Wild ANYmal: Learning robust perceptive locomotion

Takahiro Miki[1*], Joonho Lee[1], Jemin Hwangbo[2], Lorenz Wellhausen[1], Vladlen Koltun[3], Marco Hutter[1]

***Abstract*— Legged robots that can operate autonomously in remote and hazardous environments will greatly increase opportunities for exploration into under-explored areas. Exteroceptive perception is crucial for fast and energy-efficient locomotion: perceiving the terrain before making contact with it enables planning and adaptation of the gait ahead of time to maintain speed and stability. However, utilizing exteroceptive perception robustly for locomotion has remained a grand challenge in robotics. Snow, vegetation, and water visually appear as obstacles on which the robot cannot step – or are missing altogether due to high reflectance. Additionally, depth perception can degrade due to difficult lighting, dust, fog, reflective or transparent surfaces, sensor occlusion, and more. For this reason, the most robust and general solutions to legged locomotion to date rely solely on proprioception. This severely limits locomotion speed, because the robot has to physically feel out the terrain before adapting its gait accordingly. Here we present a robust and general solution to integrating exteroceptive and proprioceptive perception for legged locomotion. We leverage an attention-based recurrent encoder that integrates proprioceptive and exteroceptive input. The encoder is trained end-to-end and learns to seamlessly combine the different perception modalities without resorting to heuristics. The result is a legged locomotion controller with high robustness and speed.**

***Paper Type* – Recent Work [20].**

Fig. 1. Robust locomotion in the wild. The presented locomotion controller was extensively tested in a variety of complex environments over multiple seasons. The controller overcame a whole spectrum of real-world challenges, often encountering them in combination. These include slippery surfaces, steep inclinations, stairs, snow, and vegetation with degraded exteroception The controller traversed these environments with zero failures.

I. INTRODUCTION

Legged robots can carry out missions in challenging environments that are too far or too dangerous for humans, such as hazardous areas and the surfaces of other planets. Legs can walk over challenging terrain with steep slopes, steps, and gaps that may impede wheeled or tracked vehicles of similar size. There has been notable progress in legged robotics [24], [13], [10], [17], [22] and several commercial platforms are being deployed in the real world [3], [8], [1], [28]. However, until now, legged robots could not match the performance of animals in traversing challenging real-world terrain. Many legged animals such as humans and dogs can briskly walk or run in such environments by foreseeing the upcoming terrain and planning their footsteps based on visual information [19]. Animals naturally combine proprioception and exteroception to adapt to highly irregular terrain shape and surface properties such as slipperiness or softness, even when visual perception is limited. Endowing legged robots with this ability is a grand challenge in robotics.

One of the biggest difficulties lies in reliable interpretation of incomplete and noisy perception for control. Exteroceptive information provided by onboard sensors is incomplete and often unreliable in real-world environments. Generally, sensors which rely on light to infer distance are prone to producing artifacts on highly reflective surfaces, since the sensors assume that light travels in a straight path. In addition, depth sensors by nature cannot distinguish soft unstable surfaces such as vegetation from rigid ones.

Conventional approaches assume that the terrain information and any uncertainties encoded in the map are reasonably accurate, and the focus shifts solely to generating the motion. Offline methods use a pre-scanned terrain map, compute a handcrafted cost function over the map, and optimize a trajectory which is replayed on the robot [31], [21]. Online methods generally employ a similar approach but use only onboard resources to construct a map and continuously replan trajectories during execution [15], [9], [18], [2], [6], [12], [14], [29]. Overall, the focus of all the approaches mentioned above is on picking footholds and generating trajectories given accurate terrain information. Data-driven methods have recently been introduced in order to incor-

[1]Robotic Systems Lab, ETH Zürich, Zürich, Switzerland.
[2]Robotics & Artificial Intelligence Lab, KAIST, Daejeon, Korea.
[3]Intelligent Systems Lab, Intel, Jackson, WY, USA.
[*]Corresponding author; E-mail: tamiki@ethz.ch

porate more complex dynamics without compromising real-time performance and it has been applied to physical robot locomotion [10], [17], [27], [23], [30], [26], [16].

In both model-based and learning-based approaches, the assumption of flawless map quality precludes the application of these methods in uncontrolled outdoor environments. Existing controllers avoid catastrophic failures by simply refraining from using visual information in outdoor environments [13], [17], [26] or by adding heuristically defined reflex rules [7], [4].

