



# Motion Control of a Hexapod Robot Over Uneven Terrain Using Heightmaps

by

Andries Phillipus Lotriet

*Thesis presented in partial fulfilment of the requirements for  
the degree of Master of Engineering (Electronic) in the  
Faculty of Engineering at Stellenbosch University*

Supervisor: Prof. J.A.A. Engelbrecht

March 2023

# Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: ..... 2023/02/10

Copyright © 2023 Stellenbosch University  
All rights reserved.

# Abstract

## Motion Control of a Hexapod Robot Over Uneven Terrain Using Heightmaps

A.P. Lotriet

*Department of Electrical and Electronic Engineering,  
Stellenbosch University,  
Private Bag X1, Matieland 7602, South Africa.*

Thesis: MEng (EE)

March 2023

In recent times great strides have been made in the field of autonomous robotics, especially with regards to autonomous navigation of wheeled and aerial drones. Legged robotics however still face numerous problems before they can become practical to use, the most egregious of these problems being balancing of the robot and optimal foot placement.

This thesis focuses on providing a solution to the latter problem of foot placement. This is achieved by using a depth camera to, in real time, construct a localised map of the environment and subsequently analysing said map for optimal foot placement locations. The system is then tested using a hexapod robot both in simulation and on a physical robot.

# Acknowledgments

# Dedication

# Table of contents

<b>List of figures</b>	<b>viii</b>
<b>List of tables</b>	<b>x</b>
<b>List of symbols</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Background . . . . .	1
1.2 Research Goal . . . . .	1
1.3 Methodology . . . . .	2
1.4 Scope and Limitations . . . . .	2
1.5 Thesis Outline . . . . .	3
<b>2 Literature review</b>	<b>4</b>
2.1 Hexapod history . . . . .	4
2.2 Control . . . . .	5
2.2.1 Traditional . . . . .	5
2.2.2 Bio Inspired . . . . .	6
2.2.3 Reinforcement Learning . . . . .	6
2.3 Sensing Methods . . . . .	6
2.4 Simulation Environment . . . . .	7
2.5 Uneven Terrain . . . . .	7
2.6 Research Decisions . . . . .	8
<b>3 System Overview</b>	<b>10</b>
<b>4 Mapping</b>	<b>13</b>
4.1 Projection . . . . .	13
4.2 Map Buffer . . . . .	16
<b>5 Motion</b>	<b>17</b>
5.1 Gait State Machine . . . . .	18
5.1.1 Choosing The Supporting And Swinging Legs . . . . .	19
5.1.2 Choosing nominal foot positions . . . . .	20

5.2	Foot Motion . . . . .	22
5.2.1	Existing System . . . . .	22
5.2.2	Improved System . . . . .	23
5.3	Kinematics . . . . .	25
5.3.1	Coordinate Frames . . . . .	25
5.3.2	Inverse Kinematics (IK) . . . . .	26
5.3.3	Forward Kinematics (FK) . . . . .	27
5.3.4	Angular Rate . . . . .	27
<b>6</b>	<b>Foot Placement Optimisation</b>	<b>29</b>
6.1	Scoring . . . . .	29
6.1.1	Gradient Score . . . . .	29
6.1.2	Terrain Proximity Score . . . . .	31
6.1.3	Constraints . . . . .	32
6.2	Method of adjustment . . . . .	32
6.3	Floor Height Adjustment . . . . .	33
<b>7</b>	<b>Hardware Implementation</b>	<b>34</b>
7.1	ROS Nodes . . . . .	34
7.2	ROS communication . . . . .	35
7.2.1	Base Station . . . . .	35
7.2.2	On Board . . . . .	36
7.2.3	ROS Data Types . . . . .	38
7.3	GPU Compute Pipeline . . . . .	39
7.3.1	A note on Graphical Processing Unit (GPU) architecture	39
<b>8</b>	<b>Results</b>	<b>41</b>
8.1	Requirements . . . . .	41
8.2	Depth Camera . . . . .	41
8.3	Heightmap . . . . .	41
8.4	Walkability Scores . . . . .	41
8.5	Walking . . . . .	45
8.5.1	Flat Plane . . . . .	45
8.5.2	Staircase . . . . .	47
8.5.3	Organic . . . . .	49
<b>9</b>	<b>Conclusions</b>	<b>51</b>
<b>A</b>	<b>Reference Frame Transforms</b>	<b>52</b>
A.1	Camera to Map . . . . .	52
A.2	World to Map . . . . .	53
A.3	Local to Map . . . . .	53
A.4	Map to Local . . . . .	53

*TABLE OF CONTENTS*

	<b>vii</b>
A.5 Body to Leg . . . . .	54
<b>List of references</b>	<b>55</b>

# List of figures

2.1	A Flesh-fly . . . . .	4
2.2	A circular hexapod . . . . .	4
2.3	Trends of hexapod control schemes (Coelho <i>et al.</i> , 2021) . . . . .	5
2.4	Exploratory terrain traversability classification from (Prágr <i>et al.</i> , 2019) . . . . .	7
2.5	Quadruped terrain scoring and navigation from (Mastalli <i>et al.</i> , 2020) . . . . .	8
3.1	Basic system operation . . . . .	10
3.2	Advanced system operation . . . . .	10
3.3	Physical system diagram . . . . .	11
3.4	Physical Hexapod . . . . .	12
4.1	Camera Projection . . . . .	14
4.2	Memory Diagram . . . . .	16
5.1	Motion System Overview . . . . .	17
5.2	Gait State Machine . . . . .	18
5.3	Leg sextants, with sextant 1 being active. . . . .	19
5.4	Supporting target choosing diagram . . . . .	20
5.5	Swinging target choosing diagram . . . . .	21
5.6	Existing arc recomputation problem . . . . .	22
5.7	End effector movement path . . . . .	23
5.8	Sigmoid Like . . . . .	24
5.9	Body and leg coordinate frames. . . . .	25
5.10	Leg coordinate frame with kinematic variables. . . . .	26
6.1	Slope Score . . . . .	30
6.2	Proximity Score Diagrammatic Representation . . . . .	31
6.3	Radial search, shown for a 5 cell search diameter. . . . .	32
6.4	Floor height diagram. . . . .	33
7.1	ROS nodes and communication. . . . .	34
7.2	Compute pipeline. . . . .	39

8.1	3D Model of Terrain and its heightmap. . . . .	41
8.2	Slope score and edge proximity score tests. . . . .	42
8.3	Full walkability score. . . . .	43
8.4	Walkability score optimisation test. . . . .	43
8.5	Flat plane walk test . . . . .	45
8.6	Body height (top). Body tilt (center). Position of one foot (bottom)	46
8.7	Stairs walk test . . . . .	47
8.8	Body height (top). Body tilt (center). Position of one foot (bottom)	48
8.9	Organic walk test. . . . .	49
8.10	Body height (top). Body tilt (center). Position of one foot (bottom)	50

# List of tables

5.1	State Definitions . . . . .	18
7.1	Base station publishers . . . . .	35
7.2	Base station subscribers . . . . .	35
7.3	Jetson publishers . . . . .	36
7.4	Jetson subscribers . . . . .	37
7.5	Teensy publishers . . . . .	37
7.6	Teensy subscribers . . . . .	38
7.7	Robot Operating System (ROS) data type descriptions . . . . .	38

# List of symbols

## Constants

$$L_0 \quad 300 \text{ mm}$$

## Variables

$Re_D$	Reynolds number (diameter)	[ ]
$x$	Coordinate	[ m ]
$\ddot{x}$	Acceleration	[ m/s <sup>2</sup> ]
$\theta$	Rotation angle	[ rad ]
$\tau$	Moment	[ N·m ]

## Vectors and Tensors

### Subscripts

a	Adiabatic
$a$	Coordinate

### Superscripts

$\mathcal{W}$	Indicates the world reference frame. <sup>1</sup>
$\mathcal{B}$	Indicates the robot body reference frame. <sup>1</sup>
$\mathcal{M}$	Indicates the map reference frame. <sup>1</sup>
$\mathcal{C}$	Indicates the camera reference frame. <sup>1</sup>
$\mathcal{L}_i$	Indicates the reference frame of leg $i$ . <sup>1</sup>

---

<sup>1</sup>See appendix ?? for reference frame definitions.

**Notation**

$x^{\mathcal{A}}$	Indicates that $x$ is in reference frame $\mathcal{A}$ .
$T_{\mathcal{AB}}(x)$	Transforms $\mathcal{A}$ to reference frame $\mathcal{B}$ . <sup>2</sup>
$\llbracket a_0, \dots, a_n \rrbracket$	Defines a integer sequence from $a_0$ to $a_n$ , with a increment of 1.
$\langle \mathbf{a}, \mathbf{b} \rangle_F$	Takes the Frobenius inner product of matrix $\mathbf{a}$ and $\mathbf{b}$ .

**Abbreviations**

<b>IK</b>	Inverse Kinematics
<b>FK</b>	Forward Kinematics
<b>MuJoCo</b>	Multi-Joint dynamics with Contact
<b>GUI</b>	Graphical User Interface
<b>ROS</b>	Robot Operating System
<b>LiDAR</b>	Light Detection and Ranging
<b>RGB-D</b>	Red Green Blue Depth
<b>SLAM</b>	Simultaneous Localisation and Mapping
<b>IMU</b>	Inertial Measuring Unit
<b>RL</b>	Reinforcement Learning
<b>ANN</b>	Artificial Neural Network
<b>GPS</b>	Global Positioning System
<b>MCU</b>	Microcontroller Unit
<b>GPU</b>	Graphical Processing Unit
<b>CPU</b>	Central Processing Unit

---

<sup>2</sup>See appendix A for transform definitions.

# Chapter 1

## Introduction

### 1.1 Background

There are many applications where vehicles are required to traverse rough terrain, such as in mines, rescue operations, agriculture, construction, etc. In many of these use cases rough terrain makes the use of wheeled, or even tracked, vehicles difficult or impractical.

Compared to wheeled robots, legged robots could perform better in many of these environments, allowing navigation over terrain that would be impossible for wheeled or tracked vehicles to navigate. While legged robots possess extreme degrees of potential terrain traversability, advanced control and sensory systems are required to realise this potential.

### 1.2 Research Goal

The overarching goal of this project is to design and implement a sensory and control system that will allow a hexapod robot to autonomously walk over rough terrain.

This goal of the project is broken up into the following sub objectives:

1. Obtain a mathematical model of the robot, its actuators and its sensors.
2. Create a model of the robot in a simulation environment for development and testing.
3. Implement a vision based Simultaneous Localisation and Mapping (SLAM) system.
4. Develop a real time vision based dense mapping system for use in anchor point selection.

5. Develop a optimisation system to select optimal end effector anchor points based on the surrounding terrain.
6. Implement and test the system in simulation and on the physical hexapod.

### 1.3 Methodology

When deciding how to determine optimal end effector placement various sensing methods were considered, such as using a Red Green Blue Depth (RGB-D) camera to view the environment, placing force sensors on the robots end effectors or measuring servo torque to determine when the end effectors were in contact with a surface. A previous paper by Erasmus *et al.* (2023) used a RGB-D camera by storing past snapshots to adjust the end effectors to the optimal height, it was decided that the primary sensing method for this thesis would also be a RGB-D camera but instead of storing snapshots, a height map would be generated of the local environment. This would allow for more advanced methods of anchor point selection.

The first step in realising this system was to construct an accurate simulation of the hexapod. The primary simulation packages that were considered are Gazebo, PyBullet and Multi-Joint dynamics with Contact (MuJoCo). Gazebo was an appealing choice due to the easy integration with ROS, however it was decided to use MuJoCo since it was found to have superior contact physics simulation (Erez *et al.*, 2015).

Once the hexapod was adequately modelled in MuJoCo a tripod gait state machine, IK system and control interface was implemented, at this stage the hexapod was capable of walking on flat terrain.

Next the system to generate the height map was implemented, this entailed sampling the RGB-D camera and comparing cells in the height map against the depth buffer. Once the height map was implemented it was possible to build the system responsible for end effector placement, this is covered in detail in ??, after which collision checking for the generated end effector motion was implemented, ensuring that the hexapod does not get stuck on pieces of terrain.

With this the system was realised in simulation, next the system was implemented and tested on the physical robot, discussed in detail in chapter 7

### 1.4 Scope and Limitations

As the hardware used was developed by Erasmus *et al.* (2023) this project will focus only on developing the necessary software to control the robot hardware.

The velocity control, tilt angle stabilisation and end effector motion planner was developed by the author, while the low level IK controller used was developed by (Erasmus *et al.*, 2023). The scope of this project does not include autonomous waypoint navigation and thus requires a human operator to provide desired velocity commands. If no solution can be found for the given velocity command the system will not attempt to adjust the velocity command, the human operator will be required to adjust the command.

The local dense height map system was developed by the author, while the SLAM system used, ORB-SLAM3 was developed by Campos *et al.* (2021). It should be noted however that ORB-SLAM3 does generate a global sparse feature map of the environment, thus implementing waypoint navigation should be a trivial addition.

The sensors used in this project are limited to a single RGB-D camera, thus even with the generation of a local map, there could be cases where the system will not have height data around a desired anchor point. No torque or touch sensors are used to augment the system, thus if a leg were to collide with the terrain the robot will not adjust its trajectory. The system will however attempt to choose a step path based on the local height tmap such that no collision occurs. Additionally no Inertial Measuring Unit (IMU) is used, pose estimation is entirely handled by the SLAM system.

## 1.5 Thesis Outline

Chapter 2 provides a literature review on the methods of control, sensing and simulation used for hexapod robots.

Chapter 3 provides a overview of the hexapod hardware and the modelling thereof. This includes the robots mechanical form, sensors, on board computers and the simulation environment that is used.

Chapter 4 describes the environment mapping systems used, this includes the local dense height map and the sparse SLAM system.

Chapter 5 covers motion related topics, this includes the walking gait, IK and effector motion planning.

Chapter 6 describes the optimisation function and its various scores used to acquire the optimal end effector anchor points during each step taken.

Chapter 7 covers the hardware implementation process and software structure on the hardware.

Chapter 8 describes the various tests preformed and results obtained thereof.

Chapter 9 provides the conclusion of the research and any recommended future additions.

# Chapter 2

## Literature review

This chapter provides an overview of past research done regarding the control of hexapod movement and sensing methods. First a brief history of hexapods is presented after which various terrain sensing and adaptation methods are presented.

### 2.1 Hexapod history

Hexapoda, Greek for "six legs" refers the group of arthropods possessing three pairs of legs. As an example see a flesh-fly in Figure 2.1.



Figure 2.1: A Flesh-fly



Figure 2.2: A circular hexapod

In the context of robotics "Hexapod" is used to refer to any robot with six legs, the most common configuration of Hexapods are either a rectangular layout with three legs on either side mimicking biological Hexapoda, or a circular design with radially symmetrical leg spacing, as seen in Figure 2.2

The hexapod possess the minimum number of legs to allow a naturally stable platform since while taking a step there can be upwards of three anchor points around the center of mass at all times. This makes the hexapod hexa-

pods an ideal platform to navigate complex terrain while maintain stability, without requiring advanced balancing control systems.

For a hexapod to walk it must lift some of its legs while bracing with others, the number of swinging to bracing legs, and how each is moved, is referred to as the walking "gait". The chosen gait influences the speed and stability of the hexapod, the tripod gait is considered to be the most well rounded, having good speed and stability. In the tripod gait three legs are bracing while the remaining three swing. A example of a more stable gait would be the One by One gait, where only one leg is moved at a time.

It is also possible to create a system where there is no predetermined gait, but rather the system determines the optimal legs to brace and swing depending on the current walking environment.

