Static Pose Reconstruction with an Instrumented Bouldering Wall

Rami Aladdin

Supervised by: Paul Kry

School of Computer Science
McGill University
Montreal, Quebec, Canada

August, 2012

A thesis submitted to McGill University in partial fulfilment of the requirements of the degree of Master of Science in Computer Science

©2012 Rami Aladdin

Acknowledgments

This thesis would not have seen the day without the contribution, help and support of many. I thank my research supervisor, Paul Kry, for giving me the opportunity to work on this exciting project involving a gratifying mix of hands-on, theoretical, and programming challenges and experimentation. I am grateful to him for placing the responsibility of bringing the instrumented wall project to fruition in my hands. I am also thankful for his availability and for always offering invaluable feedback, advice and ideas.

Many thanks go to our system administrator, Ron Simpson, for making the project a reality. Without his contribution, there would be no climbing wall or calibration tool, and his exceptional carpentry skills brought the idea behind the thesis to life.

I thank my lab-mate, Simon Courtemanche, for contributing, with his expert climbing skills, to designing different climbing patterns, and to collecting data by performing trials as well as attracting numerous climbers to the lab. Working together to record the trials needed for our respective research purposes made the succession of days after days of data collection bearable. I am also thankful for our discussions that brought great perspectives to the table.

Thanks and appreciation go to lab-mates for bearing with the invasive capture trials, as well as to the capture subjects for their courage in wearing our smelly motion capture suit and their patience in enduring the several hours of work needed for a capture trial.

Last but not least, I am grateful to my room and soul mate, Raphaël Mannadiar, for his support, patience and feedback.

Abstract

As the level of realism required by modern video games and movies increases, so does the need for natural-looking simulations. Reproducing the physical behaviour of humans has been one of the pillars of modern research in the field of computer animation. This is a challenging problem because it is difficult to describe what constitutes a natural pose or motion. This problem can be addressed through optimizations that build on assumptions and approximations about properties such as energy, head orientation, or centre of mass position. Alternatively, instead of attempting to create realistic models for humans, motion capture can be used to record (and replay) the pose and motion of human subjects. Most existing approaches rely on some form of combination of physics-based optimization and motion capture. Additionally, some approaches augment and/or attempt to replace motion capture data with contact force measurements. The contributions of this thesis impact many of the above. First, we introduce a setup and a calibration technique for synchronously recording the pose as well as the contact forces in the special case of climbing. Our setup consists of an instrumented bouldering wall that records contact forces and an array of motion capture cameras that record posture. Our calibration technique uses an iterative and alternating least-squares optimization to reconcile force measurements and captured poses in a single space and time frame. Second, we study the relationship between captured motion and forces in an effort to fully describe poses from contact forces. Eliminating the need for motion capture is especially desirable in the context of climbing because of occlusions. To estimate static poses from forces, we use a physics-based optimization. The optimal solution for our objective function is the pose with the highest physical plausibility given the forces

and additional constraints such as climber anatomy and hold positions. To eliminate local minima and speed up our optimization, a simpler "hint objective" is used to guide our solver toward a promising region. Comparison between poses reconstructed from forces and the corresponding poses obtained via motion capture shows that our objective function is a good model for human posture.

Abrégé

Le réalisme toujours plus accru des films et des jeux vidéo modernes nécessite des simulations paraissant de plus en plus naturelles. La reproduction des comportements physiques d'êtres humains est l'un des piliers de la recherche moderne dans le domaine de l'animation. C'est un problème complexe parce qu'il est difficile de décrire en quoi consiste une pose ou un mouvement naturel. Pour adresser ce problème, il est possible d'utiliser des optimisations portant sur des hypothèses et approximations de propriétés telles que l'énergie du système, l'orientation de la tête, ou la position du centre de masse. Plutôt que de tenter de modéliser des êtres humains de manière réaliste, la capture de mouvement peut être utilisée pour enregistrer (et rejouer) la pose ou le mouvement de sujets capturés. La plupart des approches existantes proposent une certaine combinaison d'optimisation basée sur la physique et de capture de mouvement. Certaines approches complémentent et/ou tentent de remplacer la capture de mouvement via des mesures de force de contact. Les contributions de cette thèse sont pertinentes pour bon nombre de ces approches. En premier lieu, nous introduisons un système et une technique de calibration pour mesurer de manière synchrone des poses et des forces de contact dans le cas particulier de l'escalade. Notre système est constitué d'un mur d'escalade équippé de capteurs qui mesurent les forces de contact, ainsi que d'un ensemble de caméras qui enregistrent la pose du grimpeur via capture de mouvement. Pour notre technique de calibration, nous utilisons une optimisation des moindres carrés qui agit de manière itérative et alternante pour exprimer les forces et les poses capturées dans un même repère spatio-temporel. En second lieu, nous étudions la relation entre le mouvement et les forces capturés dans le but d'exprimer des poses exclusivement en terme des forces

de contact mesurées par les capteurs. Éiminer ainsi la capture de mouvement est particulièrement avantageux dans le contexte de l'escalade en raison des occlusions. Nous utilisons une optimisation basée sur la physique afin d'estimer des poses statiques à partir de forces de contact. La solution optimale pour notre fonction objectif est la pose la plus plausible pour les forces de contact et autres contraintes physiques prises en compte, telles que l'anatomie du grimpeur et l'emplacement des prises. Pour éliminer les minima locaux et pour accélérer l'optimisation, nous utilisons une fonction objectif simple qui guide notre résolveur vers une région prometteuse. En comparant des poses reconstruites à partir des forces de contact avec des poses mesurées via capture de mouvement, nous montrons que notre fonction objectif permet de modéliser de manière adéquate le choix de posture d'êtres humains.

