

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

Department of Physics and Astronomy

University of Heidelberg

Master thesis

in Physics

submitted by

(name and surname)

born in (place of birth)

(year of submission)

(Title)
(of)
(Master thesis)

This Master thesis has been carried out by (Name Surname)

at the

(institute)

under the supervision of

(Frau/Herrn Prof./Priv.-Doz. Name Surname)

Verhaltensklonung zur autonomen Navigation humanoider Roboter mit Nichtlinearer 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 Variante des iCub, durchgeführt, welcher speziell für Optimalsteuerung in der Fortbewegung am Istituto Italiano di Tecnologia in Genua 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	7
2	State of the Art	8
3	Background	9
3.1	Humanoid Walking	10
3.1.1	Zero Moment Point	12
3.1.2	Nonlinear Model Predictive Control	16
3.1.3	Interpolating Trajectories	16
3.1.4	Kinematics	16
3.2	Machine Learning	16
3.2.1	Behavioral Cloning	16
3.2.2	Reinforcement Learning	16
3.2.3	Image Processing	16
4	Methods	18
4.1	Software	18
4.2	Implementation	18
5	Experiments	19
5.1	User Controlled Walking	19
5.2	Autonomous Walking	19
6	Conclusion	20
I	Appendix	21
A	Lists	22
A.1	List of Figures	22
A.2	List of Tables	22
B	Bibliography	23

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}_{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

The key metric in this work, for the generation of a dynamically balanced gait, is the zero moment point. The concept was first introduced by Miomir Vukobratović and Davor Juričić in 1968 [1][2] and first utilized 1984 to generate walking trajectories for the WL-10RD robot [3]. The most intuitive understanding for the ZMP arises by thinking about the realization of the simplest arbitrary possible walking motion for which a humanoid robot will not fall. This motion is achieved by ensuring the feet's whole area, and not only the edge, is in contact with the ground [4], or put in other words, we require the robot not to rotate about its feet edges. This constraint can be met by having a reaction force \mathbf{F}_r between the foot and the ground, which compensates for all external moments \mathbf{M}_x , and \mathbf{M}_y around the x-, and y-axis at any time (fig. 3.3). The point \mathbf{r} , at which the reaction force acts, is physically

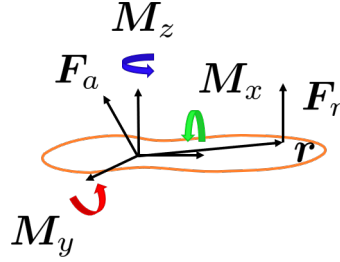


Figure 3.3: Forces acting on the sole.

only meaningful if it lies within the support polygon of the foot. Not only can it not exist outside of the support polygon, since there was no point of interaction between the foot and the ground then, but also was the robot to overturn under these circumstances. Therefore, the ZMP is defined as that point on the ground at which the net moment of the inertial forces has no component along the horizontal axes [5][6]. We now came to appreciate the importance of the support polygon for the definition of the zero moment point. The support polygon is defined as the

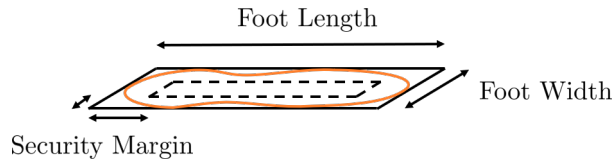


Figure 3.4: Full support polygon, and the resulting support polygon with security margin (dashed lines).

convex hull of all contact points of the feet with the ground, so the minimal number

of points to fully contain all of them. As the most restrictive case for balance, in this work we will only consider the support polygon of one foot at a time. Since the convex hull of a foot is well described by a rectangle, we only rely on the foot width ([link](#)), and foot length ([link](#)) to fully describe it. Also, to ensure that the zero moment point never comes close to the edges of the feet and therefore to provide balance, we define a security margin to their borders ([link](#)). The respective values are robot specific and can be set in the configurations file by following the provided links.

As already pointed out, within this work, we will use a simplified physical model of the humanoid solve the optimal control problem in real time. We will deal with this approximation in the following paragraph - Zero Moment Point of a Linear Inverted Pendulum.