## 2.2 Control

Walking over rough terrain requires a control system to correctly actuate the hexapods legs. Various types of control schemes exist, the primary schemes are traditional controllers, bio-inspired controllers and Reinforcement Learning (RL). These three schemes are discussed below. Control trends can be seen in Figure 2.3

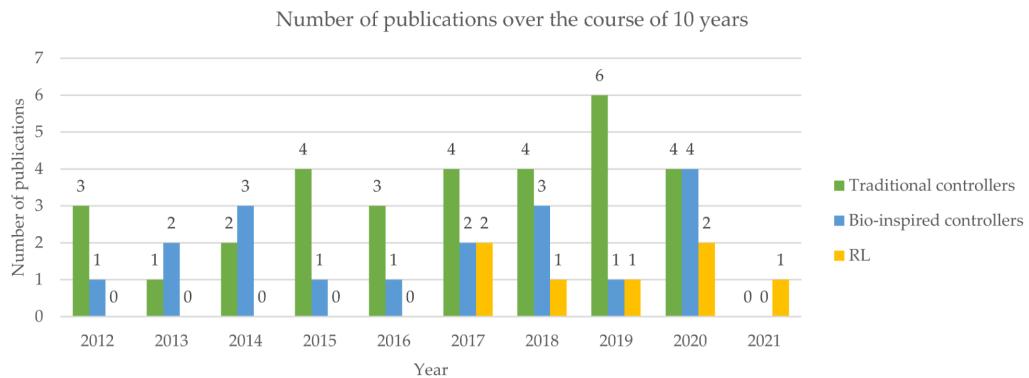


Figure 2.3: Trends of hexapod control schemes (Coelho *et al.*, 2021)

### 2.2.1 Traditional

Traditional controllers rely on an exact mathematical model of the robot and IK to calculate angular commands for all leg joints. This method of control is purely kinematic does not take into account external forces applied to the robot, thus it does not inherently adjust to the environment.

Instead of a purley kinematic model, a dynamic model can also be used. Using a dynamic model the forces acting on the robots legs are taken into account usually acquired through torque measurements from servos. By taking applied torque into account dynamic model controllers will intrinsically detect a deviation when an external force is applied to the robot or its legs and compensate appropriately.

It should be noted that it is possible for a kinematic model controller to also adjust to external disturbances, but this is not intrinsic to the control model and requires additional control logic.

### 2.2.2 Bio Inspired

Bio inspired controllers attempt to mimic the neural structure of animals to achieve the same locomotion methods that they use. This is implemented through the use of a Artificial Neural Network (ANN) If implemented successfully a bio inspired controller can be highly adaptable to the surrounding environment and is even able to adapt to damaged or missing legs.

### 2.2.3 Reinforcement Learng

RL controllers are created through using trial and error to construct a neural net that minimises a cost function for a specific goal. This theoretically allows RL controllers to adapt to any circumstances given enough time, allowing a very hight level of autonomy, as no prior direction is required. RL controllers are though notoriously difficult to train properly, especially when the amount of sensors and control outputs grow large, increasing the feature space. And event he most well trained RL agent still has the possibility to exhibit inexplicable behaviour.

## 2.3 Sensing Methods

No matter the control scheme used, to know where to place its feet the robot requires sensor(s) to sense its environment in some way, this could be achieved through simple sensors such as servo torque or touch. More advanced methods such as vision or Light Detection and Ranging (LiDAR) are also used. Homberger *et al.* (2017) uses stereoscopic vision to adjust end effector height and to classify surface materials.

Depending on the terrain navigation system it might be required to localise the robot in 3D space, for this it is possible to use external sensors such as a type of beacon (RF, Reflective, Ultrasonic), Global Positioning System (GPS) or, through the use of a SLAM system, internal sensors such as vision could be used.

## 2.4 Simulation Environment

The most popular physics simulators for robotics in recent times are Gazebo, MuJoCo and CoppeliaSim (previously V-REP) (Collins *et al.*, 2021). Gazebo and CoppeliaSim both have easy to use Graphical User Interface (GUI) interfaces and easy integration with ROS. MuJoCo on the other hand does not have a full GUI interface, only a simulation viewer, and does not have native ROS integration. Having said this MuJoCo was found to be the most accurate and fastest simulator when considering the use case of robotics (Erez *et al.*, 2015).

## 2.5 Uneven Terrain

As hexapods are inherently stable, they are good candidates for use in uneven terrain, and this tends to be where much hexapod research is focused. Such as in Homberger *et al.* (2017), which uses a stereoscopic camera to extract features and classify terrain by factors such as material type, slope and granularity. These parameters are then used to classify various gait parameters, such as step height and dampening, to improve stability while walking.

This is similar to Xu *et al.* (2023), which aims to classify terrain based on material properties, however, Xu *et al.* (2023) focuses on characterising terrain through touch sensors instead of visual sensors. This is achieved through the use of machine learning.

Moving away from gait optimisations and local terrain classification, Prágr *et al.* (2019) aims to characterize large areas of terrain in terms of navigational efficiency through the use of gaussian processes and a incrementally constructed spatial map. An example of this exploratory traversability classification can be seen in figure 2.4.

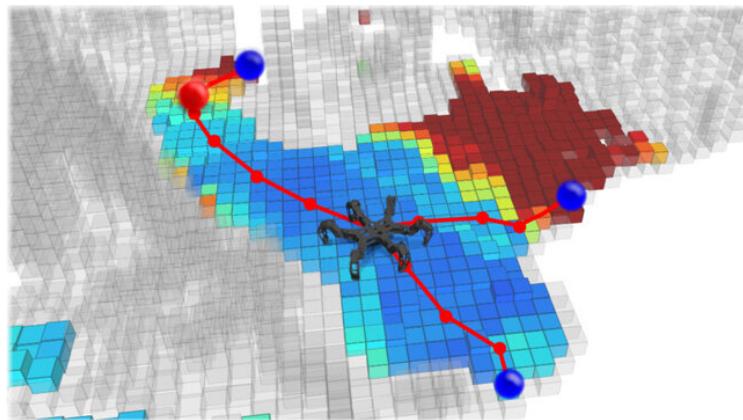


Figure 2.4: Exploratory terrain traversability classification from (Prágr *et al.*, 2019)

While this paper is on hexapods, it should be noted that much work has also been done on quadrupeds. For example Mastalli *et al.* (2020) uses various sensors to assign a cost to surrounding terrain, thus constructing a cost map, this cost map is then used to select optimal foot end positions for the next step. This can be seen in figure 2.5. Of course, being a quadruped, in addition

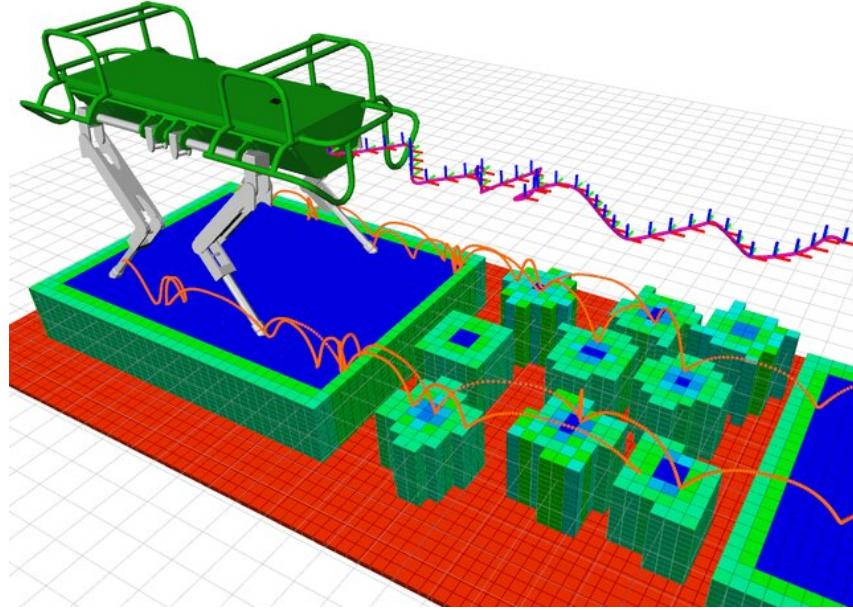


Figure 2.5: Quadruped terrain scoring and navigation from (Mastalli *et al.*, 2020)

to finding optimal foot end positions, significant consideration had to be given to maintaining the stability of the robot, something that is much less of a concern with hexapods.

## 2.6 Research Decisions

This research paper focuses on using vision based mapping, terrain classification and the optimisation of foot end positions, thus it was decided that traditional kinematic control is used as it is the simplest and most predictable method of control. Thus allowing the focus to remain on terrain classification and foot end position optimisation.

The camera that is used is the Intel Realsense D435i RGB-D camera, incorporating stereoscopic cameras and a LiDAR sensor. This camera also has a extensive existing codebase and support that ensures accurate depth readings are attained.

A system to localise the robot within its environment is required, as the primary sensor used is an RGB-D camera, various visual SLAM systems were

considered. ORB-SLAM 3, a optimisation-based, sparse map SLAM system was chosen to be used. ORB-SLAM 3 maintains a sparse map, an atlas, of both active and dormant features. This atlas is used to localise in the sparse map (Macario Barros *et al.*, 2022). The implementation of a dense map to be used for end effector placement is discussed in chapter 4. **Compare this more to others?**

Finally a simulation environment was chosen. MuJoCo has exceptional contact physics simulation and the only relevant downside is the lack of native ROS integration and the lack of a comprehensive GUI, which seeing as MuJoCo has good python bindings, could be seen as a advantage. Considering this, MuJoCo was chosen as the simulator.

# Chapter 3

## System Overview

The proposed system aims to improve the walking capability of the robot, as reference see figure 3.1

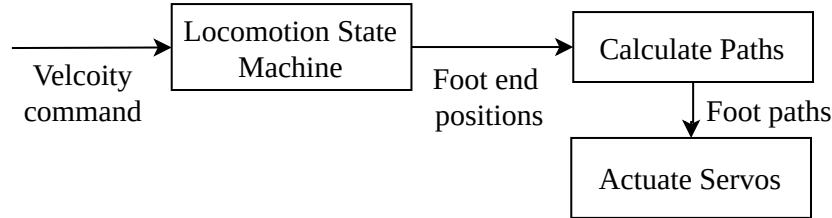


Figure 3.1: Basic system operation

The basic system will work well enough for walking over flat terrain, but will struggle once any deviation in terrain height is present. Thus the proposed, more advanced system, operates with the flow shown in figure 3.2.

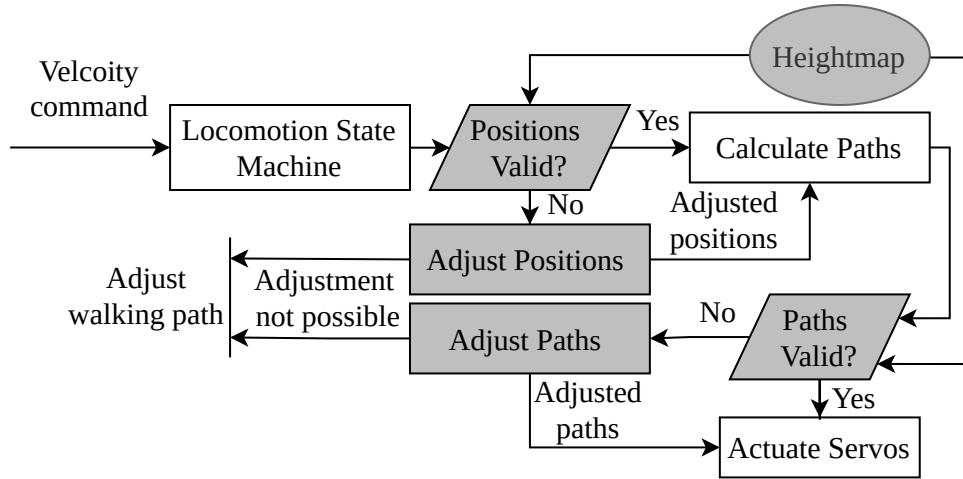


Figure 3.2: Advanced system operation

The advanced system uses similar components from the basic system, with some differences discussed in later chapters, but additionally incorporates checks against a heightmap and a score map to validate, and if necessary, adjust the nominal foot end positions and paths to move the feet to said positions. If the nominal positions are invalid and no valid adjustment can be made the robot will have to change its overall path, however this is not discussed in this paper.

The goal of maneuvering rough terrain is achieved by a combination of 4 primary systems, namely, a mapping, foot placement optimisation, motion controller and a SLAM system. A high level overview of the system implementation can be seen in figure 3.3.

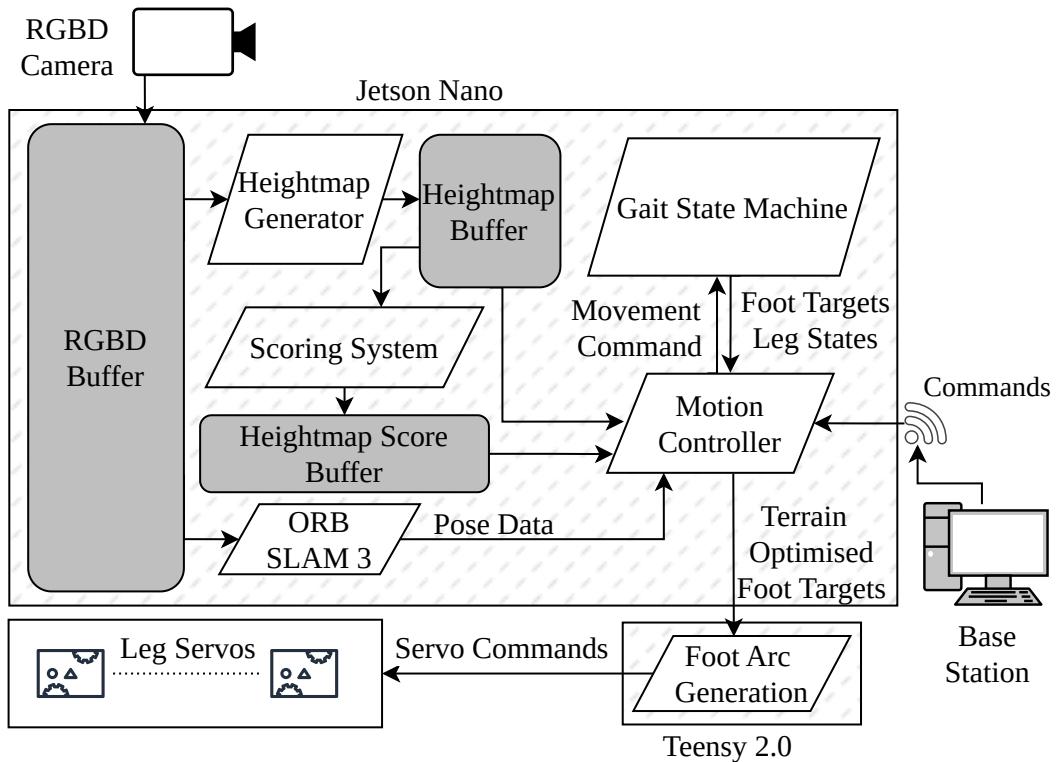


Figure 3.3: Physical system diagram

The mapping system utilises the RGB-D camera to construct a dense heightmap of the immediate surroundings of the robot, as the robot moves around old data is erased to make way for new data. The size and resolution of the heightmap is adjustable to the available memory and computational power. The heightmap system is further covered in chapter 4.

The foot placement optimisation system takes the heightmap as input and produces another map of equal size to the heightmap, this new map is the score map and is found by assigning a score to each cell of the heightmap. The

score is dependant on how stable of a place the cell would be for the robot to place its feet. The score map can then be used to evaluate, and adjust if necessary, the initial foot placement proposed by the motion system. The foot placement optimisation system is further covered in chapter 6.

All the movement of the robot is handled by the motion system, it is comprised of a gait state machine, a foot target proposal system, a foot arc generator, and IK. The gait state machine selects the swinging and supporting legs during for each step, to achieve a tripod gait. The target proposal system proposes an initial target for all the feet based on the current step parameters. Simple linear motion is not acceptable for the swinging feet, thus the foot arc generator produces a movement vector based on the remaining distance to a foot's destination, which if followed, results in a arc like motion to the destination. Finally to execute any movements, positions must be converted to angles, and velocity must be converted to angular rate, thus, and IK component is required. The motion system is further covered in chapter 5.

