Fakultät für Physik und Astronomie

Ruprecht-Karls-Universität Heidelberg

Masterarbeit

Im Studiengang Physik

vorgelegt von

Martin Huber

geboren in Frankenthal

2019

Behavioral Cloning for Autonomous Navigation

of Humanoid Robots

with Nonlinear Model Predictive Control

Die Masterarbeit wurde von Martin Huber

ausgeführt am

Institut für Optimierung, Robotik und Biomechanik

unter der Betreuung von

Frau Prof. Katja Mombaur

Verhaltensklonung zur autonomen Navigation humanoider Roboter mit Nichlinearer Modellprädiktiver Regelung:

In dieser Arbeit erkunden wir die Möglichkeiten der Verhaltensklonung zur autonomen Navigation humanoider Roboter durch bloße Bilder. Hierfür wird eine nichtlineare, Modellprädiktive Regelung, die es ermöglicht, stabile Lauftrajektorien in Echtzeit zu erzeugen, implementiert und evaluiert. Es wird demonstriert, dass minimale Veränderung in der Bildverarbeitung genügen, um vielseitige Bewegungsstrategien in vielfältigen dynamischen und statischen Umgebungen zu erlernen. Diese Einfachheit der Lösung wird als passende Ergänzung zur Meidung von Konvexen Hindernissen identifiziert, welche durch Randbedingungen die Lösungen der nichtlinearen Modellprädiktiven Regelung einschränken. Alle Experimente werden an Heicub, einer Variente des iCub, durchgeführt, welcher speziell für Optimalsteuerung in der Fortbewegung am Istituto Italiano di Tecnologiia in Genove entwickelt wurde. Die Auswertung von Stabilitätskriterien zeigt weiterhin, dass ein menschlicher Kontrolleur, einem künstlichen Agenten gegenüber, nicht überlegen ist. Um die präsentierte Methode schließlich auf tauschende Aufgaben zu erweitern, vereinfachen wir die wechselnden Umgebungen auf ein gut gelöstes Klassifizierungsproblem.

Behavioral Cloning for Autonomous Navigation of Humanoid Robots with Nonlinear Model Predictive Control:

In this work we investigate the capabilities of behavioral cloning for autonomous navigation of humanoid robots from raw image input. Therefore, a nonlinear model predictive control that allows for real time generation of stable walking trajectories is implemented and evaluated. It is demonstrated that minor modifications in the vision pipeline are sufficient for the learning of versatile motion strategies in various dynamic and static environments. This simplicity is identified as a well suited addition to the avoidance of convex obstacles, which are represented by constraints to the solution of the implemented nonlinear model predictive control. All of the experiments are carried out on Heicub, a descendant of the iCub, which was especially designed for optimal control in locomotion at the Istituto Italiano di Tecnologia in Genova. The evaluation of stability criteria further reveals that there is no superiority of a human controller over an artificial agent. Finally, to extend the proposed approach to changing tasks, we boil the variation of environments down to a well solved classification problem.

Contents

1 Introduction

2 State of the Art

3 Background

To generate dynamically balanced walking trajectories for humanoid robots and to let them navigate the environment autonomously, there are several posed challenges that we need to cover. As the logical starting point, in section 3.1 - Humanoid Walking, we want to address the real time generation of walking trajectories for humanoid robots first, and then think of ways to replace the human user by an artificial agent in the control loop (fig. 3.1). The generation of patterns in real time becomes feasible by treating the robot's physics in a simplified way as those of an inverted pendulum (sec. 3.1.1). The zero moment point of the linear inverted pendulum will therefore serve as the balance criteria for the solution of a sequentially quadratic problem (sec. 3.1.2). Resulting positions and orientations for the center of mass and the feet will then be interpolated (sec. 3.1.3) and passed as constraints to the inverse kinematics (sec. 3.1.4) so to transform them into joint angles that can be sent to the humanoid's motor controllers.