Contents

1 Introduction **1**

2 Related Work **4**

2.1 Climbing and Grasping 5

2.2 Combining Forces and Motion Capture 7

2.3 Pose Space Constraints and Optimization 8

3 Design, Calibration, and Capture **10**

3.1 Wall Design 10

3.2 Calibration 14

3.2.1 Converting the Raw Data to Forces 15

3.2.2 Optimization-Based Space Calibration 17

3.2.3 Pocket-Based Space Calibration 21

3.2.4 Measurement Synchronization 21

3.3 Capture 22

4 Reconstruction of Static Poses **26**

4.1 Statistical Analysis 27

4.2 Physics-Based Approach 31

4.2.1 Physical Simulation and CMA-ES 32

4.2.2 Physical Plausibility as an Objective 35

4.2.3 Hint objective for improved optimization 37

5 Results and Discussion **41**

5.1 Calibration 41

5.1.1 Contact Points 42

5.1.2 Sensor Frames 45

5.2 Reconstruction of Static Poses 47

5.2.1 Regression 47

5.2.2 CMA-ES 47

5.2.3 Reconstruction 49

6 Conclusions **53**

6.1 Summary and Contributions 53

6.2 Limitations and Future Work 54

A Space Calibration Optimization Detail **57**

A.1 Contact Point 57

A.2 Sensor Frame 58

A.2.1 Origin of the Frame 59

A.2.2 Orientation of the Frame 59

Bibliography **62**

List of Figures

3.1 The instrumented bouldering wall. 11
3.2 Close-up on a "sensor sandwich". 12
3.3 The bouldering wall in construction. 13
3.4 Picture of an amplifier board [AMT06]. Jumper pins are highlighted. 14
3.5 Voltages at each channel recorded for an example session. 15
3.6 The calibration tool in action. 19
3.7 Front (a) and back (b) view of marker placement on the climber. 23
3.8 Climber during a trial (a) and obtained visualization (b). 25

4.1 Subject generating arbitrary forces while maintaining a pose. 27
4.2 Variance explained against number of components for forces (a) and poses (b). 28
4.3 Correlation between forces and centre of mass position for static poses. . . . 29
4.4 Root position regression error using first-order polynomials. 30
4.5 Joint angle regression error with first- (a) and second-order polynomials (b). 30
4.6 Joint angle prediction error using canonical covariance analysis. 31
4.7 Skeleton model used for simulation. 33
4.8 Illustration of the behaviour of CMA-ES (used with permission) [cma]. . . . 34
4.9 Illustration of the shape of the fitness function. 39
4.10 Visualization of the optimization process. 40

5.1 First (a) and second (b) bouldering puzzles. 42
5.2 Point cloud illustrating contact point estimation error. 44

5.3 Reconstruction (left) and ground truth (right) for varied examples. 52

List of Tables

3.1 Sensitivity settings for sensors. 14
3.2 Excitation voltages (in V) and gains for an example trial. 24

5.1 Averaged errors and standard deviations for three contact point estimates. . 44
5.2 Errors for sensor frame estimates. 46
5.3 Root position regression error. 48
5.4 Durations and reconstruction errors for optimization examples. 51

Chapter 1

Introduction

Producing physically plausible computer animation of virtual humans is difficult because of the complexity and subtleties of how real humans control posture and motion. While it is easy to write down the equations of motion for an articulated character, it is difficult to model how this character should move or what constitutes a natural pose. One way to deal with this is to devise an optimization problem and design a set of terms for the objective function based on reasonable assumptions and approximations, such as terms that minimize metabolic energy, that keep the head level, or that guide the centre of mass to a location above the feet. Another way to generate plausible postures for virtual humans is to simply use motion capture and record the pose of a real person. Captured motion can be challenging to modify, however, and finding computational models to edit the motion for new purposes is difficult.

In this thesis, we build on the idea of bringing together physics, optimization, and motion capture. This is a popular approach in research on virtual humans. However, we also recognize that motion capture alone is only part of the picture. Contact forces are essential for building up a clear understanding of how posture and motion are produced, as they let us resolve the ambiguity in determining the torques applied at different joints. Contacts and contact forces are critical, whether in the context of standing balance, object manipulation,

or in this research, during climbing. We see climbing motion and posture as an interesting example to focus on because it combines locomotion and object manipulation, where the object being manipulated is the body as a whole.

With our desire to study posture and movement of humans in interaction with their surroundings, and the goal of developing improved virtual humans, we designed an instrumented climbing wall. Our climbing wall permits forces and torques to be measured at the holds while motion capture records the pose of the climber. It also allows a variety of different experiments to be conducted: the pitch of the wall is adjustable, and hold positions are easily reconfigurable in predefined pockets. Also, the wall is easily disassembled for transport or long term storage. To enable the expression of forces and poses in the same coordinate frame, we developed two calibration processes. The calibration process for force sensors mounted in arbitrary locations is presented, in addition to a convenient calibration based on known mounting locations.

In this investigation, we focus on static and slow-moving, near-static postures in order to simplify the problem[1]. While we can capture both motion and contact forces simultaneously, another goal of this work is to have a method for reconstructing the posture of a climber from the force capture alone. Reconstruction of poses from forces is useful because motion capture of climbing can be challenging due to occlusions, and because it is convenient to be able to produce posture estimates for a climb without requiring the capture subject to wear a motion capture suit. We use optimization to find a pose that meets the contact constraints while being valid for the measured forces and satisfying additional plausibility requirements, such as wall contact, joint limits, and facing direction. The objective function is quite complex, and uses a physics-based simulation as a black box for its evaluation. As such, we use Covariance Matrix Adaptation Evolution Strategy [Han06] (CMA-ES) to compute the solution because a derivative-free, sample-based optimization method is well suited

[1] We defer dynamics to future work.

in this case. We accelerate the optimization process with a simple objective function that guides the hips of the character to a location predicted from the forces by a linear regression. This helps bring our optimization to the neighbourhood of the optimal solution quickly, and reduces the number of computationally expensive evaluations of our full objective function. Our estimates of posture from recorded forces combine physics, capture, and optimization to produce plausible poses which we validate with ground truth motion capture measurements.

Our three main contributions are: an instrumented climbing wall design, a calibration process for this wall, and a new optimization-based method for estimating posture from contact forces. The remainder of this thesis is structured as follows. We discuss the related work in Chapter 2. In Chapter 3, we describe our instrumented bouldering wall design, we detail the calibration procedure to relate motion capture and force measurements, and we describe the capture process. In Chapter 4, we discuss the challenges in estimating poses from forces and detail our optimization-based method. In Chapter 5, we present the results for the described calibration and optimization. Finally, Chapter 6 contains a summary of our work as well as directions for future applications and extensions.