Zero Moment Point of a Linear Inverted Pendulum

Dynamically balanced walking trajectories can be generated by simplifying the dynamics of humanoid robots to those of a linear inverted pendulum [7]. A rigorous derivation for the analytic relation between the center of mass and the zero moment point of a linear inverted pendulum can be found in [8], but for the sake of simplicity we rather explain the physics in terms of cutting forces, for which a short introduction can be found in the summary of the lecture Robotics 1 ([link](#)). The

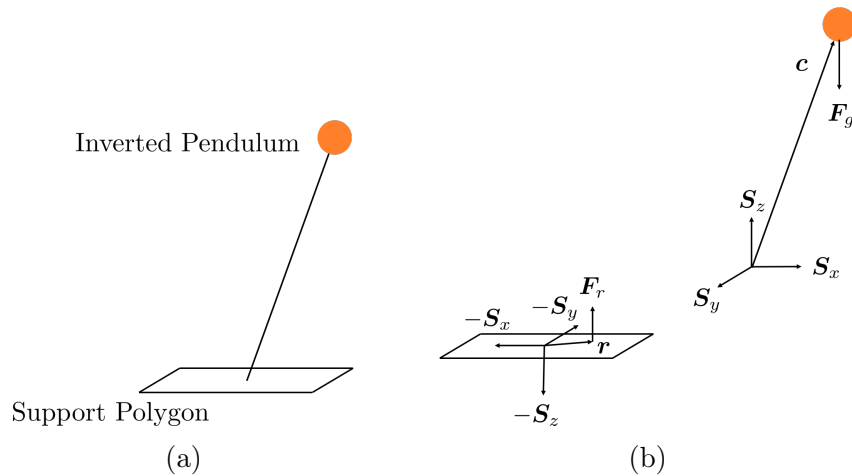


Figure 3.5: Linear inverted pendulum with a support polygon (a), and the corresponding free body diagram with cutting forces $S_{x/y/z}$ (b).

system of interest is shortly depicted in figure 3.5. We assume the support polygon of the shown linear inverted pendulum to have zero mass. By introducing cutting forces $S_{x/y/z}$ for each degree of freedom in which the motion of the linear inverted pendulum is restricted, we obtain the free body diagram (fig. 3.5), for which the

acting forces are

$$m\ddot{\mathbf{c}} = \mathbf{S} - \mathbf{F}_g \quad (3.1)$$

$$\mathbf{0} = -\mathbf{S} + \mathbf{F}_r \quad (3.2)$$

where $\mathbf{S} = \mathbf{S}_x + \mathbf{S}_y + \mathbf{S}_z$. The respective moments, since we do not take any inertias into account, are given by

$$\mathbf{0} = (\mathbf{0} - \mathbf{c}) \times \mathbf{S} + \mathbf{M} \quad (3.3)$$

$$\mathbf{0} = (\mathbf{r} - \mathbf{0}) \times \mathbf{F}_r - \mathbf{M} \quad (3.4)$$

where the transfer of the moment \mathbf{M} may for example be induced by friction. If we replace $\mathbf{S} = \mathbf{F}_r$ from eq. 3.2, equations 3.3 and 3.4 yield

$$\mathbf{0} = (\mathbf{r} - \mathbf{c}) \times \mathbf{S} = \begin{pmatrix} (r_y - c_y)S_z - (r_z - c_z)S_y \\ -(r_x - c_x)S_z + (r_z - c_z)S_x \\ (r_x - c_x)S_y - (r_y - c_y)S_x \end{pmatrix} \quad (3.5)$$

Since our goal is to have a robot that does not fall, we want to achieve that the acceleration along the z-axis becomes zero, hence $\ddot{c}_z = 0$. Given this assumption, we can infer from eq. 3.1 that $S_z = mg$, as well as $S_x = \ddot{c}_x m$, and $S_y = \ddot{c}_y m$. Furthermore, our foot shall not lift off the floor, and therefore we have $r_z = 0$. If we take these assumptions and plug them into the first to rows of eq. 3.5, we find

$$r_x = c_x - c_z \frac{\ddot{c}_x}{g} \quad (3.6)$$

$$r_y = c_y - c_z \frac{\ddot{c}_y}{g} \quad (3.7)$$

Therein, r_x , and r_y are the x-, and y-coordinates of the zero moment point, given the assumption of a linear inverted pendulum. We can see that the position is dependent on the height of the point mass, which is in turn dependent on the robot. The specific values can be set in the configurations file ([link](#)).

We have now found a simple analytic expression for the relationship of the zero moment point and the center of mass, which will help us to formulate an optimal control problem that we can solve in real time. This simplification is of course only true to some extent, and we need to find a way to verify its accuracy. The easiest way to do so is to measure the real zero moment point. We will further elaborate on this within the next paragraph - Measurement of the Zero Moment Point, and we will derive a method that only relies on force torque sensors in the ankle.