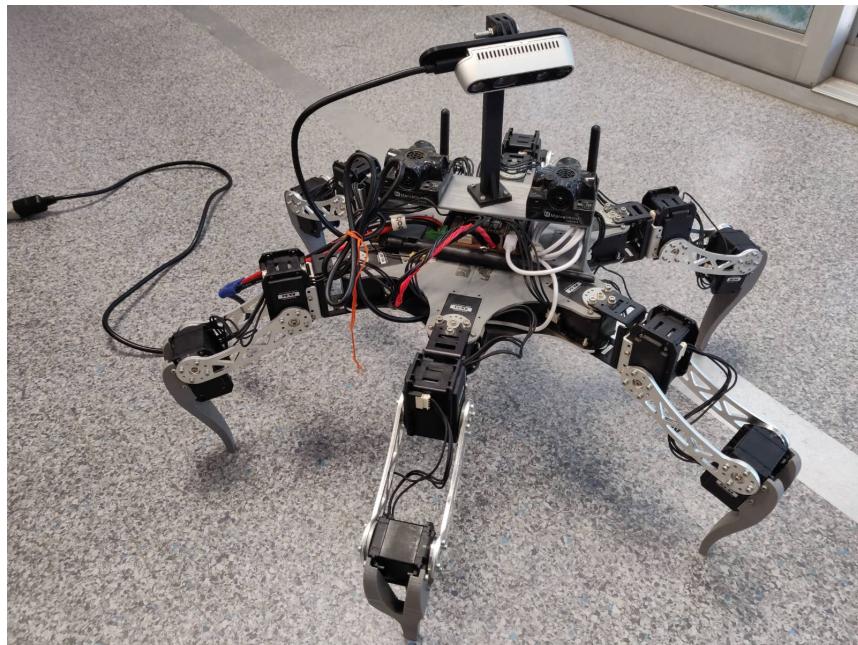


Figure 3.4: Physical Hexapod

# Chapter 4

## Mapping

For accurate foot placement and localisation purposes the robot makes use of two maps, a sparse map covering a large area, and a dense map covering a small area around the robot. The primarily use of the sparse map is for localisation and extracting pose data, i.e. orientation, velocity and rate. While the dense map is used to analyse the terrain and find an appropriate point to place the three supporting feet. It is possible to also use the sparse map for autonomous navigation, however this use case is not covered in this paper. This chapter covers the design of the mapping system.

The localisation, sparse mapping and pose estimation is handle by ORB-SLAM3 as described in Campos *et al.* (2021). Since ORB-SLAM3 is not a system designed by the author, its design will not be covered in this chapter. Implementation and operation details will however be covered in chapters 7 and ??.

### 4.1 Projection

In order to generate a heightmap from a RGB-D image, it is first required to project the RGB-D image into 3D space, this is necessary because a heightmap is essentially a 3D environment, that can be represented as a image due to the assumption of purely convex geometry.

The camera can be described by its intrinsic and extrinsic parameters. Extrinsic parameters characterise the cameras position in 3D space, and intrinsic parameters characterise the relationship between the image plane 3D space, assuming the camera is at the world origin and an zero rotation. Hartley and Zisserman (2003).

Refer to figure 4.1 as a visual aid regarding projection. Note that this figure is drawn from the perspective of projecting from the image plane into the world, if the objective was to project from the world onto the image plane the projection center and image plane would swap places, causing the image to be inverted, thus, this figure assumes that the image rotation has been corrected.

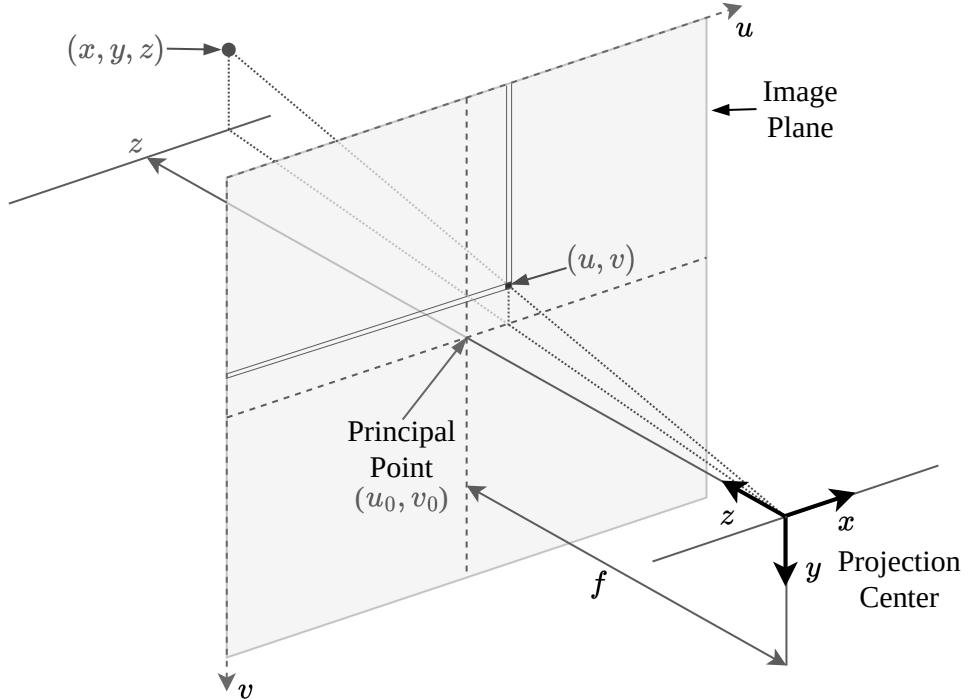


Figure 4.1: Camera Projection

Together the extrinsic and intrinsic matrices form the projection matrix, as shown in equation 4.1,

$$\mathbf{P} = \mathbf{K} [\mathbf{R} \ \mathbf{T}] \quad (4.1)$$

where  $\mathbf{K}$  is the intrinsic matrix and  $[\mathbf{R} \ \mathbf{T}]$  the extrinsic matrix, these are described in equation 4.3 and 4.4.

The projection matrix can be used to project a point on the image plane into world space as shown in equation 4.2.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.2)$$

where  $u, v$  are the pixel coordinates on the image plane and  $x, y, z$  are the coordinates in world space.

$$\mathbf{K} = \begin{bmatrix} \alpha_x & \gamma & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.3)$$

where the focal length is represented by,

$$\alpha_x = f \cdot m_y$$

$$\alpha_y = f \cdot m_x$$

with  $m_x$  and  $m_y$  being the inverse of the width and height of a image plane pixel,  $f$  the focal length and  $u_0, v_0$  being the principal point, optimally the center of the image plane. The skew coefficient,  $\gamma$ , is often, and in this case, 0.

The extrinsic matrix is as shown below,

$$[\mathbf{R} \ \mathbf{T}] = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{T}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \quad (4.4)$$

where  $\mathbf{R}$  characterises the camera's heading in world space and  $\mathbf{T}$  the world origin expressed in the camera coordinate frame.

For ease of preprocessing points are first projected into the camera coordinate frame, in other words, the extrinsic matrix is omitted from equation 4.2. The resultant matrix equation is shown in equation 4.5.

$$\begin{bmatrix} u \\ v \\ 1 \\ 1/z \end{bmatrix} = \frac{1}{z} \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (4.5)$$

From equation 4.5  $x, y, z$  are found to be show in equation 4.6 to 4.8.

$$z = I_{u,v} \quad (4.6)$$

$$x = \frac{z(u - u_0)}{\alpha_x} \quad (4.7)$$

$$y = \frac{z(v - v_0)}{\alpha_y} \quad (4.8)$$

where  $I_{u,v}$  is the depth image value at pixel coordinates  $u, v$ . For later use the variable  $\mathbf{p}^c$  is defined as in equation 4.9,

$$\mathbf{p}^c = [x \ y \ z]^T \quad (4.9)$$

## 4.2 Map Buffer

Once the depth image is projected into map space, their x and y coordinates are used as indices to write their z value into the heightmap. The heightmap is stored in a 2D circular buffer, this means that as the robot moves around, old map data is erased to make room for new data. This structure is illustrated in figure 4.2.

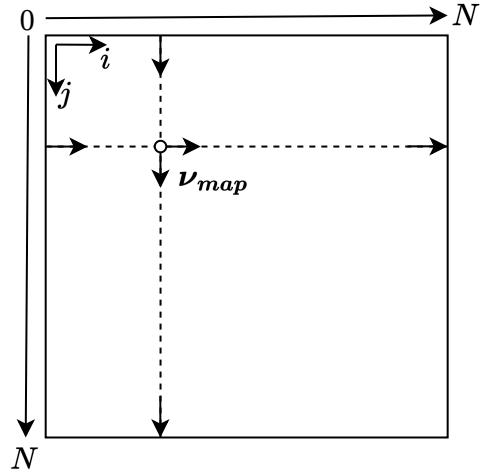


Figure 4.2: Memory Diagram

The heightmap buffer,  $\mathbf{M}$ , is of size  $N \times N$  and is indexed by  $i, j$ . In order to write the height value into the heightmap, the camera point in camera space,  $\mathbf{p}^c$ , is translated into map space,  $\mathbf{p}^{\mathcal{M}}$ , as shown in equation 4.10,

$$\mathbf{p}^{\mathcal{M}} = T_{\mathcal{CM}}(\mathbf{p}^c) \quad (4.10)$$

see appendix A for transformation definitions.

# Chapter 5

## Motion

The basic operation of the motion system is as follows, first the robot is commanded to walk in a certain direction, at a certain speed and body height. These commands are sent from the base station to Jetson Nano on the robot, the motion controller node on the Jetson Nano then sends these commands to the gait state machine node, at a fixed frequency. The gait state machine uses the received direction stride length to generate leg states (swinging or supporting) and the ideal final position of each foot. These states and positions are sent back to the motion controller node where the positions are adjusted based on the heightmap data to ensure stable footing. The leg states and adjusted feet positions are then sent to the servo controller node, this node controls the servos to move the robots feet to their final positions, either in a arc or linearly, depending on their state (swinging or supporting).

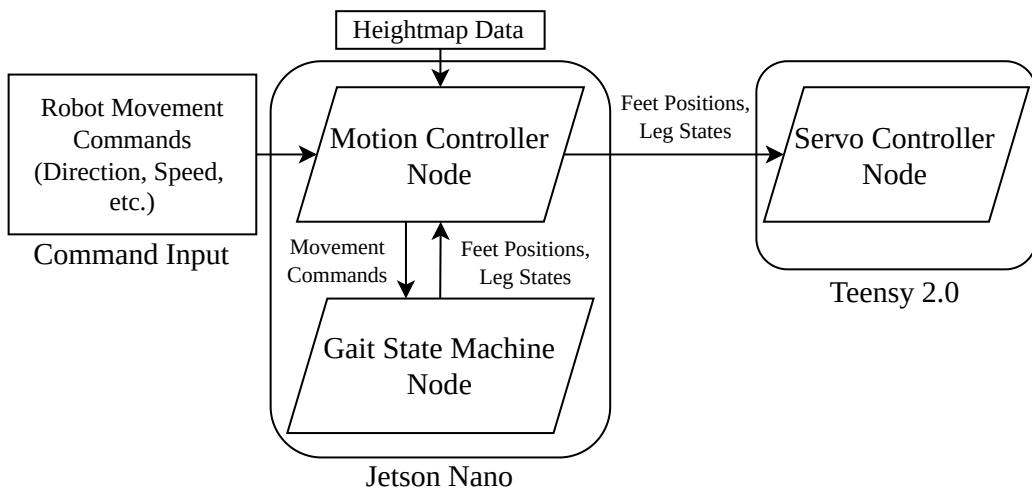


Figure 5.1: Motion System Overview

## 5.1 Gait State Machine

The state machine used to realise the tripod gait used in the robot is quite simple, comprised of only two states, stepping and resting, as can be seen from figure 5.2. Table 5.1 defines the actions that should be taken during each state.

The primary computation done by this state machine is calculating which legs are supporting and which are swinging, which occurs on entering the "Stepping" state. This function is described in section 5.1.1.

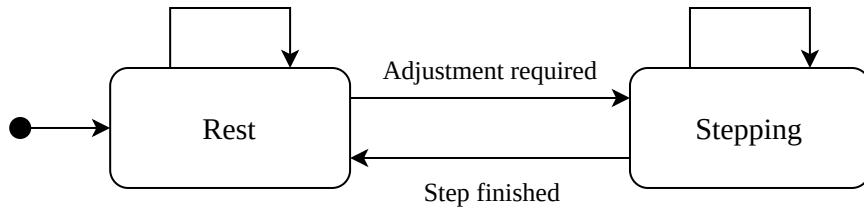


Figure 5.2: Gait State Machine

Rest State Definition	
Enter Condition	Has all feet reached their targets?
On Entering	Set all leg states as supporting.
While Active	Do nothing

Stepping State Definition	
Enter Condition	Is there a mismatch between feet targets and current position?
On Entering	Calculate and set the leg states based on walking direction, see section 5.1.1. Choose and optimise nominal targets for the current step, see chapter 6
While Active	Adjust feet targets based on direction, stride length and robot height, see chapter 6

Table 5.1: State Definitions

### 5.1.1 Choosing The Supporting And Swinging Legs

The robot body is divided up into sextants, centered around the nominal leg positions. When calculating the swinging legs it is first determined in which sextant the movement direction vector falls, this is called the active sextant. The leg related with the active sextant, and the two opposite, are then chosen as swinging, with the remaining three legs chosen as supporting. The states of the legs are encapsulated by the boolean array,  $\mathbf{S}_i$ , defined by equation 5.1, where a 1 indicates swinging and 0 supporting.

$$\mathbf{S}_{i-\xi} = [i \text{ is even}] \quad (5.1)$$

where  $\xi \in i$  is the active sextant/leg number, and,

$$i = [\![0, \dots, 5]\!]$$

Figure 5.3 illustrates and example with sextant 1 being active. Thus legs 1, 3 and 5 are swinging, while legs 0, 2 and 4 are supporting.

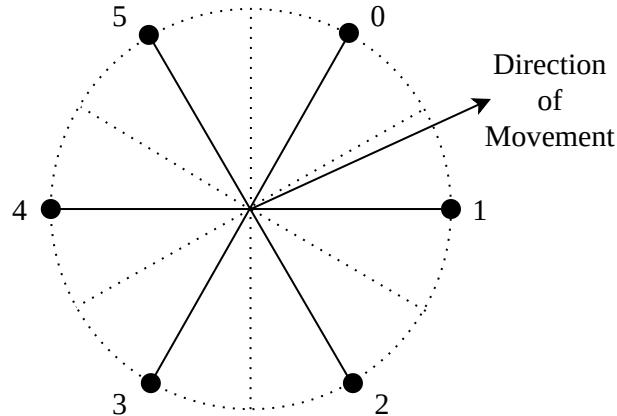


Figure 5.3: Leg sextants, with sextant 1 being active.

This of course would not be sufficient to define a walking gait, as at the end of each step  $\mathbf{S}_i$  does not invert. Thus an additional step after equation 5.1 is added. The current horizontal length of leg  $\xi$ , defined as  $l_\xi$ , is compared to its nominal horizontal length,  $L_\xi$ . If  $l_\xi > L_\xi$ , invert  $\mathbf{S}_i$ . As shown in equation 5.2.

$$\mathbf{S}_{i-\xi} = \begin{cases} \mathbf{i} \setminus \mathbf{S}_{i-\xi} & l_\xi > L_\xi \\ \mathbf{S}_{i-\xi} & l_\xi \leq L_\xi \end{cases} \quad (5.2)$$

### 5.1.2 Choosing nominal foot positions

Once the active and inactive legs have been selected it is required to choose nominal targets for all the feet. Sections 5.1.2.1 and 5.1.2.2 describe the process for finding the support and swinging leg targets; It should be noted that target matrices, denoted by  $\mathbf{t}$ , are common between these two sections, but are computed differently depending on whether the leg is swinging or supporting.