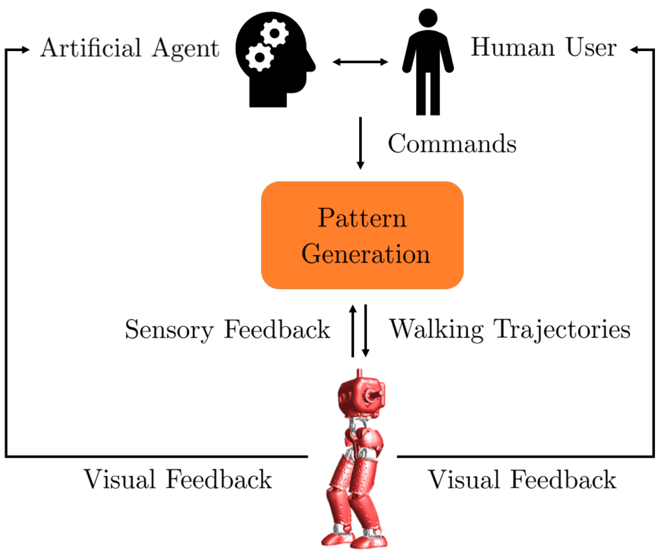

Figure 3.1: Simplified version of the proposed control loop to navigate the robot with either a human user or an artificial agent.

To close the control loop and to steer the robot towards desired goals, whilst avoiding obstacles, requires some sort of high level command that arises from visual feedback. As discussed in section 2 - State of The Art, there are several ways to achieve this, among them human users. Of particular interest to us are novel methods that evolved from the toolbox of machine learning techniques, as they decrease the computational cost into non existence. Let alone this fact enables us to run them onboard on light weight hardware with low energy usage, which is critical in the domain of humanoid robots. Center to these new methods will be neural nets that we will train on solving the task of autonomous navigation in two different ways. One of which clones the behavior of a human user (sec. 3.2.1), whereas the second presented method (sec. 3.2.2) explores policies and tries to find solutions on its own.

As a side note, within the following chapters there will always be made references to the actual implementation of the presented concepts. This shall enable future readers to bridge the gap between theory and application.

3.1 Humanoid Walking

To get started with and to understand the presented concepts that generate dynamically balanced walking trajectories, we shall have a look at figure 3.1 once more. The pattern generation therein (orange box), consists of four main building blocks: Forward kinematics, nonlinear model predictive control (NMPC), interpolation, and inverse kinematics. The relation between these four building blocks is shown in fig. 3.2. The natural entry point, to this otherwise closed control loop, is given by the