Measurement of the Zero Moment Point

There are several methods that enable us to measure the position of the zero moment point, among them the utilization of pressure sensitive soles, as outlined in [8].

Furthermore, there exist approximate approaches that involve the knowledge of all acting external forces [9], which can for example be obtained from unconstrained inverse dynamics [10]. Since we can rely on measurements of force torque sensors that are located at the ankles, we will infer the position of the zero moment point from them [8]. If we consider the force torque sensor to be located at a position \mathbf{p}_i (fig. ??), then we can obtain the moment about any point \mathbf{p} according to eq. 3.8.

$$\boldsymbol{\tau}(\mathbf{p}) = (\mathbf{p}_i - \mathbf{p}) \times \mathbf{f}_i + \boldsymbol{\tau}_i \quad (3.8)$$

by definition, the moment about the zero moment point vanishes along the horizontal axes, therefore we can then set $\tau_x = \tau_y = 0$ in eq. 3.8 and then solve for the position to obtain the zero moment point (eq. 3.9 and 3.10).

$$p_x = \frac{[-\tau_{i,y} - (p_{i,z} - p_z)f_{i,x} + p_{i,x}f_{i,z}]}{f_{i,z}} \quad (3.9)$$

$$p_y = \frac{[-\tau_{i,x} - (p_{i,z} - p_z)f_{i,y} + p_{i,y}f_{i,z}]}{f_{i,z}} \quad (3.10)$$

If we further choose our coordinate system to lie along the z-axis of the force torque sensor (fig. ...), we can simplify equations 3.9 and 3.10 to find

$$p_x = \frac{(-\tau_{i,y} - f_{1,x}d)}{f_{1,z}} \quad (3.11)$$

$$p_y = \frac{(\tau_{i,x} - f_{1,y}d)}{f_{1,z}} \quad (3.12)$$

We can use equations 3.11 and 3.12 to determine the position of the zero moment point for the left and the right foot with respect to coordinate frames that are attached to the respective foot. These circumstances change once not only one, but both feet are in contact with the ground. What still holds true, in the case of a dynamically balanced gait, is the fact that the positions which we just obtained from equations 3.11 and 3.12 represent points where the interaction of the robot with the environment can solely be described by a single force along the z-axis. All other forces or torques cancel out. Therefore, to determine the position of the zero moment point for the double support phase, we need to modify equation 3.8 slightly. This yields

$$\boldsymbol{\tau}(\mathbf{p}) = \sum_{i \in \{L,R\}} (\mathbf{p}_i - \mathbf{p}) \times \mathbf{f}_i \quad (3.13)$$

where the individual torques are now zero and the only forces \mathbf{f}_i that exist between the robot and the environment can be described by the z-component which are measured at the ankles' force torque sensors. Yet again, to obtain the position of the zero moment point, we have to set the x-, and y-components of the torque in

equation 3.13 to zero and find

$$p_x = \frac{\sum_{i \in \{L,R\}} p_{i,x} f_{i,z}}{\sum_{i \in \{L,R\}} f_{i,z}} \quad (3.14)$$

$$p_y = \frac{\sum_{i \in \{L,R\}} p_{i,y} f_{i,z}}{\sum_{i \in \{L,R\}} f_{i,z}} \quad (3.15)$$

These expressions of course only hold true in a shared coordinate system and therefore we need to transform the position of the zero moment point which we obtained from equations 3.11 and 3.12 to the world frame. Finally, we can write down the formulation for the zero moment point which holds equally true for the single and double support phase

$$p_x = \frac{p_{R,x} f_{R,z} + p_{L,x} f_{L,z}}{f_{R,z} + f_{L,z}} \quad (3.16)$$

$$p_y = \frac{p_{R,y} f_{R,z} + p_{L,y} f_{L,z}}{f_{R,z} + f_{L,z}} \quad (3.17)$$

At this point we are now equipped with a general understanding for the zero moment point, as well as with the knowledge of simplified models to compute it analytically, and a method to measure it so that we can evaluate the performance of a potential pattern generator which is based upon the zero moment point. Therefore, in the next chapter - Nonlinear Model Predictive Control, we will try to understand a method that allows us to generate dynamically balanced center of mass and feet trajectories, which satisfy the just introduced concepts optimally, given a weighting factor.