Here we present a terrain-aware locomotion controller for quadrupedal robots that overcomes limitations of previous approaches and enables robust traversal of harsh natural terrain at high speeds (Fig 1). The key component is a recurrent encoder that combines proprioception and exteroception into an integrated belief state. It learns to take advantage of the foresight afforded by exteroception to plan footholds and accelerate locomotion when exteroception is reliable, and can seamlessly fall back to robust proprioceptive locomotion when needed. The learned controller thus combines the best of both worlds: the speed and efficiency afforded by exteroception and the robustness of proprioception.

Our contribution is a method for combining multi-modal perception and demonstrating with extensive hardware experiments that the resulting control policy is robust against various exteroceptive failures. Handling exteroception failures has been a challenging problem in robotics. Our approach constitutes a general framework for robust deployment of complex autonomous machines in the wild.

II. Materials and Methods

A. Overview

We train a neural network policy in simulation and then perform zero-shot sim-to-real transfer. Our method consists of three stages, illustrated in Fig. 2.

First, a teacher policy is trained with Reinforcement Learning (RL) to follow a random target velocity over randomly generated terrain with random disturbances with access to privileged information. Then, a student policy is trained to reproduce the teacher policy's actions without using this privileged information. Lastly, we transfer the learned student policy to the physical robot and deploy it in the real world with onboard sensors. The robot constructs an elevation map by integrating depth data from onboard sensors, and samples height readings from it.

B. Training environment

We use RaiSim [11] as our simulator to build the training environment. There, we simulate multiple ANYmal-C robots on randomly generated rough terrain in parallel with an integrated actuator model [10] to close the reality gap. We randomize the masses of the robot's body and legs, the initial joint position and velocity, and the initial body orientation and velocity in each episode. In addition, external force and torque are applied to the body of the robot and the friction coefficients of the feet are occasionally set to a low value to introduce slippage.

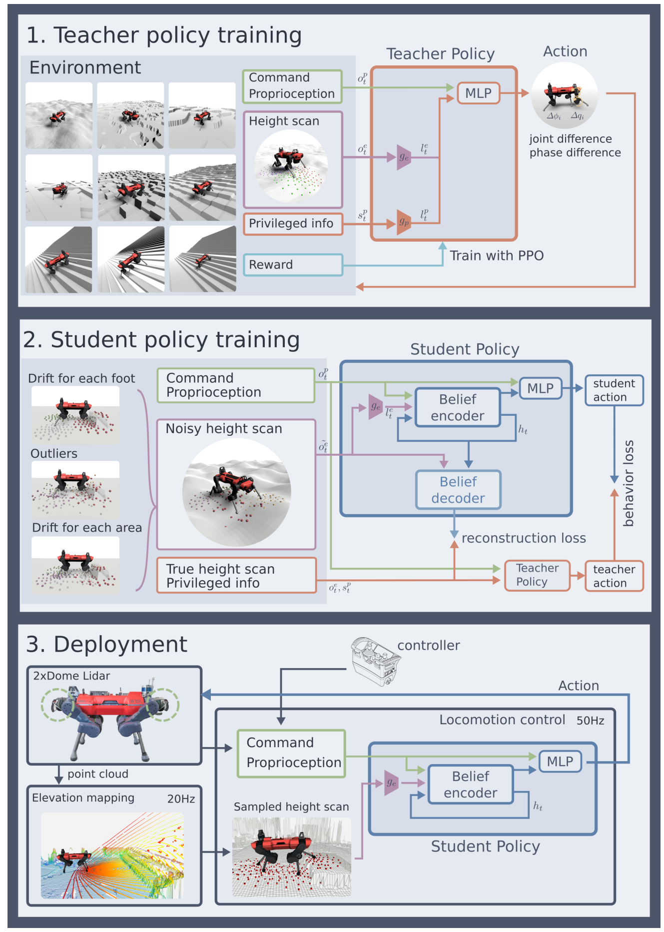

Fig. 2. Overview of the training methods and deployment. We first train a teacher policy with access to privileged simulation data using reinforcement learning (RL). This teacher policy is then distilled into a student policy, which is trained to imitate the teacher's actions and to reconstruct the ground-truth environment state from noisy observations. We deploy the student policy zero-shot on real hardware using height samples from a robot-centric elevation map.