#### 5.1.2.1 For Supporting Legs

The supporting leg nominal targets are chosen with using equations 5.3 and 5.4, with reference to figure 5.4.

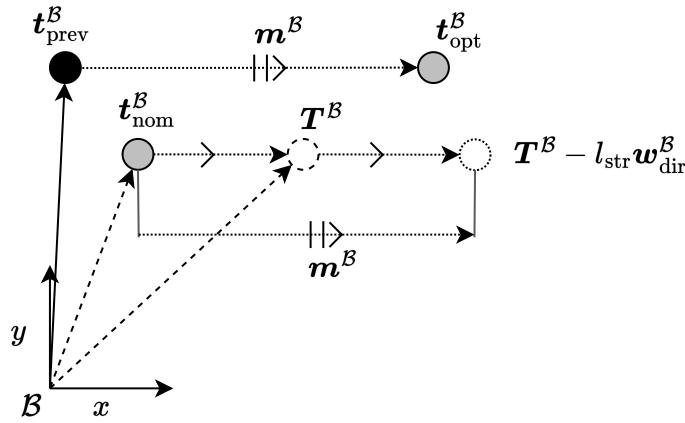


Figure 5.4: Supporting target choosing diagram

First the required move vectors,  $\mathbf{m}^{\mathcal{B}}$ , are calculated in equation 5.3,

$$\mathbf{m}^{\mathcal{B}} = (\mathbf{T}^{\mathcal{B}} - l_{\text{str}} \mathbf{w}_{\text{dir}}^{\mathcal{B}}) - \mathbf{t}_{\text{nom}}^{\mathcal{B}} \quad (5.3)$$

where  $\mathbf{T}^{\mathcal{B}}$  contains the constant rest positions the feet,  $l_{\text{str}}$  is the desired stride length,  $\mathbf{w}_{\text{dir}}^{\mathcal{B}}$  contains the desired walk directions, and  $\mathbf{t}_{\text{nom}}^{\mathcal{B}}$  contains the nominal targets as calculated in equation 5.6.

Then the optimised targets,  $\mathbf{t}_{\text{opt}}^{\mathcal{B}}$ , are calculated as the addition of the previous targets and the move vector in equation 5.3,

$$\mathbf{t}_{\text{opt}}^{\mathcal{B}} = \mathbf{t}_{\text{prv}}^{\mathcal{B}} + \mathbf{m}^{\mathcal{B}} \quad (5.4)$$

where  $\mathbf{t}_{\text{prv}}^{\mathcal{B}}$  contains the previous, optimised, targets. Note that equation 5.4 does not include the optimisation function **XXXX**, this is because the supporting feet will not move relative to the terrain, and thus do not need optimising.

### 5.1.2.2 For Swinging Legs

The swinging leg nominal targets are chosen with reference to figure 5.5.

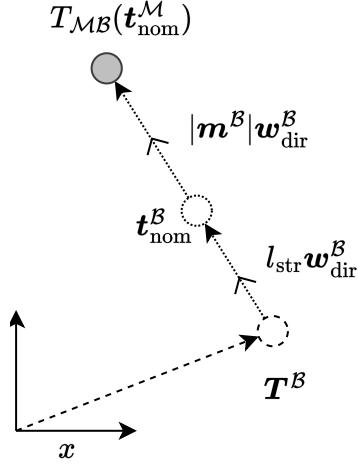


Figure 5.5: Swinging target choosing diagram

The nominal targets in map space,  $t_{\text{nom}}^{\mathcal{M}}$ , are calculated as in equation 5.5,

$$\mathbf{t}_{\text{nom}}^{\mathcal{M}} = \mathbf{T}^{\mathcal{B}} + (l_{\text{str}} + |\mathbf{m}^{\mathcal{B}}|) \mathbf{w}_{\text{dir}}^{\mathcal{B}} \quad (5.5)$$

where  $\mathbf{T}^{\mathcal{B}}$  contains the constant rest positions the feet,  $l_{\text{str}}$  is the desired stride length,  $\mathbf{w}_{\text{dir}}^{\mathcal{B}}$  contains the desired walk directions.  $\text{avg } \mathbf{m}^{\mathcal{B}}$  is the average move vector

For use in equation 5.3, the swinging legs nominal targets,  $t_{\text{nom}}^{\mathcal{B}}$ , are calculated as in equation 5.6,

$$\mathbf{t}_{\text{nom}}^{\mathcal{B}} = \mathbf{T}^{\mathcal{B}} + l_{\text{str}} \mathbf{w}_{\text{dir}}^{\mathcal{B}} \quad (5.6)$$

Note that  $T_{ML}(t_{\text{nom}}^{\mathcal{M}})$  is quite similar to  $t_{\text{nom}}^{\mathcal{B}}$ , the difference is that  $t_{\text{nom}}^{\mathcal{M}}$  is set such that it will align with  $t_{\text{nom}}^{\mathcal{B}}$  at the end of the step. Thus, the omission of the average supporting move vector,  $\text{avg } \mathbf{m}^{\mathcal{B}}$ , in calculating  $t_{\text{nom}}^{\mathcal{B}}$  in equation 5.6. Meaning at until the end of the step  $t_{\text{nom}}^{\mathcal{M}}$  and  $t_{\text{nom}}^{\mathcal{B}}$  will point to different positions if world space. This is due to needing to account for body movement in world space, but not in local space.

## 5.2 Foot Motion

When taking a step the foot can not simply be moved to its destination in a straight line, as doing so will cause the foot to be dragged on the terrain, impeding the movement of the robot. Thus it is required to move the foot in an arc like motion to clear any obstacles that might be in its path.

### 5.2.1 Existing System

The existing system will, at the start of each step, compute an arc for each foot to follow, this arc is then sent to the servo controller to be executed. While efficient and effective in ideal conditions, this method of defining the arc has poor performance when considering external influences. If for example the robot has to adjust the final target of its feet mid step, this arc would have to be recomputed in its entirety, thus leading to possible performance concerns.

In addition to this Text, the current system is designed with the assumption that the starting position of the foot is grounded, thus if the arc is recomputed mid step the arc will be undesirable, as it will rise with the desired step height for a second time. This is illustrated in Figure 5.6.

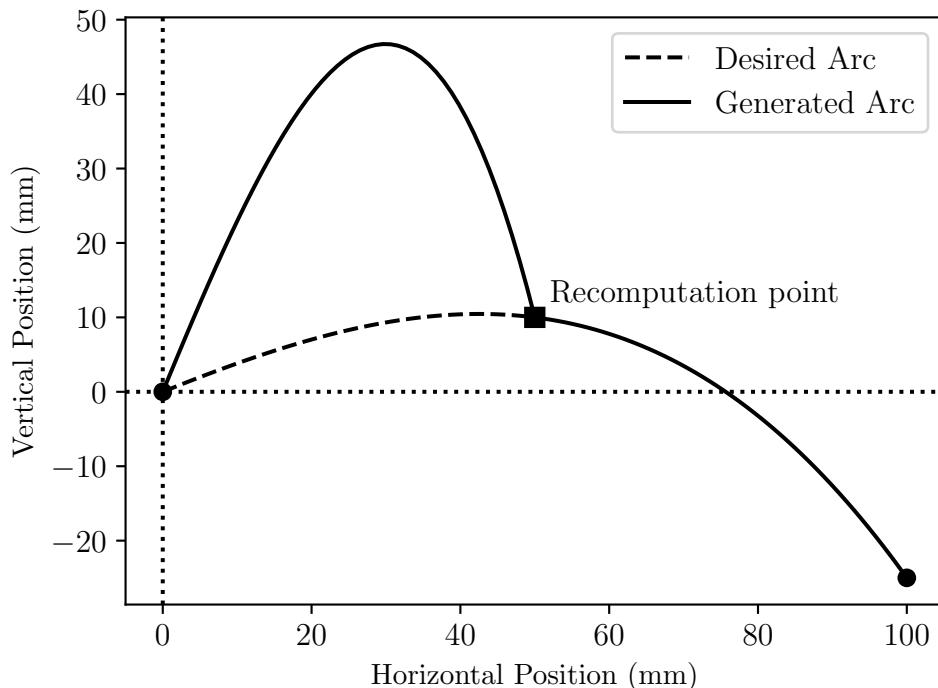


Figure 5.6: Existing arc recomputation problem

### 5.2.2 Improved System

The improved system solves this problem by utilising a flow function. During a step, this function will continuously calculate the direction that the foot must move to reach its destination. Thus this system is resilient to external disturbances and is capable of adjusting to varying destination and step height requirements.

The flow field is designed to first move the foot vertically upwards until horizontal coplanar with the destination, and then to follow a arc to the destination with a defined step height, this can be adjusted to make the arc start before or after coplanar. The step height can be adjusted at any point in time and the flow field will adjust accordingly. Figure 5.7 illustrates the field function and is described in section 5.2.2.1.

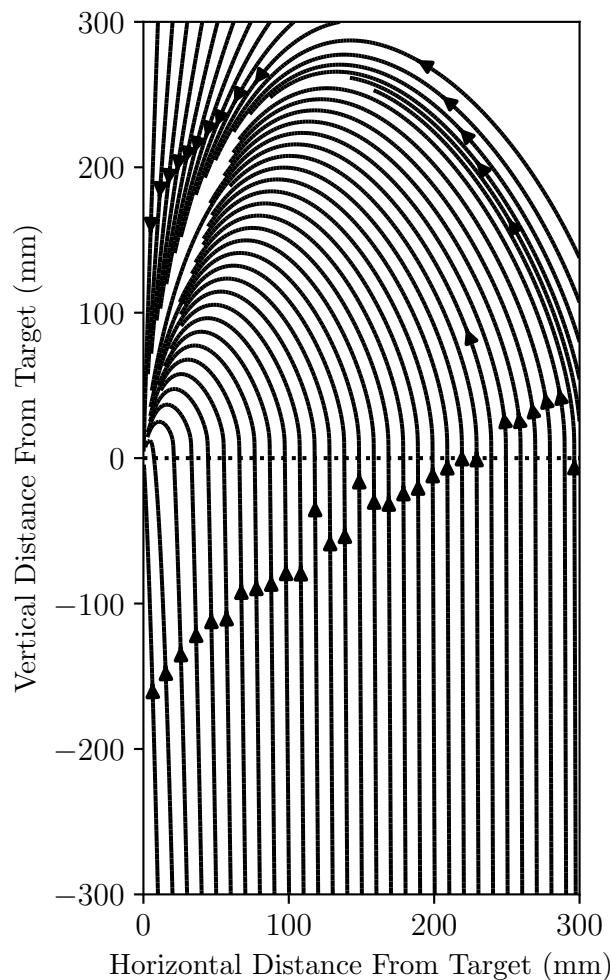


Figure 5.7: End effector movement path

### 5.2.2.1 Flow Function Description

The flow function,  $\rho(x, y)$ , uses the gradient function of a parabola passing through the point  $[0, 0]$  and  $[x, y]$  as a basis, where point  $[x, y]$  is the current point that is being evaluated and  $x$  is the horizontal distance between the destination and the current point and  $y$  the vertical distance. The final function is described by equations 5.7 to 5.10, for the process of designing the flow function please see appendix ??.

$$\begin{aligned}\rho(x, y) &= \frac{\delta}{\delta x \delta y} f_a(x, y)x^2 + f_b(x, y)x + C \\ &= 2f'_a(x, y)x + f'_b(x, y)\end{aligned}\quad (5.7)$$

where,

$$f'_a(x, y) = -\left|\frac{v_h}{x}\right| - |S(y)| \quad (5.8)$$

$$f'_b(x, y) = \frac{y}{x} - f_a(x, y) \quad (5.9)$$

with  $v_h$  being the variable describing the step height and  $S(y)$  being a sigmoid like function responsible for the initial vertical rise.  $S(y)$  is defined in Equation 5.10.

$$S(y) = \frac{0.515(y - q)}{1 + |y - q| - 0.505} \quad (5.10)$$

where  $q$  is the variable that determines at which vertical displacement the leg path transitions from primarily an vertical motion to an arc motion. Figure 5.8 illustrates the sigmoid like function for different values of  $q$ .

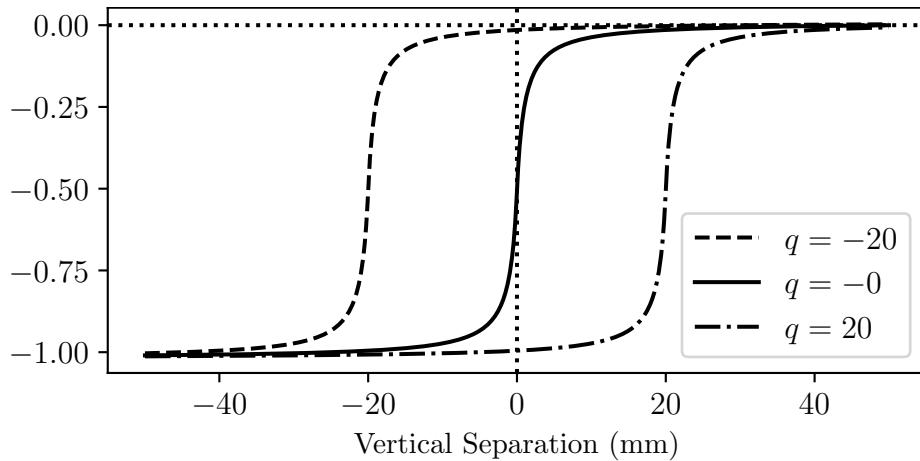


Figure 5.8: Sigmoid Like

Note that the 0.515 and 0.505 values in equation 5.10 are set to make its output range from roughly -1 to 0 over the active range.

## 5.3 Kinematics

When commanding a foot position, the servo controller requires a function to calculate servo angles. While the foot arc planner, see section 5.2, requires the current position of the feet to function. The IK and FK functions described in this section provide this functionality.

### 5.3.1 Coordinate Frames

The coordinate frames relevant to the kinematics of the robot are the body coordinate frame,  $\mathcal{B}$ , and the leg frame,  $\mathcal{L}_i$ . All foot targets/positions are specified in the body coordinate frame, while the kinematic systems operate in the leg coordinate frame. Thus a conversion from the body to leg frames is required. Figure 5.9 show the body and multiple leg coordinate systems. The

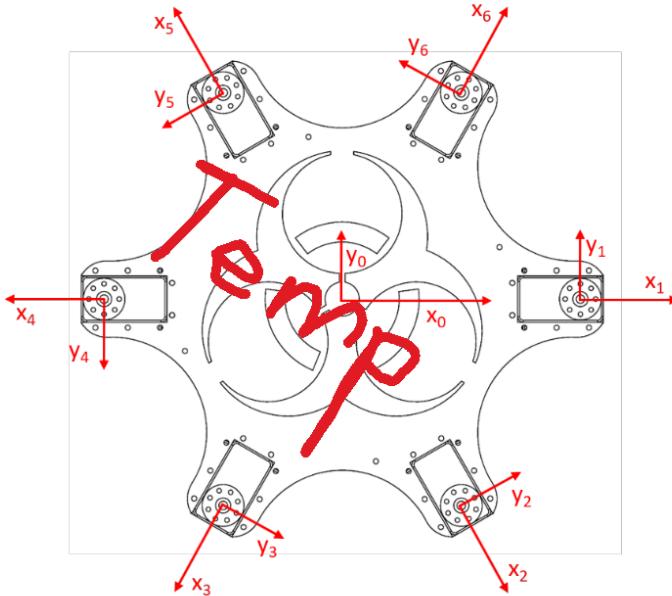


Figure 5.9: Body and leg coordinate frames.

leg coordinate systems are simply rotated and shifted from the body coordinate system. This transformation,  $T_{\mathcal{BL}_i}(\mathbf{x}^{\mathcal{B}})$  is defined by equation 5.11,

$$\begin{aligned}\mathbf{x}^{\mathcal{L}_i} &= T_{\mathcal{BL}_i}(\mathbf{x}^{\mathcal{B}}) \\ &= \mathbf{Q}_i^{\mathcal{B}} \cdot \mathbf{x}^{\mathcal{B}} \cdot \mathbf{Q}_i^{\mathcal{B}^{-1}} - \mathbf{R}_i\end{aligned}\quad (5.11)$$

where the  $\mathcal{L}_i$  is the coordinate frame of leg  $i$ ,  $\mathbf{Q}_i^{\mathcal{B}}$  is the rotation of said coordinate frame and  $\mathbf{R}_i^{\mathcal{B}}$  is the root position of said leg coordinate frame in the

body coordinate frame. For more detail on this transformation, such as the values of  $\mathbf{Q}_i^B$  and  $\mathbf{R}_i^B$ , please see appendix A.

