Preview Optimization for Learning Locomotion Policies on Rough Terrain

Carlos Mastalli[1], Ioannis Havoutis[2], Andreea Radulescu[1],
Michele Focchi[1], Darwin G. Caldwell[1], Claudio Semini[1]

***Abstract*— Legged robots promise a clear advantage in unstructured and challenging terrain, scenarios such as disaster relief, search and rescue, forestry and construction site. Dynamic locomotion on rough terrain has to guarantee stability and maximizing the *cross-ability* of a local set of candidate footholds. Trajectory optimization improves such performance metric while satisfying locomotion stability. Terrain conditions increase significantly the dimensionality of the optimization problem. Moreover, decoupling footstep selection and Center of Mass (CoM) motion generation may limit the success of the task. We are inspired by the observation that humans solve complex problems through intensive reasoning in the initial phases, which allows them to solve faster and naturally similar problems. In the same vein, the preview optimization allows the robot to infer the locomotion skills required on challenging terrain, and then use the data to build a locomotion policy that can be the used in real-time. A set of preview model allows us to reduce the dimensionality of the problem, which is desirable for trajectory optimization and policy reconstruction.**

I. MOTIVATION

One promising application domain for legged systems is search and rescue. Such environments can be dangerous for humans and inaccessible to wheeled vehicles. From the legged locomotion point of view, natural disaster scenarios require planning and execution of complex behaviours in challenging environments (see Fig. 1). Zero Moment Point (ZMP)-based approaches have been proposed for quadrupedal locomotion on rough terrain [1][2]. Those approaches decouple the footstep selection and CoM generation problems [3][4]. Moreover, they assumed a fixed step duration which allows posing the CoM trajectory generation as a Quadratic Programming (QP) problem. In rough terrain, it can limited the set of possible footsteps. For instance in those approaches, the footsteps are chosen based on robot kinematic, and fixed step duration limits the step variability and transitions between different gaits.

II. PROBLEM OVERVIEW

Preview models are low-dimensional representations that describe and capture different locomotion behaviours such as walking and trotting. Reducing the dimensionality of the optimization problem helps to generate complex locomotion behaviours and their transitions. Recently, it has been proposed a preview optimization approach for bipedal locomotion on an animated character [5]. In this approach, they defined a simple preview schedule that allows the character to generate three kind of behaviours: standing, walking and running. Extending this technique to quadrupedal locomotion on rough terrain is challenge. For instance the number of legs, transitions between phases and the combination of behaviours is higher. Moreover the adaptation of these preview models and terrain information for on-line motion generation have never been done. Note that the terrain information increases significantly the dimensionality of the optimization problem, making it hard to solve on-line. Thus, a policy reconstruction of the optimized preview control sequence may allow the robot to re-plan in real-time, which improves the robustness of the locomotion.

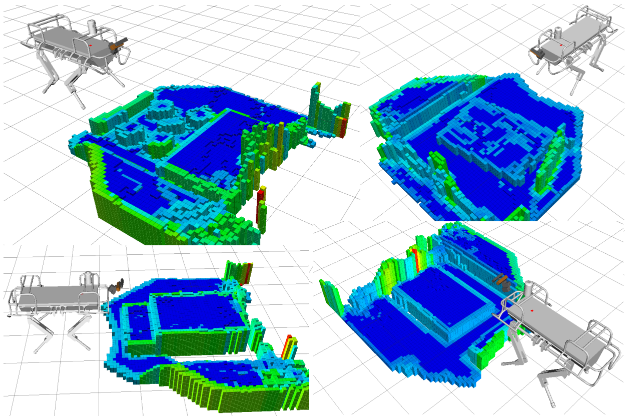

Fig. 1. Different challenging terrains that can be encountered on natural distance scenarios. Top left: stepping stones; top right: pallet; bottom left: stair; bottom right: gap.