C. Teacher policy training

In the first stage of training we aim to find an optimal reference control policy which has access to perfect, privileged information and enables ANYmal to follow a desired command velocity over randomly generated terrain. We used Proximal Policy Optimization (PPO) [25] to train the teacher policy. The teacher observation consists with proprioceptive observation o_t^p, exteroceptive observation o_t^e, and privileged state s_t^p. o_t^p contains the body velocity, orientation, joint position and velocity history, action history, and each leg's phase. o_t^e is a vector of height samples around each foot with five different radii. The privileged state s_t^p includes contact states, contact forces, contact normals, friction coefficient, thigh and shank contact states, external forces and torques applied to the body, and swing phase duration. Our action space is inspired by central pattern generators (CPGs) which keeps phase variables per each leg and defines a nominal trajectory based on the phase [17]. The action from the policy is the phase difference $\Delta\phi_l$ and the residual joint position target Δq_i. We model the teacher policy π_θ as an MLP. It consists of three MLP components: exteroceptive encoder, privileged encoder, and the main network, as shown in Fig. 2. We define a positive reward for following the command velocity and a negative reward for violating some imposed constraints. Please refer to [20] for details.

D. Student policy training

After we train a teacher policy, we distill it into a student policy that only has access to information that is available on the real robot. We use the same training environment as for the teacher policy, but add additional noise to the student height sample observation. The student policy consists of a recurrent belief state encoder and an MLP, as shown in Fig. 2.2. The MLP structure remains the same as for the teacher policy and reuse the weight of teacher policy to speed up training. Training is performed in supervised fashion by minimizing two losses: a behavior cloning loss and a reconstruction loss. The behavior cloning loss is defined as the squared distance between the student action and the teacher action. The reconstruction loss is the squared distance between the noiseless height sample and privileged information (o_t^e, s_t^p) and their reconstruction from the belief state. During student training, we inject random noise into the height samples as shown in Fig. 2.2.

E. Belief state encoder

To integrate proprioceptive and exteroceptive data, we introduce a gated encoder. The encoder learns an adaptive gating factor that controls how much exteroceptive information to pass through. First, proprioception o_t^p, exteroception $l_t^e = g_e(\tilde{o_t^e})$, and hidden state s_t are encoded by the Recurrent Neural Network (RNN) module into the intermediate belief state b'_t. Then, the attention vector α is computed from b'_t.

$$\begin{aligned} b'_t, h_{t+1} &= \mathrm{RNN}(o_t^p, l_t^e, h_t) \\ \alpha &= \sigma(g_a(b'_t)) \\ b_t &= g_b(b'_t) + l_t^e \odot \alpha \end{aligned}$$

Here, g_a and g_b are fully-connected neural networks and $\sigma(\cdot)$ is the sigmoid function. We use the Gated Recurrent Unit (GRU) [5] as our RNN architecture.

III. Results

We deployed our controller in a wide variety of terrain, as shown in Fig. 1. This includes alpine, forest, underground, and urban environments[1]. The controller was consistently robust and had zero falls during all deployments. Because of the exteroceptive perception, the robot could anticipate the terrain and adapt its motion to achieve fast and smooth walking. This was particularly notable for structures that require high foot clearance, such as stairs and large obstacles. The robot was able to leverage exteroceptive input to conquer terrain that was beyond the capabilities of prior work that did not utilize exteroception [17]. ANYmal successfully

Fig. 3. Our locomotion controller perceives the environment through height samples (red dots) from an elevation map (A). The controller is robust to many perception challenges commonly encountered in the field: missing map information due to sensing failure (B, C, G) and misleading map information due to non-rigid terrain (D, E) and pose estimation drift (F).

traversed challenging natural environments with steep inclination, slippery surfaces, grass, and snow (Fig. 1 A-J). The robot was robust in these conditions, even when occlusion and surface properties such as high reflectance impeded exteroception. Our controller was also robustly deployed in underground environments with loose gravel, sand, dust, water, and limited illumination (Fig. 1 K-N). Urban environments also present important challenges (Fig. 1 O-R).

Throughout our experiments, we encountered many circumstances in which exteroception provides incomplete or misleading input. As shown in Fig. 3 B-G, the estimated elevation map can unreliable due to sensing failures, limitations of the 2.5D height map representation, or viewpoint restrictions due to onboard sensing. Overall, our controller could handle all of these challenging conditions gracefully, without a single failure. The belief state estimator was trained to assess the reliability of exteroceptive information and made use of it to the extent possible. When exteroceptive information was incomplete, noisy, or misleading, the controller could always gracefully degrade to proprioceptive locomotion.