Now that positions can be transformed into the leg coordinate frame the kinematic equations, as described in the following sections, can be applied. The kinematic equations are defined with reference to the variables in the leg frame as shown in figure 5.10.

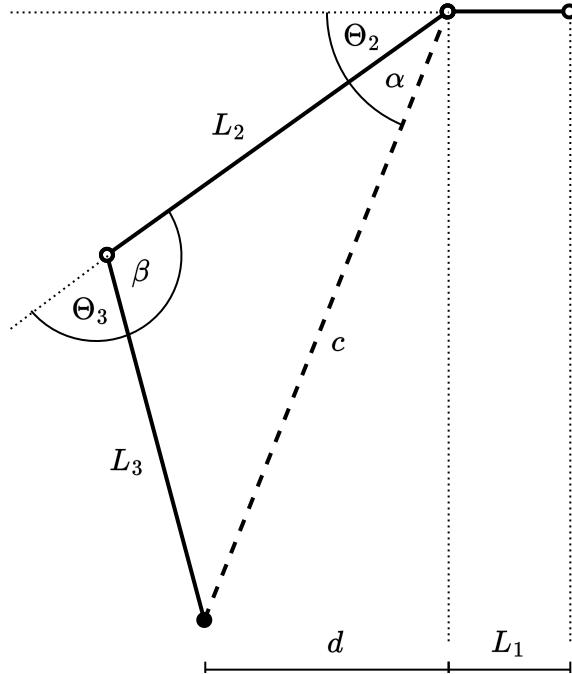


Figure 5.10: Leg coordinate frame with kinematic variables.

### 5.3.2 Inverse Kinematics (IK)

The IK function calculates the leg servo angles,  $\Theta = [\Theta_1, \Theta_2, \Theta_3]^T$  required to move the foot to the given target position vector,  $\mathbf{t} = [x_t, y_t, z_t]^T$ . Equation 5.12 describes the IK function.

$$\Theta(x_t, y_t, z_t) = \begin{bmatrix} \arctan\left(\frac{x_t}{y_t}\right) \\ \frac{\pi}{4} - \alpha - \arctan\left(\frac{y_t}{d - L_1}\right) \\ \frac{\pi}{2} - \beta \end{bmatrix} \quad (5.12)$$

where  $\alpha$ ,  $\beta$ ,  $c$  and  $d$  are calculate as shown in equations 5.13 to 5.16. For derivations of these variable please see ??.

$$\alpha = \arcsin\left(\frac{L_3 \sin \beta}{c}\right) \quad (5.13)$$

$$\beta = \arccos\left(\frac{L_1^2 + L_2^2 - c^2}{2L_1L_2}\right) \quad (5.14)$$

$$c = \sqrt{(d - L_1)^2 + z_t^2} \quad (5.15)$$

$$d = \sqrt{x_t^2 + y_t^2} \quad (5.16)$$

### 5.3.3 Forward Kinematics (FK)

The FK function calculates the position vector of a foot,  $\mathbf{p}_f = [x_f, y_f, z_f]^T$ , given the current angles of the leg servos,  $\boldsymbol{\theta} = [\theta_1, \theta_2, \theta_3]^T$ .

$$\mathbf{p}_f(\theta_1, \theta_2, \theta_3) = \begin{bmatrix} d \cos \theta_1 \\ d \sin \theta_1 \\ L_2 \sin \theta_2 + L_3 \sin (\theta_2 + \theta_3) \end{bmatrix} \quad (5.17)$$

where  $d$  is calculated as shown in in equation 5.18.

$$d = L_1 + L_2 \sin \theta_2 + L_3 \sin (\theta_2 + \theta_3) \quad (5.18)$$

### 5.3.4 Angular Rate

To move a foot on a desired path it is important to not only know the absolute angle of the three leg servos, but also the angular rates of all three servos. If the servos are all moved at the same rate, the shape of the path that the foot follows will not be linear, but rather dependant on the current foot position. This is undesirable, thus equations 5.19 define the derivative of the IK equations (5.12), i.e. the angular rate, given the target movement speeds of a foot,  $\dot{x}$ ,  $\dot{y}$  and  $\dot{z}$ .

$$\boldsymbol{\omega}(\dot{x}, \dot{y}, \dot{z}) = \begin{bmatrix} \frac{-x\dot{y} + y\dot{x}}{x^2 + y^2} \\ \frac{[(L_1 - d)\dot{z} + z\dot{d}] \alpha + [(L_1 - d)^2 + z^2] \arctan\left(\frac{L_1 - d}{z}\right) \dot{\alpha}}{(L_1 - d)^2 + z^2} \\ -\dot{\beta} \end{bmatrix} \quad (5.19)$$

where  $\dot{\alpha}$ ,  $\dot{\beta}$ ,  $\dot{c}$  and  $\dot{d}$  as shown in equations 5.20 to 5.23.

$$\dot{\alpha} = \frac{L_3 [c \cos(\beta) \dot{\beta} - \sin(\beta) \dot{c}]}{c^2 \sqrt{-\frac{L_3^2 \sin^2(\beta)}{c^2} + 1}} \quad (5.20)$$

$$\dot{\beta} = \frac{2c\dot{c}}{L_2 L_3 \sqrt{4 - \frac{(L_2^2 + L_3^2 - c^2)^2}{L_2^2 L_3^2}}} \quad (5.21)$$

$$\dot{c} = \frac{-(L_1 - d)\dot{d} + z\dot{z}}{\sqrt{(L_1 - d)^2 + z^2}} \quad (5.22)$$

$$\dot{d} = \frac{x\dot{x} + y\dot{y}}{\sqrt{x^2 + y^2}} \quad (5.23)$$

---



---



---

# Chapter 6

## Foot Placement Optimisation

The best possible anchor points for the supporting feet, given an initial point from the walking gait state machine, must be found based on the heightmap. This is done by applying various scores to the heightmap and selecting a point with the best score within an allowable deviation from the initial point received from the gait state machine.

### 6.1 Scoring

The scores considered are the gradient and proximity scores. The gradient score aims to reject points with high gradients while the proximity score rejects points close to other parts of the terrain with steep inclines, i.e. to reject points inside holes or close to ledges.

#### 6.1.1 Gradient Score

The gradient score is simply taken as the slope of the terrain at the current point. The aim of this score is to prevent the robot from slipping due to selecting anchor points with too steep of a gradient.

As the heightmap slope is not defined by a known function, the gradient is calculated using the Sobel operator (Sobel (2014)), a combination of a central finite difference and a smoothing operator.

Equation 6.1 and 6.2 describe the two separable x and y direction kernels,  $G_x$  and  $G_y$ , of the Sobel operator. These kernels are a combination of central finite difference and smoothing operator, see Appendix ?? for a breakdown. Equation 6.3 combines  $G_x$  and  $G_y$  to produce the gradient score,  $S_g$ . Note that these equations represent a single scalar result of the Sobel operator using the Frobenius inner product, Horn and Johnson (2012). If evaluated over the whole heightmap this is equivalent to convolution.

$$G_x(x, y) = \left\langle \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix}, \mathbf{h}_{i,j} \right\rangle_F \quad (6.1)$$

$$G_y(x, y) = \left\langle \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, \mathbf{h}_{i,j} \right\rangle_F \quad (6.2)$$

$$S_g(x, y) = C_g \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (6.3)$$

where,

$$i = \llbracket x - 1, x + 1 \rrbracket$$

$$j = \llbracket y - 1, y + 1 \rrbracket$$

with  $C_g$  being the weighting constant associated with the gradient score,  $S_g$ .

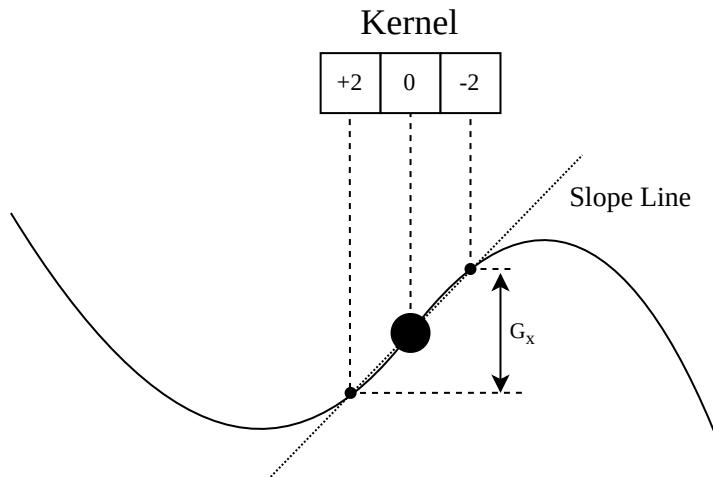


Figure 6.1: Slope Score

### 6.1.2 Terrain Proximity Score

The proximity score aims to bias the selected anchor point away from points near steep inclines in the terrain. This score is defined as the average of the height difference of the current point weighted by the gaussian kernel  $\mathbf{K}$ , of size  $n$  by  $n$ . The distance around inclines that is rejected depends on the chosen size of the kernel  $\mathbf{K}$ , the standard deviation of  $\mathbf{K}$  and the height differences. This score is described in Equation 6.4.

$$S_p(x, y) = C_p \left| \frac{\sum \langle \mathbf{K}, (\mathbf{h}_{i,j} - \mathbf{h}_{x,y}) \rangle_F}{n^2} \right| \quad (6.4)$$

where,

$$i = [\lfloor \frac{1}{2}n \rfloor, x + \lfloor \frac{1}{2}n \rfloor]$$

$$j = [\lfloor \frac{1}{2}n \rfloor, y + \lfloor \frac{1}{2}n \rfloor]$$

with  $x$  and  $y$  being the indices of the cell whose score is currently being evaluated,  $\mathbf{h}$  the heightmap and  $C_p$  the score's weighting constant.

A diagrammatic, sliced, representation of the proximity score can be seen in Figure 6.2. The slice is taken along the x axis of the heightmap.

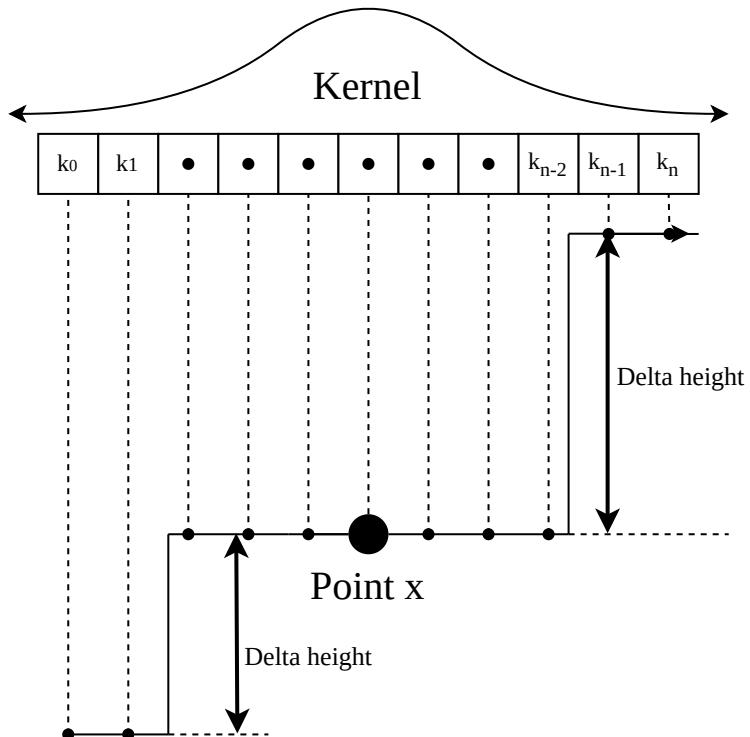


Figure 6.2: Proximity Score Diagrammatic Representation

### 6.1.3 Constraints

It is important to constrain the possible anchor points to confer to the stability triangle, meaning that the centre of mass of the robot must be inside the triangle formed by the three anchor points. Additionally, it is important that the points selected are not too far away, both in the horizontal and vertical direction for the robot to reach.

## 6.2 Method of adjustment

Once the heightmap has been processed into the score map, which is done by adding the slope score and terrain proximity score, the point with the best score must be found for every initial anchor point. The resolution of the heightmap is not very high and the adjusted anchor point is not allowed to deviate too far from the initial anchor point, additionally due to the parallel nature of the heightmap generation it is possible to score the entire heightmap with minimal cost. Thus, it was decided to not employ an optimisation algorithm, such as gradient decent or Bayesian, but rather to use a radial search algorithm. This algorithm progressively expands its searching radius over the square score grid until a valid score is found, at which point it terminates, thus ensuring that the closest valid point to the initial anchor point is selected. See Figure ?? for a diagram representing the search pattern for a 5 grid squares search area. Note that this search pattern will become inaccurate for large search areas, as the pattern steps in a square manner. This however is not much of a concern for smaller search areas.

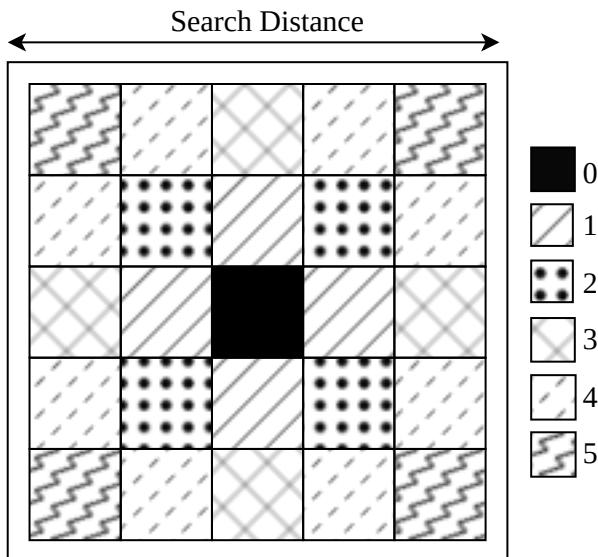


Figure 6.3: Radial search, shown for a 5 cell search diameter.

### 6.3 Floor Height Adjustment

After the horizontal position of the feet have been optimised, the height of the foot is simply set to the height of the terrain at the target horizontal coordinates. This ensures that the robots body remains level and at a constant global height.

It is however important to adjust the reference floor height that the robot uses, if this is not done, then the robot will be incapable of surmounting terrain higher than its commanded standing height, as the robot will simply maintain said height and walk into the tall terrain. There are various methods to choosing a floor height, from using a time of flight sensor underneath the robots main body to using height data from the heightmap.

The method that was implemented in this paper uses the average of three highest target positions out of all the robots legs. This allows the robot to preemptively increase or decrease its floor height as it targets to step onto higher or lower terrain.

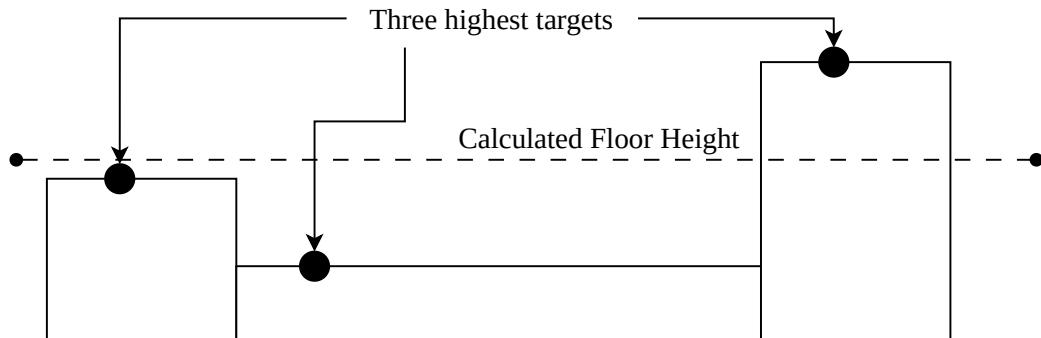


Figure 6.4: Floor height diagram.