Chapter 2

Related Work

In this chapter, we discuss related works from the fields of biomechanics, robotics, and animation. First, we explore works that are related to climbing. Then, we examine works that focus on pose reconstruction by combining force and motion or by defining and exploiting a low-dimensional pose manifold. Finally, we discuss works that use optimization methods for physical simulations.

Climbing is a rich combination of locomotion and body manipulation. It is the focus of many works in the field of biomechanics and is a challenge for many researchers in robotics. Climbing-capable robots are useful as they can perform tasks that can be hazardous to human operators, but they need to have a high power-to-weight ratio and they need to be able to generate stable grips with the surface.

While climbing has not been extensively studied in animation, a closely related topic has received more attention: grasping synthesis and analysis. Combining motion and forces, and the study of their underlying relationship is also a popular topic, and is often driven by pose reconstruction from forces. Finding a relationship between forces and pose is interesting because it is a first step in understanding how posture is chosen and it offers an additional perspective in understanding what constitutes a natural, human-like pose. Other authors

reconstruct poses from different types of lower-dimensional data through the incorporation of physical constraints. Finally, the use of optimization for physics-based simulation is a common approach in dealing with space-time constraints.

2.1 Climbing and Grasping

Although not the focus of this research, the interest for climbing in the field of robotics shows that it is a fundamental type of locomotion, alongside swimming and flying, that deserves to be distinguished from the more traditional walking and running. While it could be argued that similar gaits are found in climbing and walking for articulated robots, many gravity-related constraints need to be taken into account to design a climbing-capable robot. Luk et al. [LCB91] described the design of an articulated climbing robot as early as 1991. Many considerations emanating from the behaviour of insects, such as keeping the centre of mass close to the wall or having a large limb to body length ratio, are taken into account in this work. To grip onto smooth surfaces, the authors use vacuum suckers at the end of the limbs as well as under the base of the robot. More recently, Spenko et al. [SHS+08] presented a design for a robot capable of climbing on a wide range of surfaces, including trees and man made materials such as brick, stucco, cinder block and crushed rock. This flexibility is achieved by using both interlocking mechanisms (using spines and claws) and bonding (using smooth adhesive patches), allowing the robot to climb both rough and smooth surfaces. A combination of these two methods is in fact used by many small animals such as lizards, frogs, and insects. In practice, human climbs are fundamentally different and offer additional constraints, as they are driven by the locations of the holds for grip.

Climbing has also been studied by various research groups in the field of biomechanics using instrumented holds. Rougier et al. [RBMB91] described the first instrumented wall designed to measure the amplitude of forces at the holds in 1991. Testa et al. [TMD99] used 3-Degrees of Freedom (3DoF) sensors. Quaine et al. [QMB97a] used two video cameras to record two 2D views of the climber while recording forces on 3DoF transducers to

include postural information. A more recent instrumented bouldering wall design is described by Fuss et al. [FN07], where eight holds are equipped with 6DoF sensors. Works using these setups focus on analyzing the performance of a climb and measuring the difficulty in climbing [RB91, FN07], on studying the impact of the climber's level of expertise [FN06, FN08, RBMB91, RB92], and on studying the effects of specific posture changes or constraints on forces at the holds [QMB97b, QMB97a, QM99, TMD99, TMD03]. Our setup is different in that it incorporates motion capture to offer full 3D reconstruction of forces, torques and posture synchronized in time and space. The objective of our approach is also different as we attempt to reconstruct poses from forces.

In computer animation, a closely related topic that has been widely covered is grasping. In grasping, objects are manipulated through contacts with fingers. Similarly, in climbing, the body is manipulated through contacts and grasps with the hands and feet. As such, many works in grasping focus on evaluating the quality of a grasp or the choice of the pose in different contexts in the same way as many works in biomechanics concentrate on evaluating the performance of a climb or the effects of certain constraints on pose and forces. The relevance of grasping in the context of this thesis is strengthened by the static nature of grasps. Miller et al. [MA99] designed a system that computes measures of quality for a grasp by determining the types of contacts given the pose of a hand and an object to grasp. Ciocarlie et al. [CDL$^+$09] studied the selection of fingertip contact point locations during grasping and manipulation through a number of disturbances. While gripping a single hold can be enough to maintain contact with the surface in the case of climbing, the stability of a grasp is determined by the joint contributions of each contact patch between the hand and object. Therefore the considerations for a quality pose are vastly different and rely more on energy efficiency than robustness to disturbances.

2.2 Combining Forces and Motion Capture

There exists a wide range of contributions to capture and synthesis of posture or force data in computer animation. Brubaker et al. [BSF09], who studied the problem of estimating forces acting on the system from a given motion, used a simulated character to model both the subject and the constraints. We follow the same approach in the context of the opposite problem of estimating poses from forces.

Motion capture and force sensing are used by Kry et al. [KP06b] to capture interactions, focusing on hands and grasping. Joint compliance is estimated from small time windows just before and after contacts, and used along with the reference trajectory to synthesize new interactions. While it would be interesting to look at how this method could be applied to climbing to synthesize new postures with different wall configurations, the focus of our research is the relationship between static poses and forces at the holds.

More intimately related to our work, Yin et al. [YP03] explored how full body poses can be estimated using force measurements from a foot-ground pressure sensor pad. Kry et al. [KP06a] explored how grasp configurations can be estimated using force measurements from a graspable device. These approaches use a database consisting of pairs of motion capture and force data to find the pose for the closest set of measured forces. We, on the other hand, use a physics-based approach to find a pose that best explains the forces.

Ha et al. [HBL11] used reduced features and a physics-based approach to reconstruct human motion from consumer-grade foot-ground pressure sensor pads. While they used a form of motion capture to obtain hand positions, one can argue that in the case of climbing, both hands and feet positions are given by the holds whereas only foot locations can be obtained from standing on a platform. Fundamental differences are that our work focuses on statics and explores how climbing-specific constraints can drive the reconstruction through the optimization of an objective function that we introduce as a means of characterizing static poses.