IV. Conclusion

We have presented a fast and robust quadrupedal locomotion controller for challenging terrain. The controller seamlessly integrates exteroceptive and proprioceptive input. Exteroceptive perception enables the robot to traverse the environment quickly and gracefully by anticipating the terrain and adapting its gait accordingly before contact is made. When exteroceptive perception is misleading, incomplete, or missing altogether, the controller smoothly transitions to proprioceptive locomotion. The controller remains robust in all conditions, including when the robot is effectively

[1] Video: `https://youtu.be/zXbb6KQ0xV8`

blind. The integration of exteroceptive and proprioceptive inputs is learned end-to-end and does not require any hand-coded rules or heuristics. The result is the first rough-terrain legged locomotion controller that combines the speed and grace of vision-based locomotion with the high robustness of proprioception.

Future work could explicitly utilize the uncertainty information in the belief state. Explicitly estimating uncertainty may allow the policy to become more careful when exteroceptive input is unreliable, for example using its foot to probe the ground if it is unsure about it. In addition, our current implementation obtains perceptual information through an intermediate state in the form of an elevation map, rather than directly ingesting raw sensor data. Appropriately folding the processing of raw sensory input into the network may further enhance the speed and robustness of the controller. Another limitation is the inability to complete locomotion tasks which would require maneuvers very different from normal walking, for example recovering from a leg stuck in narrow holes or climbing onto high ledges.

REFERENCES

[1] Agility Robotics. Robots. `https://www.agilityrobotics.com/robots`, 2021. [Online; accessed June-2021].

[2] Dominik Belter, Przemysław Łabecki, and Piotr Skrzypczyński. Adaptive motion planning for autonomous rough terrain traversal with a walking robot. *Journal of Field Robotics*, 33(3):337–370, 2016.

[3] Boston Dynamics. Spot. `https://www.bostondynamics.com/spot`, 2021. [Online; accessed March-2021].

[4] Boston Dynamics. Spot user guide release 2.0 version A. `https://www.generationrobots.com/media/spot-boston-dynamics/spot-user-guide-r2.0-va.pdf`, 2021. [Online; accessed June-2021].

[5] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. In *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, page 1724–1734, 2014.

[6] Péter Fankhauser, Marko Bjelonic, C Dario Bellicoso, Takahiro Miki, and Marco Hutter. Robust rough-terrain locomotion with a quadrupedal robot. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5761–5768. IEEE, 2018.

[7] Michele Focchi, Romeo Orsolino, Marco Camurri, Victor Barasuol, Carlos Mastalli, Darwin G Caldwell, and Claudio Semini. Heuristic planning for rough terrain locomotion in presence of external disturbances and variable perception quality. In *Advances in Robotics Research: From Lab to Market*, pages 165–209. Springer, 2020.

[8] Christian Gehring, Péter Fankhauser, Linus Isler, Remo Diethelm, Samuel Bachmann, Marcel Potz, Lars Gerstenberg, and Marco Hutter. ANYmal in the field: Solving industrial inspection of an offshore HVDC platform with a quadrupedal robot. In *Field and Service Robotics*, pages 247–260. Springer, 2021.

[9] Ioannis Havoutis, Jesus Ortiz, Stephane Bazeille, Victor Barasuol, Claudio Semini, and Darwin G Caldwell. Onboard perception-based trotting and crawling with the hydraulic quadruped robot (HyQ). In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 6052–6057. IEEE, 2013.

[10] Jemin Hwangbo, Joonho Lee, Alexey Dosovitskiy, Dario Bellicoso, Vassilios Tsounis, Vladlen Koltun, and Marco Hutter. Learning agile and dynamic motor skills for legged robots. *Science Robotics*, 4(26), 2019.

[11] Jemin Hwangbo, Joonho Lee, and Marco Hutter. Per-contact iteration method for solving contact dynamics. *IEEE Robotics and Automation Letters*, 3(2):895–902, 2018.

[12] Fabian Jenelten, Takahiro Miki, Aravind E Vijayan, Marko Bjelonic, and Marco Hutter. Perceptive locomotion in rough terrain–online foothold optimization. *IEEE Robotics and Automation Letters*, 5(4):5370–5376, 2020.

[13] Benjamin Katz, Jared Di Carlo, and Sangbae Kim. Mini cheetah: A platform for pushing the limits of dynamic quadruped control. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 6295–6301. IEEE, 2019.