While this method works well to preemptively adjust height and to optimise the body height against foot height, it does present one problem. If there is a tall piece of terrain in the path of the robots body without any feet positioned on it, then the robot will not adjust its body height to clear this piece of terrain. Thus, the bounding box of the robots body, excluding the legs and feet, is checked against the heightmap and if anywhere in the bounding box is higher than the previously set floor height, the floor height is adjusted to this new value.

# Chapter 7

## Hardware Implementation

This chapter describes the process of implementing the system built in previous chapters on the physical robot using ROS and parallel computation.

### 7.1 ROS Nodes

There are various ros nodes spread out across the base station, the Jetson Nano and the Teensy Microcontroller Unit (MCU), figure 7.1 provides and overview of the nodes and how they communicate with each other. Section 7.2.1 and 7.2.2 go into detail on the publishers and subscribers on board the robot and on the base station. Figure 7.1 provides and overview of the ROS nodes and communication used to construct the system. Further detail pertaining to communication using ROS is provided in section 7.2.

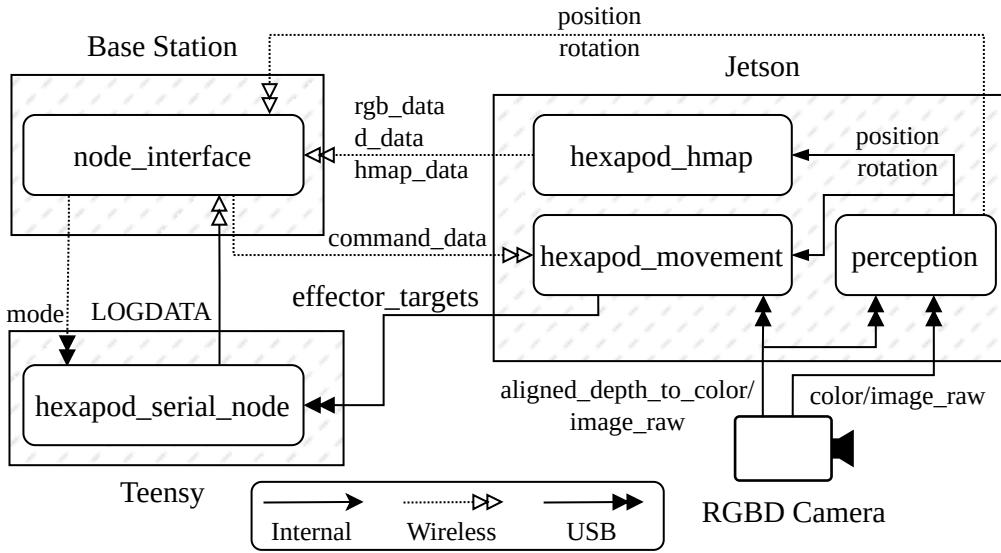


Figure 7.1: ROS nodes and communication.

## 7.2 ROS communication

This section provides a detailed description of the ROS communication scheme used in the system as shown in figure 7.1.

### 7.2.1 Base Station

The base station wirelessly communicates with the robot to send commands and receive data. For a description of the base station publishers see table 7.1 and for its subscribers see 7.2. Table 7.7 describes data types used.

Base Station Publishers			
Name	Data Type	Description	Frequency
command_data	HexapodCommands	Various robot command parameters.	On change
mode	Int32	Specifies the operating mode of the robot.	On change.

Table 7.1: Base station publishers

The robot currently only has two operating modes, torque cutoff mode and walking mode. The torque cutoff mode is the initial mode the robot is in, while in this mode the leg servos disable all torque control, thus entering a relaxed state. In walking mode the robot walks in the commanded direction whilst optimising its foot positions according to the terrain. If the robot encounters a piece of terrain for which no optimisation can be found the human controller will have to adjust the walking direction from the base station.

Base Station Subscribers		
Name	Data Type	Description
rgb_data	Image	The processed color image from the robot.
d_data	Image	The processed color depth from the robot.
hmap_data	Image	The heightmap generated on the robot.
LOGDATA	String	General logs from the robot.

Table 7.2: Base station subscribers

The only subscribers present on the base station are the processed camera images, heightmap and logs. These are all used to provide a interface from where the operator can control the robot.

### 7.2.2 On Board

The hexapod has two computational units on board, first the Jetson Nano, which handles all high level operations, including heightmap generation and scoring, foot optimisation, maintain a walking gait and localisation using ORB-SLAM3. Secondly a Teensy2.0 MCU handles low level operations, including interpolating feet movement paths and servo control. Table 7.3 to 7.6 describe the ROS publishers and subscribers present on these two computational units.

Jetson Publishers			
Name	Data Type	Description	Frequency
effector_targets	EffectorTargets	Data indicating which feet to move where, and what type of interpolation to use.	On change
rgb_data	Image	The processed color image from the RGB-D camera.	15Hz.
d_data	Image	The processed depth image from the RGB-D camera.	15Hz.
hmap_data	Image	The heightmap generated on the robot.	15Hz.
position	Vector3	The localised position of the robot	15Hz
rotation	Quat	The localised rotation of the robot	15Hz

Table 7.3: Jetson publishers

The camera data, heightmap data, position and rotation are published for display at the base station. While the effector targets are published for use on the Teensy to move the robot's feet to the optimised positions.

As can be seen from table 7.4 the only subscribers required on the Jetson is the raw camera feed, for constructing the heightmap, and the commands from the base station.

Table 7.5 show that the Teensy publishes the current feet positions these are the positions calculated through FK. Additionally log data is also published for use on the base station.

Lastly, from table 7.6 it can be seen that the Teensy subscribes to the command data, effector targets and mode. The walking speed component

Jetson Subscribers		
Name	Data Type	Description
command_data	HexapodCommands	Commands from the base station.
color/image_raw	Image	Raw color image from the RGB-D camera.
aligned_depth_to_color/image_raw	Image	Raw depth image from the RGB-D camera.

Table 7.4: Jetson subscribers

Teensy Publishers			
Name	Data Type	Description	Frequency
LOGDATA	String	General logs.	10Hz
effector_current_position	Eigen::Vector3d	Current feet positions.	10Hz

Table 7.5: Teensy publishers

from the command data is used to set the rotational rate of the servos, as discussed in section 5.3.4. The effector targets are used to interpolate a curve for the feet to move along, as described in section 5.2. The mode is required as some modes could integrate directly with the servo control, namely the torque cutoff mode.

Teensy subscribers		
Name	Data Type	Description
command_data	HexapodCommands	The color image from the RGB-D camera.
effector_targets	EffectorTargets	Data indicating which feet to move where, and what type of interpolation to use.
mode	Int32	Receives mode data from the base station

Table 7.6: Teensy subscribers

### 7.2.3 ROS Data Types

Various custom ROS data types are defined to assist with communication, these data types are described in table 7.7.

Name	Type Definition	Description
Vector3	float[3] data.	A vector in 3D space.
EffectorTargets	Vector3[6] targets bool[6] swinging.	Data describing the targets of the robot's feet and which feet are swinging.
HexapodCommands	float32[2] walk_dir float32 speed float32 height	Data packet containing various command parameters for the robot.

Table 7.7: ROS data type descriptions

## 7.3 GPU Compute Pipeline

As the heightmap generation and heightmap scoring systems are essentially image manipulation processes, parallelisation of the algorithms is a very efficient way to increase computational speeds, thus, these algorithms are run in parallel on the Jetson nano's GPU using OpenGL compute shaders. This section describes the compute pipeline used to build and score the heightmap. The GPU pipeline can be seen in figure 7.2.

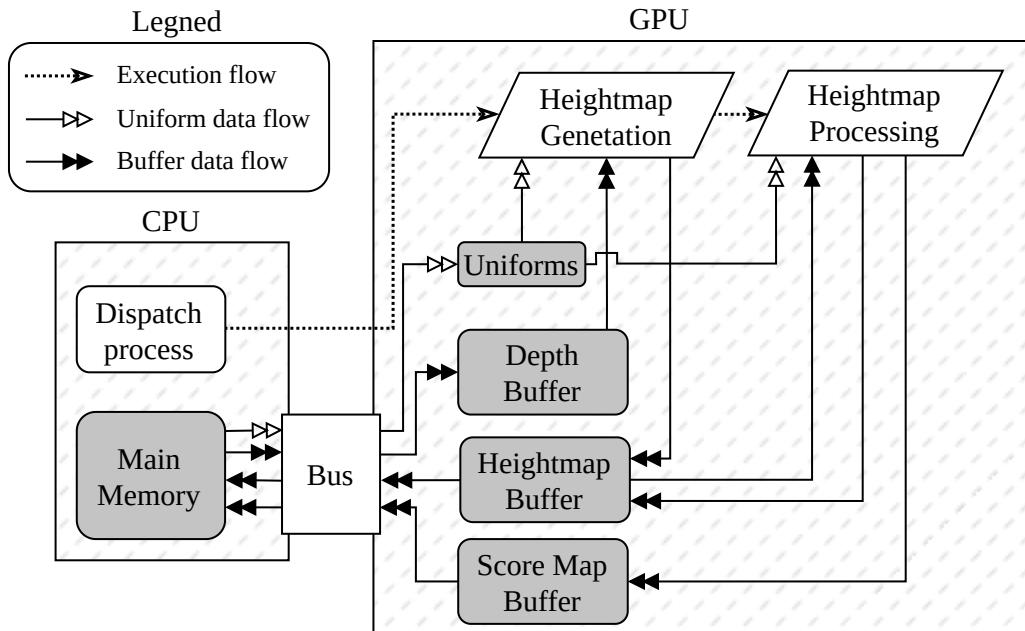


Figure 7.2: Compute pipeline.

As seen from figure 7.2 the GPU pipeline is relatively simple, being comprised of only two stages, the heightmap generation stage and the heightmap processing stage. The heightmap generation stage executes for each pixel on the depth image, constructing the heightmap as described in chapter 4.

After the heightmap has been generated, the heightmap processing stage operates over the heightmap buffer. This stage has two tasks, to erase the old height data as the robot walks around, and to generate the score map, as described in section 6.1.

### 7.3.1 A note on GPU architecture

A process on a GPU operates fully in parallel and the GPU is highly optimised for parallelisation, thus there is a very specific execution structure that a GPU

process abide by to perform at maximum efficiency, or to even function at all. This execution structure is as follows.

When writing compute code a 3D size is specified, the localgroup size,  $\mathbf{N}_l = [X_l, Y_l, Z_l]$ , next when the Central Processing Unit (CPU) dispatches a compute task, the workgroup count, is fed as parameter,  $\mathbf{n}_w = [x_w, y_w, z_w]$ . The GPU then initialises  $\mathbf{n}_w$  workgroups, and each workgroup initialises  $\mathbf{N}_l$  threads. These threads are grouped into warps, or waves, which, depending on the architectures processor count, can be either 32 or 64 threads. To ensure maximum efficiency it is important that  $\mathbf{N}_l$  is divisible by the warp size. Originally Nvidia GPUs utilised a warp size of 32 and AMD 64, however with AMDs latest RDNA architecture, the warp size could be either 32 or 64. This system assumes a warp size of 32.

For the heightmap generation stage  $\mathbf{N}_l = [32, 32, 0]$  for a total of 1024 threads per workgroup, or in other words, 32 warps per workgroup. As for the heightmap processing stage,  $\mathbf{N}_l = [32, 32, 32]$ , meaning 32768 threads, or 1024 warps per workgroup. As such it is important that camera images, the heightmap and the score map are of a size divisible by 32.

While threads can directly communicate with each other within the same workgroup, direct communication across workgroups is impossible. If cross workgroup communication is desired, this must be performed using GPU buffers, which are slower to access. Thus processes are designed to operate fully independently from each other. Of course, data transfer between the CPU and the GPU is orders of magnitudes slower than accessing local buffers, as such, all CPU, GPU communication is kept to a minimum.

Finally it should be noted that while it is possible to vary  $s_w$  with each execution cycle,  $S_l$  is a constant specified at compile time, and as such the number of threads per workgroup cannot be altered during operation.

---

# Chapter 8

## Results

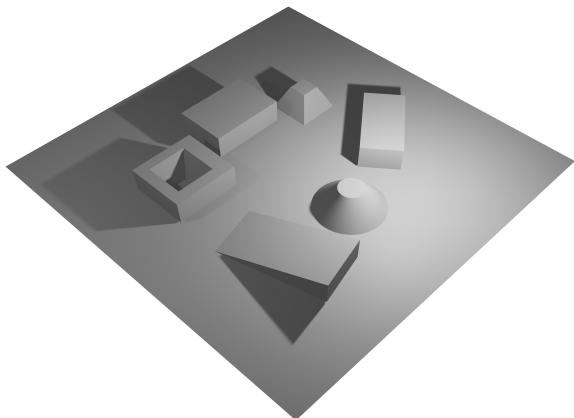
### 8.1 Requirements

### 8.2 Depth Camera

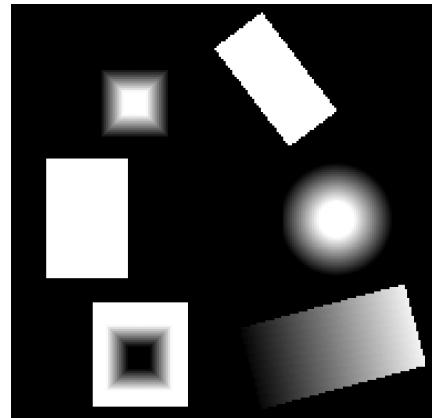
### 8.3 Heightmap

### 8.4 Walkability Scores

For testing the walkability scores described in section 6.1, a piece of terrain that demonstrates various types of obstacles that the robot might encounter was chosen, this terrain and its heightmap is shown in figure 8.1.



(a) 3D Terrain



(b) Heightmap of 3D Terrain

Figure 8.1: 3D Model of Terrain and its heightmap.

The heightmap in figure 8.1b is used to test the various walkability scores,

first, in order to show individual results, the slope and edge proximity score is applied to the heightmap separately from each other, this is shown in figure 8.2. Next the combined score is shown in figure 8.3, this is the score that is used to optimise foot end positions. The slope score can be seen in figure 8.2a,

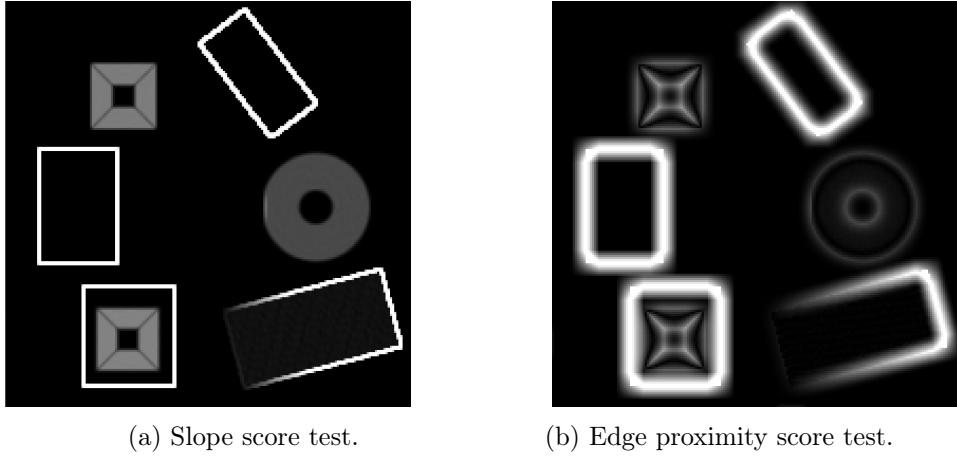


Figure 8.2: Slope score and edge proximity score tests.

from this it can be seen that the slope score successfully indicates sloped areas as less suitable foot end positions. It is also clear that vertical inclines are strongly discouraged, and while not the primary purpose of the slope score, this is expected and does not have a negative impact on scoring.