In the field of computer vision, Rajalingham et al. [RVC10] presented a probabilistic approach to determining the position of the feet for users in movement using an instrumented floor surface. Their results were validated using synchronously recorded motion capture and force data by placing markers on the subject's feet. Using Bayesian filtering, the authors were able to track the feet with reasonable precision using a coarse sensor array. While it would be interesting to see how probabilistic approaches behave for predicting poses from forces in our setup, our work is directed toward a physical approach to relating forces and posture.

2.3 Pose Space Constraints and Optimization

In an effort to retrieve motion without the use of markers, several authors have focused on combining video and physical constraints to reconstruct motion. Rosenhahn et al. [RSB+08] used simulation to enforce constraints found through analysis of capture of special interactions to take advantage of the lower dimensionality of the pose manifold. Bregler et al. [BMP04] started from a hand-initialization applied to a video frame to track a skeleton and use exponential maps and twist motion to perform a differential motion estimation over a video clip. Hasler et al. [HRT+09] extracted body shape information using 3D scanners and used frames from moving cameras, synchronized through sound, to find a least-squares reconstruction of the pose. They used climbing, which is difficult because of the occlusion caused by the wall, as an example to prove the robustness of their method. While these works do not rely on any force capture, they are highly relevant because they show how physical constraints can be used to reduce and solve an otherwise under-constrained problem.

In order to compress motion capture data, Tournier et al. [TWC+09] presented a method for retrieving an entire pose through end effector positions using an approximation of the pose manifold found through principal geodesics analysis. Their work introduces an interesting approach to approximating the pose manifold and is therefore extremely relevant in

the context of pose reconstruction. However, of all body positions, those of end effectors are least relevant in distinguishing between poses in the case of climbing, as they are fixed to hold locations. Important differences with our work are that we focus mainly on finding a correlation between poses and the forces at the holds, and that our goal is to predict an unknown pose for a given set of forces.

Another body of relevant research in animation concerns the use of optimization for physics-based simulations. De Lasa et al. [dLMH10] described a means of expressing locomotion using a small number of features. Jain et al. [JYL09] also looked at constrained optimization for virtual characters to design controllers by formulating high-level objectives. They used this method to build a number of controllers, including a climbing controller. A method we share with Nunes et al. [NCNV$^+$12] is the use of CMA-ES with simple penalties around the main objective. However, all these works look at generating plausible motion through the optimization of arbitrary objectives, while we are interested in optimizing the physical quality of a static pose reconstructed from a set of forces.

Chapter 3

Design, Calibration, and Capture

In order to study the process of climbing, we designed an instrumented climbing wall to measure forces and torques at the holds, and use motion capture to record the pose of the climber. In this chapter, we first describe the wall and discuss the advantages of our design that include portability and flexibility in capturing different climbing scenarios. Then, we investigate the challenges in relating the force measurements captured by the wall with the motion capture pose data, and offer a general solution for a complete calibration of the system. Finally, we present the full capture procedure, elaborate on the main difficulties and concerns for completing a capture trial, and present a visualization of the result of a successful trial.

3.1 Wall Design

The instrumented bouldering wall's main structure is an 8'-by-8' torsion box, as shown in Figure 3.1. The torsion box design was chosen in order to have a rigid yet light core, and will ultimately allow us to hide sensors and wires inside and behind the pockets, leaving sensitive material out of harm's way and giving climbers the same freedom they would have on a regular wall. Specifically, additional slits will be cut into the wall and additional plates will be built to cover the pockets. The wall does not need to be fixed to a second supporting wall

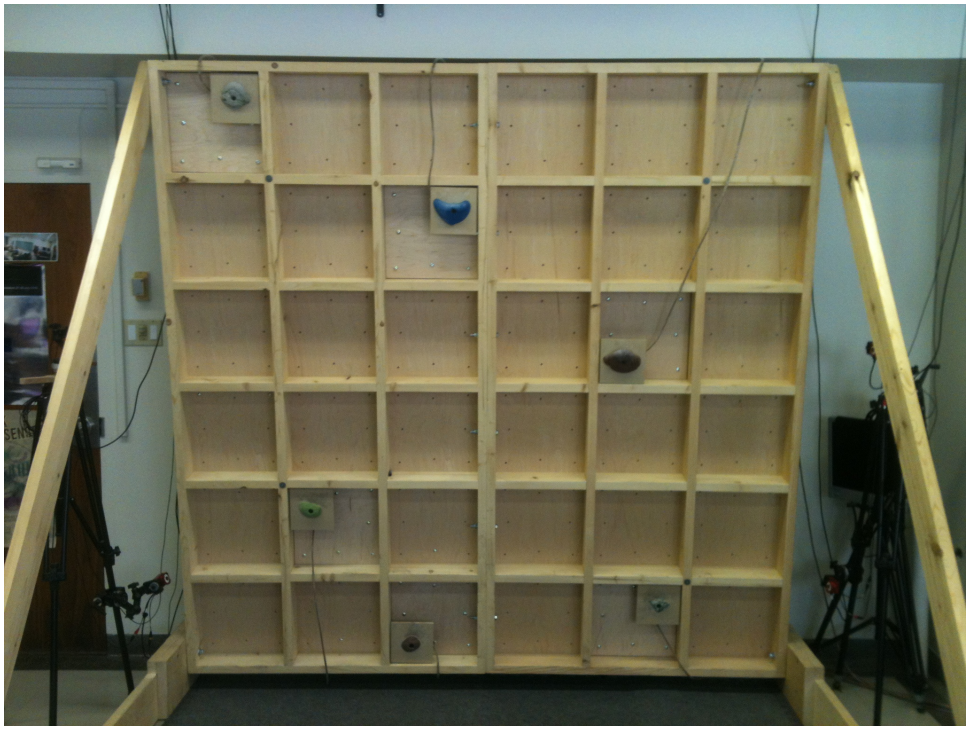


Figure 3.1: The instrumented bouldering wall.

as four additional beams offer support for the box and climber's weight. This configuration allows for easy movement of the wall and offers control over the desired tilt of the climbing path.