III. PRELIMINARY RESULTS

Describing the quadrupedal locomotion can be done through a sequence of different preview models or phases, i.e. preview schedule, such as: stance and flight models. The preview schedule allows the robot to generate different patterns of locomotion such as walking, trotting, and their transitions (see Fig. 2). For the stance phase, a cart-table model can be used to describe the balancing dynamics. Decoupling and linearising the cart-table model of [6] into the horizontal CoM motion allows us to derive the following Equation of Motions (EoM):

$$\mathbf{x}(t) = \boldsymbol{\beta}_1 e^{\alpha t} + \boldsymbol{\beta}_2 e^{-\alpha t} + \frac{t}{T}(\mathbf{p}_T - \mathbf{p}_0) + \mathbf{p}_0 \qquad (1)$$

---

[1]Department of Advanced Robotics, Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova, Italy. *email*: {carlos.mastalli, andreea.radulescu, michele.focchi, darwin.caldwell, claudio.semini}@iit.it.

[2]Robot Learning and Interaction Group, Idiap Research Institute, Martigny, Switzerland *email*: ioannis.havoutis@idiap.ch

This work was in part supported by the DexROV project through the EC Horizon 2020 programme (Grant #635491).

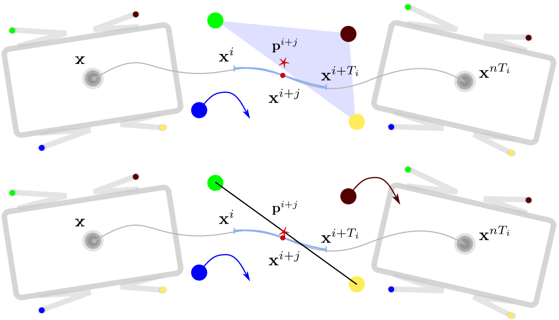

Fig. 2. Different quadrupedal gaits can be described using a sequence of preview phases. In every preview phase, a cart-table model allows us to capture the balancing dynamic when stability constraints are imposed. Depends on the gait, those constraints can be posed as inequality or equality constraints.

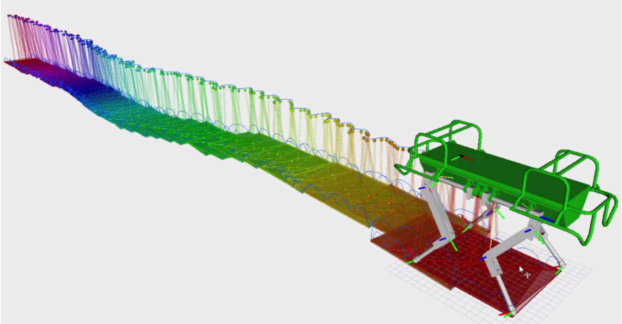

Fig. 3. The optimized trajectory from a sequence of 52 preview optimization problems for flat terrain. The color describes different preview phases of the optimization problems, where the spheres represent CoP and CoM positions.

where the model coefficients depend on duration of the phase T, and Center of Pressure (CoP) displacement $\delta\mathbf{p} = (\mathbf{p}_T - \mathbf{p}_0)$. Additionally, the parameters are mass of the pendulum m, height of the pendulum h, and gravity acceleration g. We assume that the base orientation is independent of the CoM motion. We can control the base yaw orientation with a constant angular acceleration $\ddot{\alpha}$

$$\alpha(t) = \alpha_0 + \dot{\alpha}_0 t + \frac{1}{2}\ddot{\alpha}t^2 \tag{2}$$

In the preview schedule, we build up a sequence of control parameters that describe the locomotion:

$$\mathbf{U} = \begin{bmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_n & \mathbf{F}\end{bmatrix} \tag{3}$$

where $\mathbf{u}_i = \begin{bmatrix}T & \delta\mathbf{p} & \ddot{\alpha}\end{bmatrix}$ and $\mathbf{F}$ the foothold targets.