Figure 8.2b shows that edge proximity score successfully discourages foot placement in areas with large height deviations, while allowing placement in areas with a locally similar slope, such as the ramp in the bottom-right. As can be seen from the larger, and more intense, discourage areas around the boxes in the top-right, bottom-right, left and bottom-left, it is clear that the magnitude of the height differential will increase the size and strength of the rejection area. Next, in figure 8.3 the combined slope and edge proximity scores can be seen. This is simply a weighted addition of the two scores.

Finally, the radial search algorithm described in section 6.2 is used to optimise various nominal foot end positions. This is shown in figure 8.4. If the nominal foot position is in an acceptable position, the position will not be further optimised, this can be seen with the left most nominal position in figure 8.4. If however the nominal position is too close to an edge, such as with the top-right example, the nominal position is optimised to the closest acceptable position. The optimised position does not need to have a perfect score, as can be seen in the bottom-right example, this example is similar to the aforementioned top-right, but it adjusts onto the ram, which has a non-perfect, however acceptable, walkability score. Next the right most example

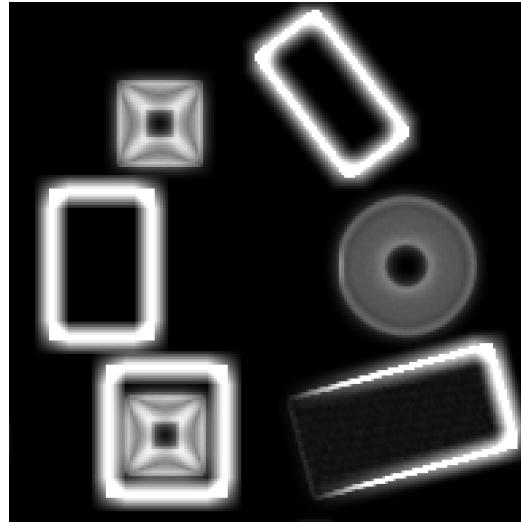


Figure 8.3: Full walkability score.

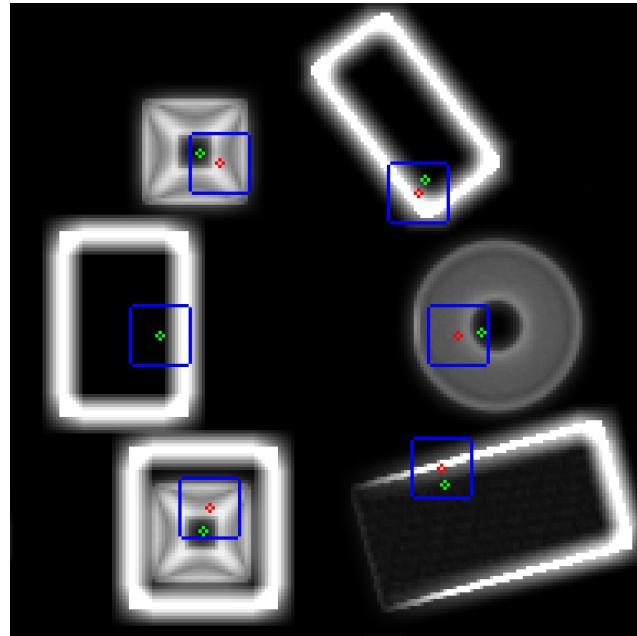


Figure 8.4: Walkability score optimisation test.

shows how the nominal position is place on a piece of terrain that is too steep, this it is optimised onto the flat surface on top of the round pedestal.

Finally, the top-left and bottom-left examples show the case of a nominal position placed on a sloped pillar and a sloped hole respectively. In both cases the nominal position is optimised to the flat surface at the top of the pillar and at the bottom of the hole.

The flat surfaces on the pillar and in the hole are large enough to not be

rejected by the edge proximity score, however, if the inclines were steeper or the flat surface smaller it is possible that these nominal positions would be unsolvable, as there is no appropriate foot end placement inside the search radius. In this case the robot would need to make a course adjustment.

## 8.5 Walking

The final walking tests are comprised of three different terrains, a simple flat plane as a baseline, then a staircase to demonstrate simple foot and body height adjustment and finally a uneven organic like surface.

When presenting testing data, a snapshot of the 3D simulation environment will be shown, with the body and foot paths traced. Additionally key data points are also graphed, however only the height of a single foot will be graphed, this is to maintain clarity.

### 8.5.1 Flat Plane

The flat plane test aims to quantify nominal body height oscillations and to provide a baseline to compare subsequent test to. Figure 8.5 and 8.6 shows the test results of walking over a simple flat plane, as expected the step sizes and arcs are uniform with no optimisation necessary.

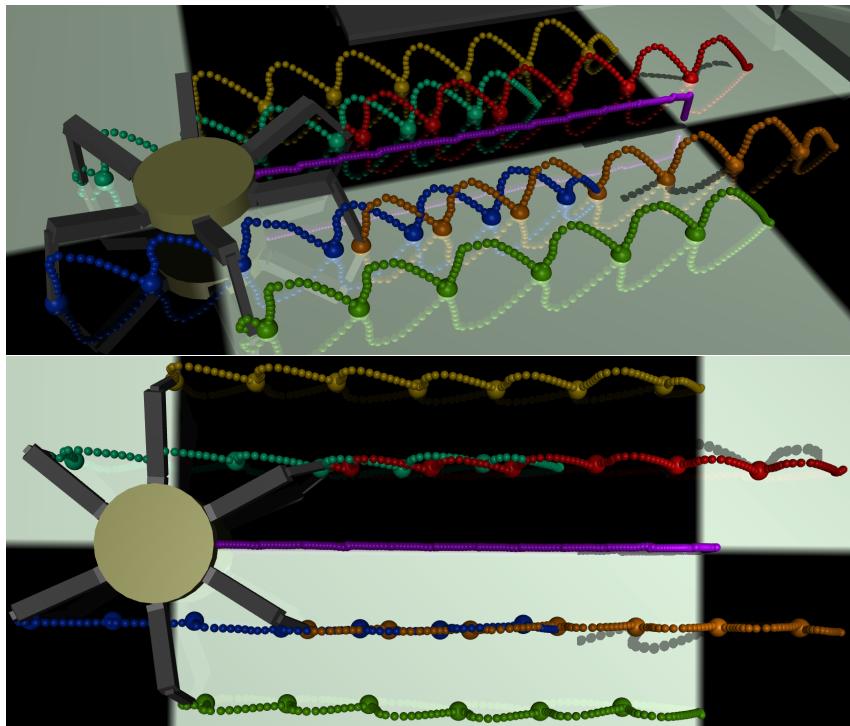


Figure 8.5: Flat plane walk test

From figure 8.6 it can be seen that there is a constant error of about 1cm in the body height of the robot, this can be easily solved by adding a constant offset or adding integral control to the leg servos. It is also clear that as the robot

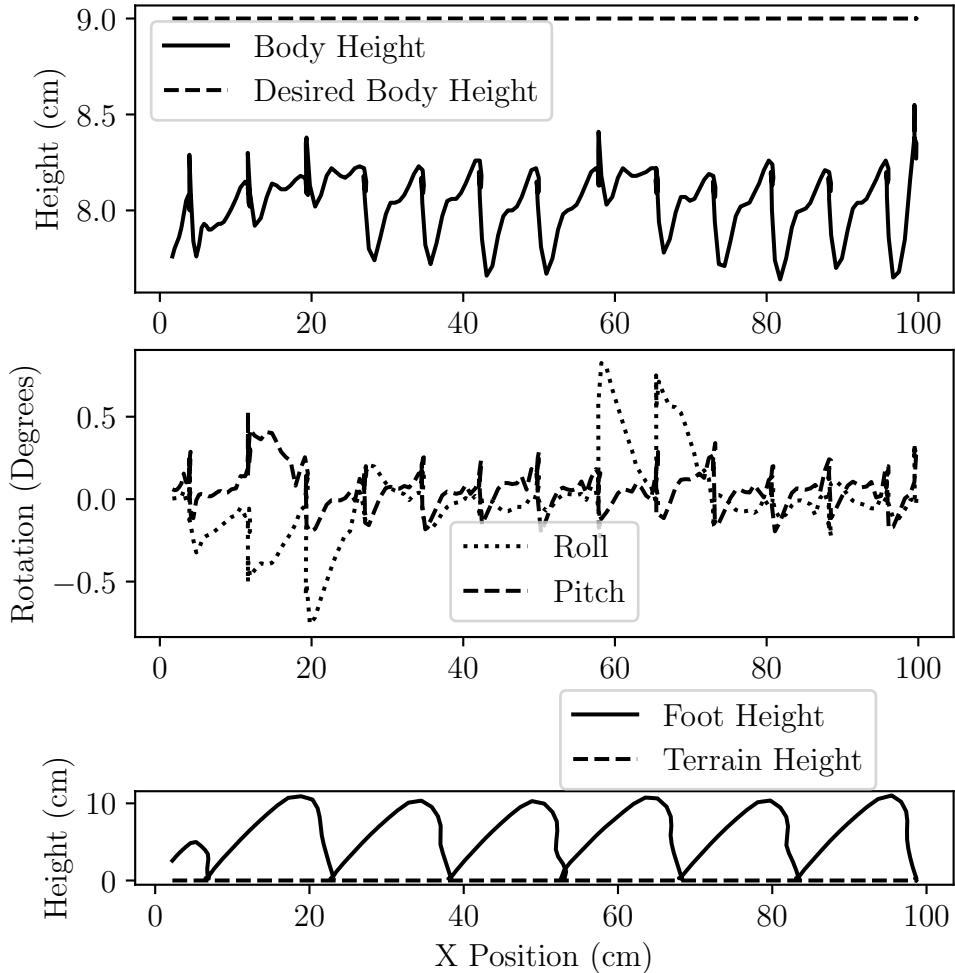


Figure 8.6: Body height (top). Body tilt (center). Position of one foot (bottom)

walks there are some oscillations in the body height, this is to be expected and the oscillations are lower than the 10% desired body height specification.

Next it can be seen that the tilt of the body, in both roll and pitch, tends to stay below  $0.5^\circ$ , which is well within specifications. It is also possible to further improve the tilt error by incorporating accelerometer based feedback control to the feet heights. Finally, the height of one of the legs are shown, from this it is clear that the robot takes uniform, stable steps.

### 8.5.2 Staircase

The staircase test demonstrates the ability of the robot to adjust the height of its feet in order to maintain a level body, additionally the ability of the robot to automatically adjust its height relative to the floor is also demonstrated. A 3D image of this test can be seen in figure 8.7 and the data plots in figure 8.8.

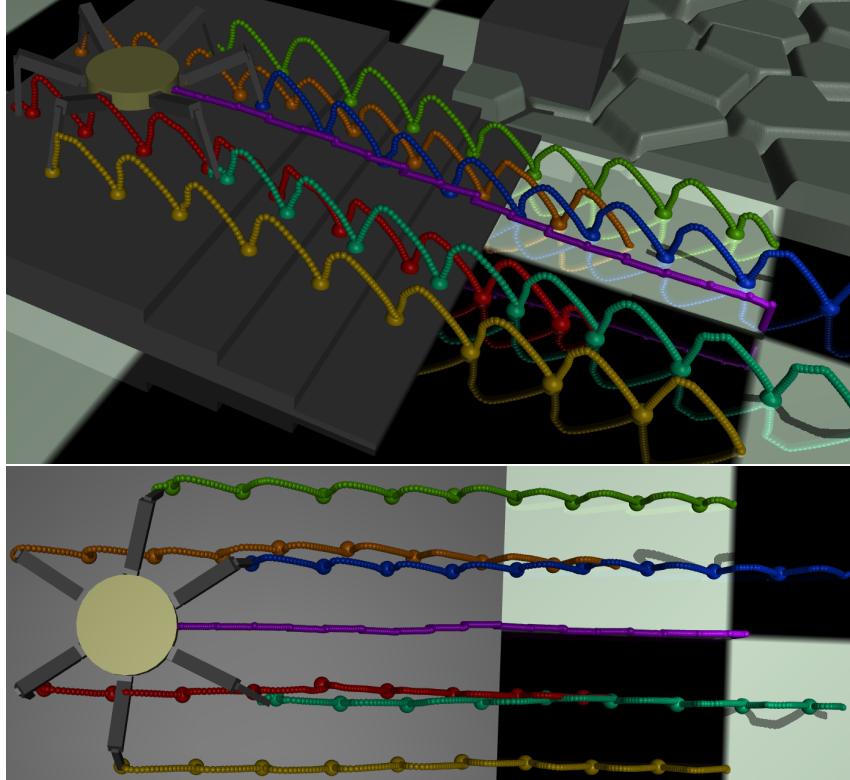


Figure 8.7: Stairs walk test

Similarly to the flat plane test, as expected, the constant body height error is still present. However, from this test it can be seen in figure 8.8 that the robot does increase its height to maintain a constant height above the terrain. It is clear that the body height is not increased in the four discrete steps as the stairs do, rather the body height is increased more smoothly using many smaller steps. This is thanks to the system described in section 6.3.

Body tilt, while more prominent than in the flat plane test, is still quite low as it largely stays below  $1.0^\circ$ , with intermittent jumps approaching  $2.0^\circ$ . This is still within the specifications, and as previously stated could easily be improved by incorporating accelerometer based feedback control on the foot height.

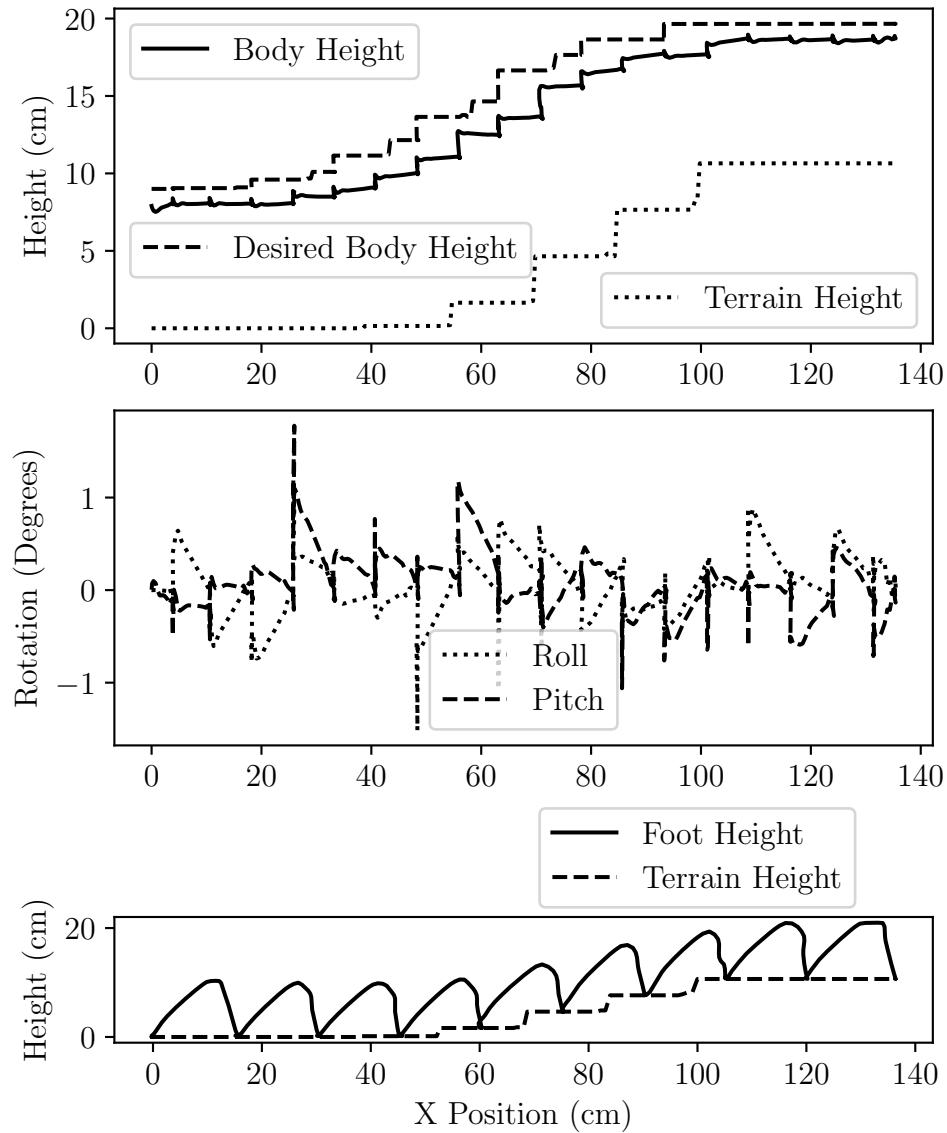


Figure 8.8: Body height (top). Body tilt (center). Position of one foot (bottom)

Finally, the foot height plot clearly shows how one of the feet of the robot adjusts its height and step arc based on the terrain. Otherwise the foot arc looks very similar to that of the flat plane test, which is optimal.