Each climbing hold is mounted onto a front plate which is in turn attached to a 6-axis force torque sensor. The sensor is fixed to a back plate that is mounted on the wall. The hold, front plate, sensor, and back plate form a "sensor sandwich", as shown in Figure 3.2. Note that the high stiffness of the sensors ensures that there is a gap of a few millimetres (and therefore no contact) between the front plates and the frame at all times. The sensor sandwiches can be inserted into each pocket in four different orientations, providing flexibility for designing climbing routes. The design also makes reconfiguration easy. Moving a hold or changing its shape can be done by respectively unscrewing the sensor from one of the 144 possible positions, and unscrewing the hold from the sensor. We built the wall from the ground up with the help of an expert carpenter, as shown in Figure 3.3.

Figure 3.2: Close-up on a “sensor sandwich”.

The sensors are multi-component transducers with six channels for linear and angular forces expressed in a local coordinate frame. Each sensor is connected to a 6-channel strain gauge amplifier with adjustable excitation voltage and gain. Excitation voltage and gain can be adjusted with jumper pins, shown in Figure 3.4. Sensors can therefore easily be calibrated for maximum sensitivity depending on the climber’s weight, or, in the case of dynamic motions, the planned motion. We use two similar types of sensors of varying sensitivity. These different types were chosen for their shape and maximum load specifications. The chosen sensors offer reasonable capacities and sensitivities. In practice, a climber weighing up to 200 lb can be completely suspended to a single hold – therefore making contact as far as 15 cm from the centre of the sensor – and perform dynamic movements (such as chin-ups) without exceeding 50% of the force and torque capacities. The sensors are also sensitive enough to capture small tremors of muscle fatigue, as well as contacts as light as 0.01 lb. Furthermore, the sensors’ high stiffness is an important feature as the holds of a bouldering wall need to be fixed. Finally, the size of the sensors is also appropriate (neither too large,

Figure 3.3: The bouldering wall in construction.

nor too small), and spacers of different thickness, depending on the sensor type, are added to the back plates to ensure the holds are situated along the surface of the wall.

A maximum of four amplifiers can be connected to a computer interface. Therefore, two sets of respectively two and four sensors are connected and used to record data separately. Since sensors possess different sensitivities and because they undergo different forces depending on their position on the wall, they need to be calibrated individually. There are three gain and three excitation voltages to choose from at each channel, as shown in Table 3.1. These settings provide a trade-off between the maximum and minimum loads that the sensor can measure. Calibration matrices to convert raw voltage data to wrenches are close to diagonal and real-time visualization is enough to evaluate excitation voltage and gain choices for sensor sensitivity, as shown in Figure 3.5. Sensor overload happens for voltages below -10 V or above 10 V. Ideally, in order to obtain the greatest range and accuracy from the sensor, these bounds should be reached when applying the maximum loads. In practice, we leave

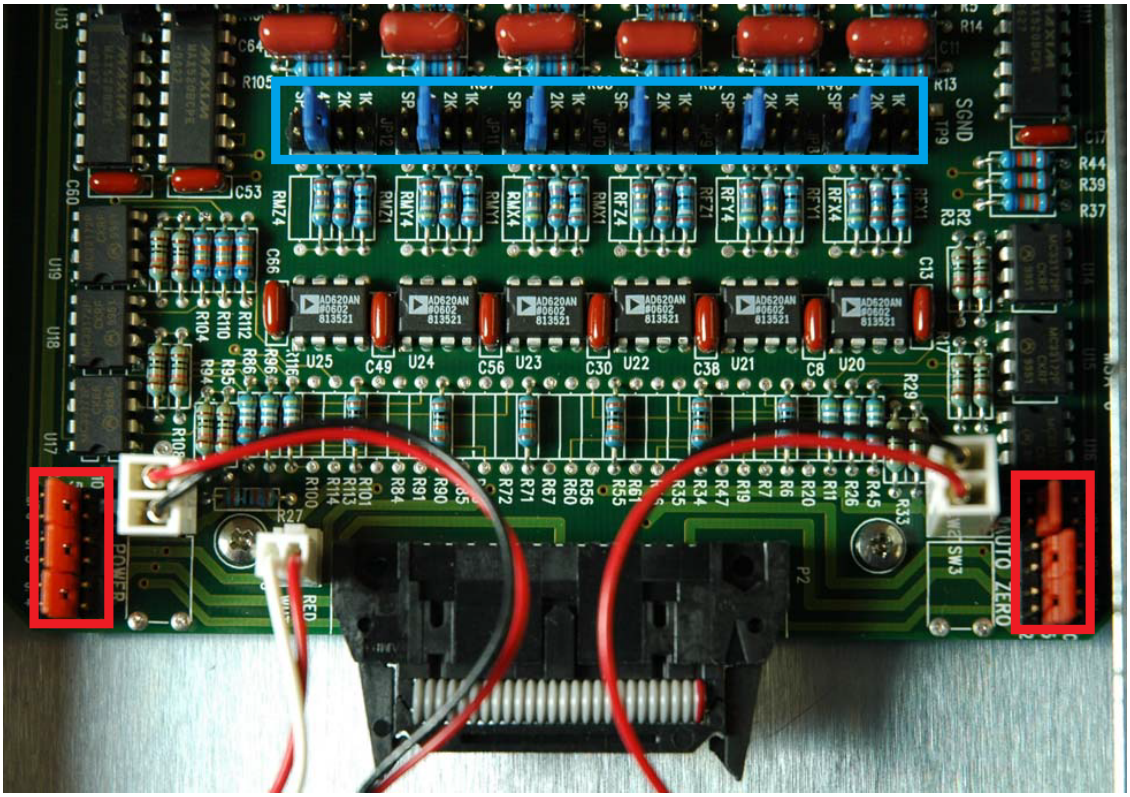


Figure 3.4: Picture of an amplifier board [AMT06]. Jumper pins are highlighted.

Type	Low sensitivity	Medium sensitivity	High sensitivity
Excitation Voltage	2.5 V	5 V	10 V
Gain	1000	2000	4000

Table 3.1: Sensitivity settings for sensors.

a margin to avoid saturating the amplifier outputs. When there is ambiguity in choosing a gain and excitation voltage for a channel, priority is always given to a higher excitation voltage to minimize noise.