[14] D Kim, D Carballo, J Di Carlo, B Katz, G Bledt, B Lim, and Sangbae Kim. Vision aided dynamic exploration of unstructured terrain with a small-scale quadruped robot. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2464–2470. IEEE, 2020.

[15] J Zico Kolter, Youngjun Kim, and Andrew Y Ng. Stereo vision and terrain modeling for quadruped robots. In *2009 IEEE International Conference on Robotics and Automation*, pages 1557–1564. IEEE, 2009.

[16] Ashish Kumar, Zipeng Fu, Deepak Pathak, and Jitendra Malik. Rma: Rapid motor adaptation for legged robots. In *Proceedings of Robotics: Science and Systems (RSS)*, 2021.

[17] Joonho Lee, Jemin Hwangbo, Lorenz Wellhausen, Vladlen Koltun, and Marco Hutter. Learning quadrupedal locomotion over challenging terrain. *Science Robotics*, 5(47), 2020.

[18] Carlos Mastalli, Michele Focchi, Ioannis Havoutis, Andreea Radulescu, Sylvain Calinon, Jonas Buchli, Darwin G Caldwell, and Claudio Semini. Trajectory and foothold optimization using low-dimensional models for rough terrain locomotion. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1096–1103. IEEE, 2017.

[19] Jonathan Samir Matthis, Jacob L Yates, and Mary M Hayhoe. Gaze and the control of foot placement when walking in natural terrain. *Current Biology*, 28(8):1224–1233, 2018.

[20] Takahiro Miki, Joonho Lee, Jemin Hwangbo, Lorenz Wellhausen, Vladlen Koltun, and Marco Hutter. Learning robust perceptive locomotion for quadrupedal robots in the wild. *Science Robotics*, 7(62):eabk2822, 2022.

[21] Peter D Neuhaus, Jerry E Pratt, and Matthew J Johnson. Comprehensive summary of the institute for human and machine cognition's experience with LittleDog. *The International Journal of Robotics Research*, 30(2):216–235, 2011.

[22] Hae-Won Park, Patrick M Wensing, and Sangbae Kim. Jumping over obstacles with MIT Cheetah 2. *Robotics and Autonomous Systems*, 136:103703, 2021.

[23] Xue Bin Peng, Erwin Coumans, Tingnan Zhang, Tsang-Wei Edward Lee, Jie Tan, and Sergey Levine. Learning agile robotic locomotion skills by imitating animals. In *Robotics: Science and Systems*, 07 2020.

[24] Marc Raibert, Kevin Blankespoor, Gabriel Nelson, and Rob Playter. Bigdog, the rough-terrain quadruped robot. *IFAC Proceedings Volumes*, 41(2):10822–10825, 2008.

[25] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[26] Jonah Siekmann, Kevin Green, John Warila, Alan Fern, and Jonathan Hurst. Blind bipedal stair traversal via sim-to-real reinforcement learning. In *Robotics: Science and Systems*, 2021.

[27] Jie Tan, Tingnan Zhang, Erwin Coumans, Atil Iscen, Yunfei Bai, Danijar Hafner, Steven Bohez, and Vincent Vanhoucke. Sim-to-real: Learning agile locomotion for quadruped robots. In *Robotics: Science and Systems*, 2018.

[28] Unitree Robotics. A1. `https://www.unitree.com/products/a1/`, 2021. [Online; accessed March-2021].

[29] Octavio Antonio Villarreal-Magaña, Victor Barasuol, Marco Camurri, Michele Focchi, Luca Franceschi, Massimiliano Pontil, Darwin G. Caldwell, and Claudio Semini. Fast and continuous foothold adaptation for dynamic locomotion through cnns. *IEEE Robotics and Automation Letters*, 4(2):2140–2147, 2019.

[30] Zhaoming Xie, Patrick Clary, Jeremy Dao, Pedro Morais, Jonanthan Hurst, and Michiel van de Panne. Learning locomotion skills for cassie: Iterative design and sim-to-real. In Leslie Pack Kaelbling, Danica Kragic, and Komei Sugiura, editors, *Proceedings of the Conference on Robot Learning*, volume 100 of *Proceedings of Machine Learning Research*, pages 317–329. PMLR, 30 Oct–01 Nov 2020.

[31] Matt Zucker, J Andrew Bagnell, Christopher G Atkeson, and James Kuffner. An optimization approach to rough terrain locomotion. In *2010 IEEE International Conference on Robotics and Automation*, pages 3589–3595. IEEE, 2010.