### 8.5.3 Organic

The organic test aims to show how the robot can place feet on an appropriate spot on the terrain, while maintaining level and a certain height. Figure 8.9 shows the 3D view of the test.

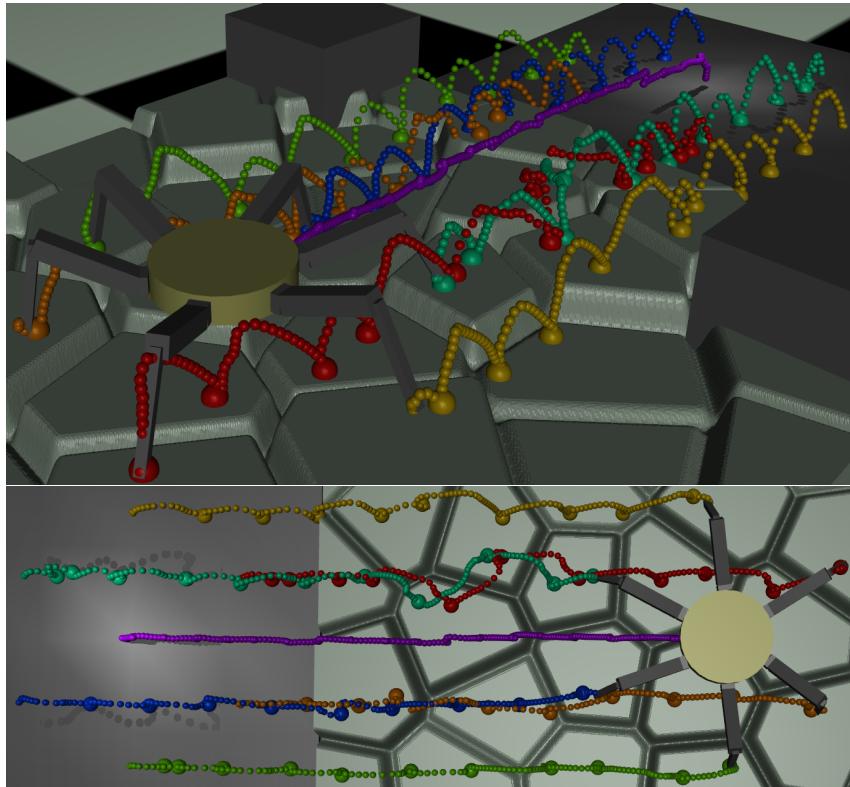


Figure 8.9: Organic walk test.

From figure 8.10 it can be seen that the body height error seen in the flat plane test is still present and more severe. But the body does stay within a reasonable margin from the desired height.

For most part the body tilt is similar to that of the stairs test, however more frequent and more sever spikes are present. This is due, in part, to the more erratic mismatch between the body height and the requested body height. The horizontal adjustments made to the foot end positions cause the swinging feet to end their steps at different times, this is the primary cause of the tilt spikes, as when one foot is placed on the floor before the other two, the robot is tilted slightly. This, however, would not occur if there was no, or little, error in the body height.

Finally, when looking at the foot position plot in figure 8.10, please note that the plot is a projection into the X-Z plane, and thus any adjustment made in the Y-axis is not shown, the foot shown has been chosen to include minimum

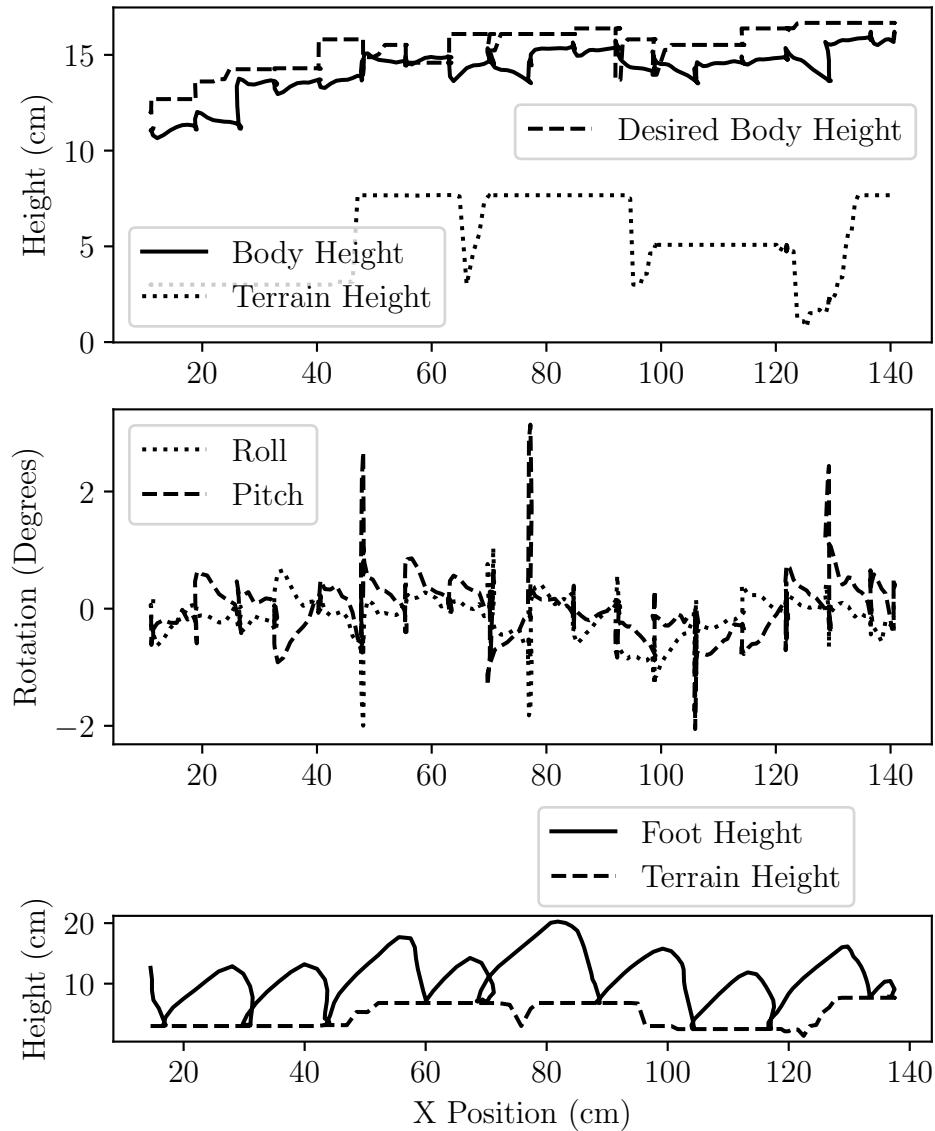


Figure 8.10: Body height (top). Body tilt (center). Position of one foot (bottom)

Y-axis variation. Thus, please also see the 3D view in figure 8.9. From these two figures it is clear that the foot end positions are, in addition to height, adjusted in the horizontal plane, the step arcs are appropriately adjusted and remain relatively smooth, similar to the flat plane test.

# Chapter 9

## Conclusions

# Appendix A

## Reference Frame Transforms

Reference frame transforms are defined by the shorthand  $T_{AB}(\mathbf{x}^A)$ , meaning vector  $\mathbf{x}^A$  is transformed from reference frame  $\mathcal{A}$  to  $\mathcal{B}$ . Reference frames used are the camera frame,  $\mathcal{C}$ , the map frame,  $\mathcal{M}$ , the body frame,  $\mathcal{B}$ , the leg reference frames,  $\mathcal{L}_i$ , and the world frame,  $\mathcal{W}$ . Transforms used are described in this appendix.

### A.1 Camera to Map

Transforming a vector from camera, to map space requires rotating the vector by the camera quaternion, adding the camera's global position to the vector, scaling the vector into the map by multiplying with the map scaling factor, and finally modulating the vector to the confines of the map. As shown in equation A.1.

$$\begin{aligned}\mathbf{x}^{\mathcal{M}} &= T_{\mathcal{CM}}(\mathbf{x}^{\mathcal{C}}) \\ &= (\mathbf{q}_{\text{cam}} \cdot \mathbf{x}^{\mathcal{C}} \cdot \mathbf{q}_{\text{cam}}^{-1} S + \mathbf{p}_{\text{rob}}^{\mathcal{M}}) \mod N\end{aligned}\tag{A.1}$$

where  $\mathbf{x}^{\mathcal{C}}$  is the vector in camera space,  $\mathbf{q}_{\text{cam}}$  is the camera quaternion that rotates  $\mathbf{x}^{\mathcal{C}}$  into world space, with its  $z$  axis pointing upwards,  $\mathbf{p}_{\text{rob}}^{\mathcal{M}}$  is the coordinate vector of the robot in map space, and  $S$  is the scaling factor used to relate heightmap cells to cm. The relationship between  $N, S$  and  $E$  are characterised by equation A.2.

$$N = E \cdot S\tag{A.2}$$

where  $E \times E$  is the size of the  $N \times N$  heightmap buffer in cm.

## A.2 World to Map

The world to map transform is similar to camera to map transform, except that the rotation is not required, as the map and world frames are already aligned. This is shown in equation A.3.

$$\begin{aligned}\mathbf{x}^{\mathcal{M}} &= T_{\mathcal{W}\mathcal{M}}(\mathbf{x}^{\mathcal{W}}) \\ &= \mathbf{x}^{\mathcal{W}} S \mod N\end{aligned}\tag{A.3}$$

where  $S$  is the scaling factor used to relate heightmap cells to cm, and the heightmap buffer is of size  $N \times N$ .

## A.3 Local to Map

Local to map space is very similar to equation A.3, with the exception of using the body quaternion instead of the camera quaternion. This can be seen in equation A.4.

$$\mathbf{x}^{\mathcal{M}} = (\mathbf{q}_{\text{bod}} \cdot \mathbf{x}^{\mathcal{B}} \cdot \mathbf{q}_{\text{bod}}^{-1} S + \mathbf{p}_{\text{rob}}^{\mathcal{M}}) \mod N\tag{A.4}$$

where  $\mathbf{x}^{\mathcal{B}}$  is the vector in local space,  $\mathbf{q}_{\text{bod}}$  is the robot body's quaternion that rotates  $\mathbf{x}^{\mathcal{B}}$  into map space, with its  $z$  axis pointing upwards,  $\mathbf{p}_{\text{rob}}^{\mathcal{M}}$  is the coordinate vector of the robot in map space, and  $S$  is the scaling factor used to relate heightmap cells to cm.

## A.4 Map to Local

Due to the circular nature of the map buffer, translating from map to local space is more challenging. First an intermediate position,  $\mathbf{x}_{\text{temp}}$ , is defined as a helper,

$$\mathbf{x}_{\text{temp}} = \mathbf{p}_{\text{rob}}^{\mathcal{M}} \frac{1}{S}\tag{A.5}$$

where  $\mathbf{x}_{\text{temp}}$  is the map coordinates with the inverse map scaling applied. Now  $\mathbf{x}_{\text{temp}}$  must be checked to see if either of its coordinates exceed half the map extents, if so, then it must be adjusted such that both coordinates fall within  $\frac{1}{2}E$ . This is expressed in equation A.6

$$\mathbf{x}_{\text{temp}} \Leftarrow \begin{cases} \mathbf{x}_{\text{temp}} & \mathbf{x}_{\text{temp}} \leq \frac{1}{2}E \\ \mathbf{x}_{\text{temp}} - E(\text{sgn } \mathbf{x}_{\text{temp}}) & \mathbf{x}_{\text{temp}} > \frac{1}{2}E \end{cases}\tag{A.6}$$

where sgn is the signum function, which return either 1 or -1 depending on the sign of the operand. Note that equation A.6 is applied to both coordinates in the vector  $\mathbf{x}_{\text{temp}}$  separately, in pursuit of readability this is not explicitly stated.

Now that  $\mathbf{x}_{\text{temp}}$  has been corrected for the circular nature of the map buffer, only the rotating into the local space remains. As shown in equation A.7.

$$\begin{aligned}\mathbf{x}^{\mathcal{B}} &= T_{\mathcal{MB}}(\mathbf{x}^{\mathcal{M}}) \\ &= \mathbf{q}_{\text{bod}}^{-1} \cdot \mathbf{x}_{\text{temp}} \cdot \mathbf{q}_{\text{bod}}\end{aligned}\tag{A.7}$$

where  $\mathbf{q}_{\text{bod}}$  is the robot body's quaternion.

## A.5 Body to Leg

# List of references

- Campos, C., Elvira, R., Rodríguez, J.J.G., Montiel, J.M. and Tardós, J.D. (2021). Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam. *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1874–1890.
- Coelho, J., Ribeiro, F., Dias, B., Lopes, G. and Flores, P. (2021). Trends in the control of hexapod robots: a survey. *Robotics*, vol. 10, no. 3, p. 100.
- Collins, J., Chand, S., Vanderkop, A. and Howard, D. (2021). A review of physics simulators for robotic applications. *IEEE Access*, vol. 9, pp. 51416–51431.
- Erasmus, S. *et al.* (2023). Guidance, control, and motion planning for a hexapod robot moving over uneven terrain.
- Erez, T., Tassa, Y. and Todorov, E. (2015). Simulation tools for model-based robotics: Comparison of bullet, havok, mujoco, ode and physx. In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4397–4404.
- Hartley, R. and Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge university press.
- Homberger, T., Bjelonic, M., Kottege, N. and Borges, P.V. (2017). Terrain-dependant control of hexapod robots using vision. In: *2016 International Symposium on Experimental Robotics*, pp. 92–102. Springer.
- Horn, R. and Johnson, C. (2012). *Matrix Analysis*. Cambridge University Press. ISBN 9780521839402.  
Available at: <https://books.google.co.za/books?id=NAffwAEACAAJ>
- Macario Barros, A., Michel, M., Moline, Y., Corre, G. and Carrel, F. (2022). A comprehensive survey of visual slam algorithms. *Robotics*, vol. 11, no. 1, p. 24.
- Mastalli, C., Havoutis, I., Focchi, M., Caldwell, D. and Semini, C. (2020 06). Motion planning for quadrupedal locomotion: Coupled planning, terrain mapping and whole-body control. *IEEE Transactions on Robotics*, vol. 0.
- Prágr, M., Čížek, P., Bayer, J. and Faigl, J. (2019 06). Online incremental learning of the terrain traversal cost in autonomous exploration.

Sobel, I. (2014 02). An isotropic 3x3 image gradient operator. *Presentation at Stanford A.I. Project 1968.*

Xu, P., Li, Z., Yang, H., Wang, Z., Gao, H., Zhou, R., Su, Y., Deng, Z. and Huang, Y. (2023 02). Learning physical characteristics like animals for legged robots. *National Science Review*, vol. 10.