3.2 Calibration

The main challenge in capturing data from both the motion capture and force sensing equipment is to find the relationship between the two independently recorded data sets, both in

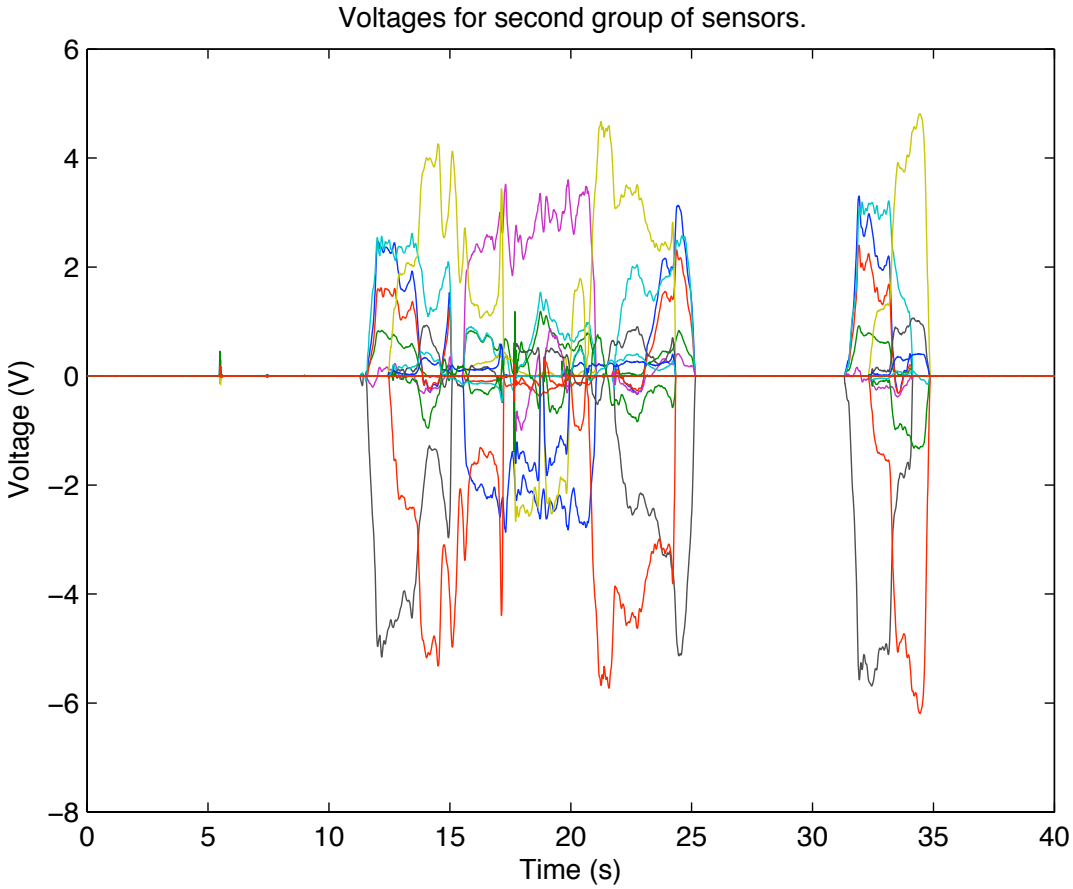


Figure 3.5: Voltages at each channel recorded for an example session.

time and space. Each sensor possesses its own local frame in which force measurements are expressed as voltages, while skeleton reconstruction is expressed in the motion capture world frame. Below, we describe how to obtain forces in the sensors' local frames. Then, we detail two techniques for calibrating the position and orientation of the sensors, which allow the force and torque measurements to be expressed in the motion capture world reference frame. Finally, we introduce a simple technique for synchronizing the two data sets in time.

3.2.1 Converting the Raw Data to Forces

As shown in Figures 3.5, sensors provide voltages for each of the six channels. However, software records data in an unsigned integer format that first needs to be converted to

voltages, and then to actual forces and torques using the sensor's sensitivity matrix. The voltage for a sensor is given by

$$V = B \circ (U - L)/(2^r - 1) + L, \tag{3.1}$$

where V is the 6-by-1 vector of voltages for the different channels, B is the corresponding 6-by-1 vector of unsigned integer values, U and L are respectively the 6-by-1 vectors for upper and lower bounds on the voltages on the different channels, $\circ$ is the Hadamard entry-wise product, and r is the resolution of the computer interface in bits. Using the inverse sensitivity matrices that are provided with each sensor, we then compute the forces and torques from these voltages while taking into consideration possible cross-talk. The forces and torques are given by

$$F = S(V \circ C), \tag{3.2}$$

where F is the 6-by-1 vector containing the forces and torques, V is the 6-by-1 vector containing the voltages for the corresponding channels, S is the manufacturer-provided 6-by-6 inverse cross-talk sensitivity matrix for the sensor, and C is the 6-by-1 vector containing the inverse of the correction factors c_i for each channel i. These correction factors are defined as

$$c_i = gi\ v_i\ 10^{-6}, \tag{3.3}$$

where g_i and v_i are respectively the chosen gain and excitation voltage for that channel, and the last term is used to scale the sensitivity terms which originally convert microvolts to the desired force units. Gains and excitation voltages can be changed at the amplifier boards for each individual channel of each sensor by moving the jumper pin from one sensitivity setting to another, as shown in Figure 3.4. Excitation voltage jumper pins for torques (on the left), and forces (on the right), are highlighted in red. Gain jumper pins for torques and forces (at the top) are highlighted in blue.

