain classification for planetary surface exploration rovers," *Journal of Field Robotics (JFR)*, vol. 29, no. 3, pp. 445–468, 2012.

[26] C.-c. Chang and C.-j. Lin, "LIBSVM : A Library for Support Vector Machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, pp. 1–39, 2013.

[27] M. R. Heinen, P. M. Engel, and R. C. Pinto, "IGMN : An Incremental Gaussian Mixture Network that Learns Instantaneously from Data Flows," *Enia*, vol. ENIA, p. 12, 2011.

[28] R. C. Pinto and P. M. Engel, "A Fast Incremental Gaussian Mixture Model," *PLoS ONE*, vol. 10, no. 10, pp. e0 139 931+, 2015.

[29] P. Domingos and G. Hulten, "Mining high-speed data streams," in *Sixth ACM SIGKDD International Conference*, 2000, pp. 71–80.

[30] V. da Silva and A. T. Winck, "Video popularity prediction in data streams based on context-independent features," in *SAC*, 2017, pp. 95–100.

[31] J. Mrva and J. Faigl, "Tactile sensing with servo drives feedback only for blind hexapod walking robot," in *RoMoCo*, 2015, pp. 240–245.

[32] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.