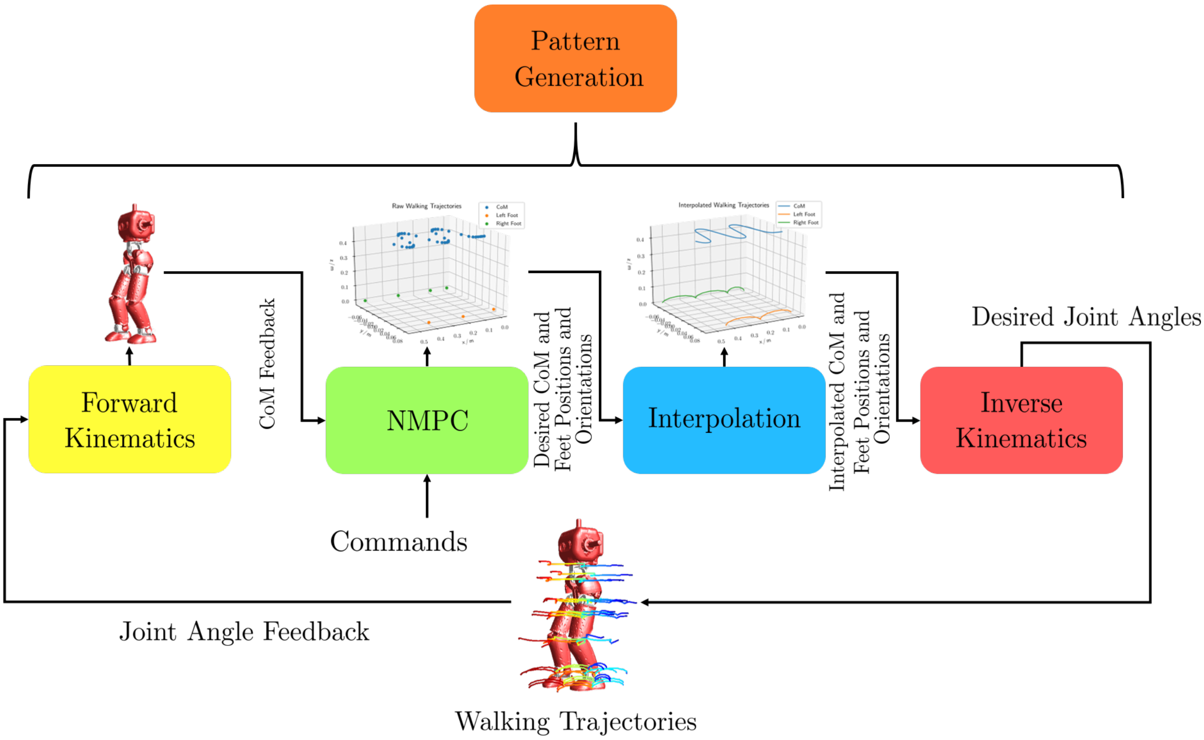

Figure 3.2: Building blocks of the pattern generation. To understand the greater picture, a connection can be drawn to fig. 3.1, where the orange box represents the one shown in this figure.

commands that enter the nonlinear model predictive control. Commands are passed in the form of a desired velocity $\mathbf{v}_{\text{ref}}$ that the robot's center of mass (CoM) shall satisfy optimally according to a cost function that also takes dynamic balance and a smooth motion into account. The future desired positions and orientations for the CoM and the feet then result from the solution to a sequentially quadratic problem that tries to minimize this cost function. The balance criteria within this problem formulation is based upon the zero moment point (ZMP) around which the whole control framework is built. It is only by simplifying the robot's model that we can solve the optimal control problem in real time. Therefore, we assume the robot to be a linear inverted pendulum, for which we have a well defined analytical relation between the CoM and the ZMP. The minimization of the distance between the analytical expression of the ZMP and the foot placement results in the desired dynamic balance. As shown in fig. 3.2, the desired CoM and the feet positions and orientation, as they are obtained from the NMPC, are sparsely distributed in space. Moreover, there is neither information about how the feet shall move along the z-axis, nor along the x-, and y-axis, but only where they should be placed in the x-y-plane. Therefore, as the subsequent step to the NMPC, we need to add an interpolation. The interpolation interpolates the trajectories of the CoM to obtain a finer sampling time. Additionally, the movement of the feet in the x-, y-, and z-direction, as well as their orientation, is computed by polynomials that we require to satisfy the initial and end conditions of the foot placement. Put together, the nonlinear model predictive control and the interpolation between the resulting subsequent solutions for the positions and orientation of both, the CoM and the feet, describe dynamically balanced trajectories, given that the humanoid robot of interest resembles the physics of an inverted pendulum. Now to bridge the gap between dynamically balanced trajectories in Cartesian space, and a humanoid robot that actually satisfies them with its CoM and its feet, the inverse kinematics problem needs to be addressed. The inverse kinematics, which follow immediately after the interpolation step, take the positions and orientations of the CoM and the feet as constraints and find a composite of joint angles that fulfill them. The continuity of subsequent solutions is therein assured by initializing the inverse kinematics with the previous solution. Resulting joint angles, once passed to the humanoid, then result in walking trajectories, as indicated in fig. 3.2 by the colored lines at the joints of the robot. Due to the inherent mismatch of the robot's physics from that of an inverted pendulum, as well as other effect like friction, there is a chance that the desired joint angles differ from the actually achieved ones. To compensate for the discrepancy, the last building block of the pattern generation is the feedback of the measured CoM to the NMPC. The CoM is computed by reading out the achieved joint angles, so that the forward kinematics can be utilized to determine the positions and orientations of the humanoid's links in space, and therefore the CoM.

As already highlighted in the previous paragraph, special attention has to be given to the zero moment point, since it defines the central concept of the presented pattern generation. We therefore will explain its theoretical foundations, as well as its analytical relationship to the CoM for simplified physical models, and ways to measure it with force torque sensors in the section that lies ahead - Zero Moment Point.

3.1.1 Zero Moment Point

Concept of the Zero Moment Point

Zero Moment Point of a Linear Inverted Pendulum

Computation of the Zero Moment Point

3.1.2 Nonlinear Model Predictive Control

3.1.3 Interpolating Trajectories

3.1.4 Kinematics

Forward Kinematics

Inverse Kinematics

3.2 Machine Learning

3.2.1 Behavioral Cloning

3.2.2 Reinforcement Learning

3.2.3 Image Processing

4 Methods

4.1 Software

4.2 Implementation

5 Experiments

5.1 User Controlled Walking

5.2 Autonomous Walking

6 Conclusion

Part I

Appendix

A Lists

A.1 List of Figures

A.2 List of Tables

B Bibliography

Erklärung:

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den (Datum) ....