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

In the previous sections we have learned about two different approaches to train neural nets on solving certain tasks. Although we came to understand that the

complexity of the task to be solved correlates strongly with the amount of data at hand, there exist domains from which it is undeniably easier to do so. To equip a neural net with some sort of prior knowledge by switching the domain may therefore not only be highly desirable but sometimes also needed if the amount or quality of data is not sufficient. One domain which is of special interest when it comes to interacting in a three dimensional environment is a domain that represents depth information. If there are any, it may sometimes be possible to extract this kind of prior knowledge from a depth camera. As for this work, we need to rely on stereo cameras and powerful algorithms that allow us to compute depth images in real time. The algorithm that helps us to do so, in terms of the extraction of weighted least squares disparity maps, will be presented in the following - Depth Map Extraction.

Depth Map Extraction

[11][12]

As a requirement for the algorithm to work properly, it is important to calibrate the robot's cameras. Therefore, the next chapter - Mono and Stereo Camera Calibration, will explain in detail why, and how to calibrate cameras.

Mono and Stereo Camera Calibration

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

3.1	Simplified version of the proposed control loop to navigate the robot with either a human user or an artificial agent.	9
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.	10
3.3	Forces acting on the sole.	12
3.4	Full support polygon, and the resulting support polygon with security margin (dashed lines).	12
3.5	Linear inverted pendulum with a support polygon (a), and the corresponding free body diagram with cutting forces $\mathbf{S}_{x/y/z}$ (b).	13

A.2 List of Tables

B Bibliography

- [1] Miomir Vukobratovic and Davor Juricic. Contribution to the synthesis of biped gait. *IFAC Proceedings Volumes*, 2:469–478, 1968.
- [2] Miomir Vukobratovic and Davor Juricic. Contribution to the synthesis of biped gait. *IEEE Transactions on Biomedical Engineering*, pages 1–6, 1969.
- [3] Jin-ichi Yamaguchi, Atsuo Takanishi, and Ichiro Kato. Development of a biped walking robot compensating for three-axis moment by trunk motion. *Journal of the Robotics Society of Japan*, 11(4):581–586, 1993.
- [4] Miomir Vukobratovi and Branislav Borovac. Zero-moment point—thirty five years of its life. *International journal of humanoid robotics*, 1(01):157–173, 2004.
- [5] Kazuo Hirai, Masato Hirose, Yuji Haikawa, and Toru Takenaka. The development of Honda humanoid robot. *Proceedings. 1998 IEEE International Conference on Robotics and Automation (Cat. No. 98CH36146)*, 2:1321–1326, 1998.
- [6] Anirvan Dasgupta and Yoshihiko Nakamura. Making feasible walking motion of humanoid robots from human motion capture data. *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, 2:1044–1049, 1999.
- [7] Shuuji Kajita, Fumio Kanehiro, Kenji Kaneko, Kiyoshi Fujiwara, Kensuke Harada, Kazuhito Yokoi, and Hirohisa Hirukawa. Biped walking pattern generation by using preview control of zero-moment point. *2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422)*, 2:1620–1626, 2003.
- [8] Shuuji Kajita, Hirohisa Hirukawa, Kensuke Harada, and Kazuhito Yokoi. *Introduction to humanoid robotics*. Springer, 2014.
- [9] Qiang Huang, Kazuhito Yokoi, Shuuji Kajita, Kenji Kaneko, Hirohiko Arai, Noriho Koyachi, and Kazuo Tanie. Stability compensation of a mo-bile manipulator by manipulator motion: Feasibility and planning. *IEEE Transactions on robotics and automation*, 17(3):280–289, 2001.
- [10] Raphael Michel. *Dynamic filter for walking motion corrections*. Bachelor’s thesis, Heidelberg University, 2017.

- [11] Rostam Affendi Hamzah, Rosman Abd Rahim, and Zarina Mohd Noh. Sum of absolute differences algorithm in stereo correspondence problem for stereo matching in computer vision application. In *2010 3rd International Conference on Computer Science and Information Technology*, volume 1, pages 652–657, 2010.
- [12] Dongbo Min, Sunghwan Choi, Jiangbo Lu, Bumsub Ham, Kwanghoon Sohn, and Minh N. Do. Fast global image smoothing based on weighted least squares. *IEEE Transactions on Image Processing*, 23(12):5638–5653, 2014.

Erklärung:

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

Heidelberg, den (Datum)