We find the sequence of control parameters $\mathbf{U}$, through an unconstrained optimization problem, given the initial state $\mathbf{s}_0$ and desired user commands (step length and duration)

$$\mathbf{U}^* = \underset{\mathbf{U}}{\operatorname{argmin}} \sum_i \omega_i g_i(\mathbf{S}(\mathbf{U})) \tag{4}$$

where $\mathbf{S} = \begin{bmatrix}\mathbf{s}_1 & \cdots & \mathbf{s}_n\end{bmatrix}$ is a sequence of preview states. The preview state is defined by the CoM position and velocity $(\mathbf{x}, \dot{\mathbf{x}})$, CoP position $\mathbf{p}$ and the stance support region $\mathbf{c}$, i.e. $\mathbf{s} = \begin{bmatrix}\mathbf{x} & \dot{\mathbf{x}} & \mathbf{p} & \mathbf{c}\end{bmatrix}$. The cost functions $g_i(\mathbf{S})$ describes:

- User command tracking (step length and duration)
- CoM energy
- Stability soft-constraint, i.e. CoP inside support polygon
- Preview model soft-constraint, i.e. pendulum length

Fig. 3 shows a generated trajectory computed from the optimized preview control sequence. This trajectory is generated from a sequence of 52 preview optimization problems, and it is computed off-line. For every problem, we optimize a 2-cycle[1] movement and prune the second one, i.e. in Model Predictive Control (MPC) fashion. As in MPC, the main advantage is the fact that it optimizes the current timeslot while keeping future timeslots in account. According to our preliminary results, it helps to discover smoother transitions between locomotion cycles. We strongly believe that considering future timeslots will become more important on rough terrain conditions. For instances, it may generate desired anticipated behaviours given the terrain condition of the next locomotion cycle.

IV. FUTURE WORK

A terrain cost map quantifies how desirable it is to place a foot at a specific location. The terrain cost c_t for each voxel in the map is computed using geometric terrain features as [4]. In future work we will integrate a local terrain cost map into our preview optimization problem (Eq. (4)). Then we will develop a suitable policy reconstruction technique for quadrupedal locomotion on rough terrain.

ACKNOWLEDGEMENTS

This research is funded by the Fondazione Istituto Italiano di Tecnologia.

REFERENCES

[1] M. Kalakrishnan, J. Buchli, P. Pastor, M. Mistry, and S. Schaal, “Learning, planning, and control for quadruped locomotion over challenging terrain,” *The International Journal of Robotics Research (IJRR)*, vol. 30, no. 2, pp. 236–258, 11 2010,

[2] A. Winkler, C. Mastalli, I. Havoutis, M. Focchi, D. G. Caldwell, and C. Semini, “Planning and Execution of Dynamic Whole-Body Locomotion for a Hydraulic Quadruped on Challenging Terrain,” in *IEEE International Conference on Robotics and Automation (ICRA)*, May 2015,

[3] J. Z. Kolter, M. P. Rodgers, and A. Y. Ng, “A control architecture for quadruped locomotion over rough terrain,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2008, pp. 811–818,

[4] C. Mastalli, A. Winkler, I. Havoutis, D. G. Caldwell, and C. Semini, “On-line and On-board Planning and Perception for Quadrupedal Locomotion,” in *IEEE International Conference on Technologies for Practical Robot Applications (TEPRA)*, May 2015,

[5] I. Mordatch, M. de Lasa, and A. Hertzmann, “Robust physics-based locomotion using low-dimensional planning,” *ACM Transactions on Graphics*, vol. 29, no. 4, p. 1, July 2010,

[6] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, and H. Hirukawa, “Biped walking pattern generation by using preview control of zero-moment point,” in *IEEE International Conference on Robotics and Automation*, 2003, pp. 1620–1626,

---

[1] A quadrupedal locomotion cycle happens when the robot swings its 4 legs