3.2.2 Optimization-Based Space Calibration

In order to visualize the force and motion data and to be able to relate the different coordinate frames in which the different measurements are made, we need to find the appropriate transforms. Our goal is to find the rigid transform, represented by the 4-by-4 transformation matrix

$$ {}^{w}_{s}E = \begin{bmatrix} {}^{w}_{s}R & {}^{w}p_s \\ \mathbf{0} & 1 \end{bmatrix} \tag{3.4} $$

that converts homogeneous coordinates expressed in the sensor's local frame to homogeneous coordinates in the world motion capture frame. The leading superscript is used to denote the frame in which a quantity is expressed, and for rigid transforms, the leading subscript denotes the frame in which the quantity the transform accepts on its right hand side is expressed. Here s denotes the sensor frame, w denotes the world frame, ${}^{w}p_s$ is the 3-by-1 vector representing the origin of the sensor frame in world coordinates, ${}^{w}_{s}R$ is the 3-by-3 rotation matrix that aligns the axis in the sensor frame with the corresponding axis in the world frame, and the zero in the bottom row is a placeholder for a 1-by-3 vector of zeros. We use a collection of force measurements at the sensor for forces applied at known locations (on the front plate or hold rigidly attached to the sensor) to estimate the rigid transform that gives us the position and orientation in the world frame. We first explain how we obtain the contact locations where force is applied, and then describe how these contacts are combined with the force measurements to obtain the position and orientation of the sensor.

To obtain contact locations, a calibration tool consisting of a rigid body with a sharp tip is tracked with motion capture as it is rotated around a fixed contact point, as shown in Figure 3.6. For a recorded set of orientations of our rigid body, the contact point p_c is found using a least-squares approach (see Appendix A for details), by computing

$$ \underset{{}^{b}p_c, {}^{w}p_c}{\operatorname{argmin}} \sum_{j=1}^{n} \left\| {}^{w}_{b}E_j \; {}^{b}p_c - {}^{w}p_c \right\|^2, \tag{3.5} $$

where b is the body frame, n is the number of collected samples, ${}^{w}p_c$ is the position of the contact point in homogeneous world coordinates, ${}^{b}p_c$ is the position of the contact point in

homogeneous body coordinates, and ${}^{w}_{b}E_j$ is given by the motion capture software for each recorded frame j. In other words, we are looking for the point p_c that is invariant for ${}^{w}_{b}E$ over the data. That is the point that best describes a centre of rotation for all individual frames, which is the location of the contact between the tip and surface. Knowing the position of the tip in body coordinates, we can thereafter find any contact location in world coordinates for any new recorded sequence of motion capture.

In order to find the position and orientation of a sensor, the calibration tool is used to apply forces on the front plate and hold rigidly attached to the sensor. Since the calibration tool contacts the wall surface at only one point, it can only generate linear forces. Assuming the torque generated at the contact point is zero, we use data from multiple contact points to estimate the location of the sensor frame. For simplicity, we define contact coordinate frames aligned with the world coordinate frame. The measured torque can be expressed as

$$ {}^{c}\tau = {}^{w}_{s}R\ {}^{s}\tau + ({}^{w}p_s - {}^{w}p_c) \times {}^{w}_{s}R\ {}^{s}f\ . \tag{3.6} $$

Since we are not applying a torque at the contact point, ${}^{c}\tau = 0$. Thus, finding the origin and orientation of the sensor's local frame given a wide variety of contacts and associated applied forces becomes an optimization problem where we are trying to find the sensor frame that can best explain force and torque measurements for every given contact point. Concretely, we solve for

$$ \underset{{}^{w}_{s}R, {}^{w}p_s}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{j=1}^{m} \| {}^{w}_{s}R\ {}^{s}\tau_{ij} + ({}^{w}p_s - {}^{w}p_{ci}) \times {}^{w}_{s}R\ {}^{s}f_{ij} \|^2 , \tag{3.7} $$

where k is the number of contacts and m is the number of recorded force samples per contact. We find ${}^{w}p_c$ using the calibration tool and motion capture, and ${}^{s}\tau$ and ${}^{s}f$ are given by the sensors. All that is needed in order to perform this calibration is therefore a large and varied set of forces applied at various contact points to ensure that the problem is not under-constrained. In practice, three different contact points and 2000 samples per contact point are enough to find good solutions.

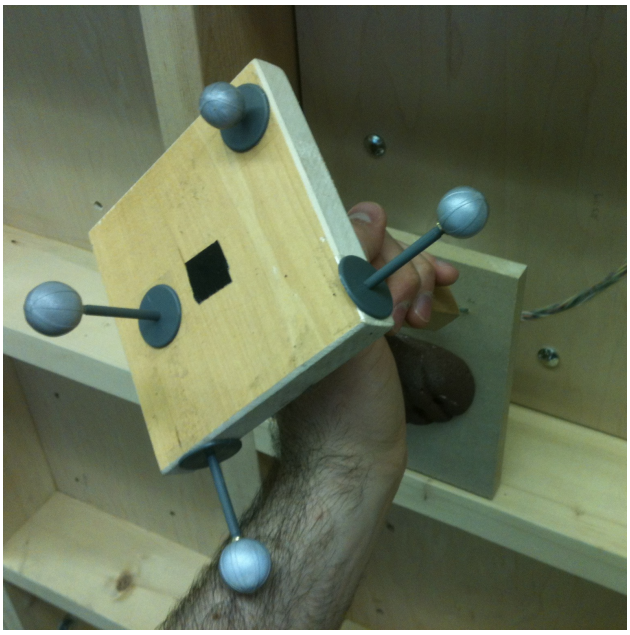


Figure 3.6: The calibration tool in action.

We use motion capture to track the calibration tool as it is used to apply linear forces at the contacts while the sensors record forces and torques. With the estimated contact point locations and recorded forces at the sensor, Equation 3.7 is used to find estimates for ${}^{w}_{s}R$ and ${}^{w}p_s$. We use an iterative and alternating approach because of the non-linearities introduced by the rotation, and because even after linearizing the rotation, the equation remains quartic due to the fact that we minimize the norm of an expression containing a cross product between the unknowns. We set ${}^{w}_{s}R$ and ${}^{w}p_s$ to an initial approximation and solve for one of the unknowns while fixing the other. Then, we alternate and repeat until we converge to an acceptable error. In order to simplify the optimization, we use a first order approximation and three variables to express the degrees of freedom of the rotation. In the following, we describe each part of the alternating least-squares approach.

While solving for ${}^{w}p_s$ using Equation 3.7 and a fixed ${}^{w}_{s}R$ can be directly expressed as a least-squares problem (see Appendix A for details), the rotational part of the problem is less straightforward. The estimate we wish to obtain needs to be a rotation matrix and it should

be expressed with fewer variables as there are only three degrees of freedom. Therefore, we solve for a vector containing those degrees of freedom by linearizing the rotation. We want to solve for a matrix $R = \exp([\omega])$, where ω is the 3-by-1 vector containing the said degrees of freedom, and $[\omega]$ is the skew symmetric matrix that computes the cross product operation $\omega\times$. To use a first-order approximation $I + [\omega]$ of $\exp([\omega])$, we in fact solve for a small change in rotation R such that our estimate at iteration $l + 1$ is updated from the estimate at iteration l using

$$ {}^{w}_{s}R_{l+1} = {}^{w}_{s}R_{l}\ R\ . \tag{3.8} $$

Substituting R, the above equation becomes

$$ {}^{w}_{s}R_{l+1} \approx {}^{w}_{s}R_{l}\ (I + [\omega]), \tag{3.9} $$

where we solve for ω. Consequently, when solving for ${}^{w}_{s}R$ with a fixed ${}^{w}p_s$, we solve for

$$ \underset{\omega}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{j=1}^{m} \| {}^{w}_{s}R_{l}(I + [\omega])\ {}^{s}\tau_{ij} + ({}^{w}p_s - {}^{w}p_{ci}) \times {}^{w}_{s}R_{l}(I + [\omega])\ {}^{s}f_{ij} \|^2, \tag{3.10} $$

where ${}^{w}_{s}R_{l}$ is the most recent estimate for ${}^{w}_{s}R$. We then use least-squares to find ω (see Appendix A for details). We use the exponential map $R = \exp([\omega])$ computed with Rodrigues' formula to update our most recent estimate for ${}^{w}_{s}R$ with Equation 3.8. Note that since ω is used to update rather than compute the orientation, we can bound the norm of ω to a small arbitrary value to perform smaller updates. Bounding the norm of ω is useful to prevent errors caused by large updates, due to the fact that the linearization is only accurate for small rotations.

In sum, the calibration process consists of two steps. First, we calibrate the calibration tool to have accurate contact positions. Then, we record motion and forces as the tool is used to apply a varied set of linear forces at different contact locations rigidly attached to the sensor (e.g., on the front plate or hold). Finally, the recorded forces and torques at the sensor and the corresponding estimated contact position are used to compute a least-squares sensor frame estimate via an iterative and alternating method, where we solve for the rotational part by looking at first-order approximations of small updates.

3.2.3 Pocket-Based Space Calibration

The previous approach is accurate and indispensable when sensors are not mounted in specific locations. However, it is time consuming as it requires several motion and force recordings for each sensor to be calibrated. In cases where sensors are attached in known pockets and orientations, a more expedient calibration method is to place markers on arbitrary locations on the wall and define a wall coordinate frame in which a sensor position and orientation can be expressed. Recall that sensor sandwiches can be inserted in 4 different orientations. The origin of each sensor frame can be computed from the orientation of the sensor sandwich and the location of the pocket where it is mounted. Similarly, the orientation can be set to the appropriate combination of 90-degree rotations about the axis of the wall frame. For example, choosing the origin of the wall frame as the top left corner of the wall, defining the x and y-axis of the frame as aligned with the horizontal and vertical borders of the pockets and using the known size of pockets, sensor position can be precisely extracted. Orientation is straightforward as well because sensors are attached perpendicularly to the wall and obtaining the rotation matrix to get from sensor to wall frame reduces to reorganizing the rows of an identity matrix. The positions and orientations are then easily expressed in the world frame using the position of the wall in that frame, determined by the markers attached to it and motion capture, which allow us to extract the transform from wall to world coordinates.

3.2.4 Measurement Synchronization

Our equipment does not permit a synchronous recording of forces and motion capture. It does not permit capturing forces from all six sensors at once either, because the two data capture cards used to interface the sensor amplifiers record data independently. Therefore, an extra step of measurement synchronization is needed. Similarly to what is done to synchronize sound and video in cinema, we hit one of the sensors from each of the two sensor groups with an object tracked with motion capture. The post-processing step consists of finding the frame with a sharp peak at the beginning of the raw voltage data visualization, and the frame where the tracked body used for synchronization has its momentum change in the

motion capture data. This synchronization peak can be seen near the fifth second of the recording in Figure 3.5. For simplicity, forces, which can otherwise be sampled at rates up to 1000 Hz, are captured at 100 Hz, the same frequency as motion.

3.3 Capture

As in any system involving motion capture, one of the main challenges when it comes to collecting quality data is placing the markers on the subject as well as positioning and directing the cameras such as to avoid occlusions. When capturing a climber, the front is mostly occluded as the body is facing the wall. For that reason, most markers on the subject are placed carefully to face away from the wall during a climbing session, as shown in Figure 3.7. Sometimes marker positions are only fixed after several trials, especially when recording new dynamic motions. For example, in some cases, a knee may be bent entirely, therefore requiring upper leg markers to be moved slightly toward the outside instead of left completely behind the hamstrings. Changing marker locations is however fast as it only requires us to record a new skeleton configuration with the software. Our motion capture setup consists of 24 cameras, six of which have wide-angle lenses. These six cameras are placed close to the wall, above and on each side, and are critical because they cover all the regions occluded from the other cameras. Changing the position and orientation of cameras is time consuming, as calibration for the motion capture setup with our software lasts more than 30 minutes.

For the force sensors, the only concern is to choose gain and excitation parameters that offer high sensitivity while preventing sensor overload. Recall that in our setup, the upper and lower bounds for voltages are respectively 10 V and -10 V. Since sensitivity can only be increased or decreased by a factor of two (see Table 3.1), we favour settings that guarantee no overload over settings that give a wider span of values (and therefore more sensitivity). As shown in Figure 3.5, maximum and minimum recorded voltages are still far from the upper and lower bounds. The choice of the sensitivity parameters is made by trial and error: a simple trial lasting less than a minute is enough to see how voltages

(a) (b)

Figure 3.7: Front (a) and back (b) view of marker placement on the climber.

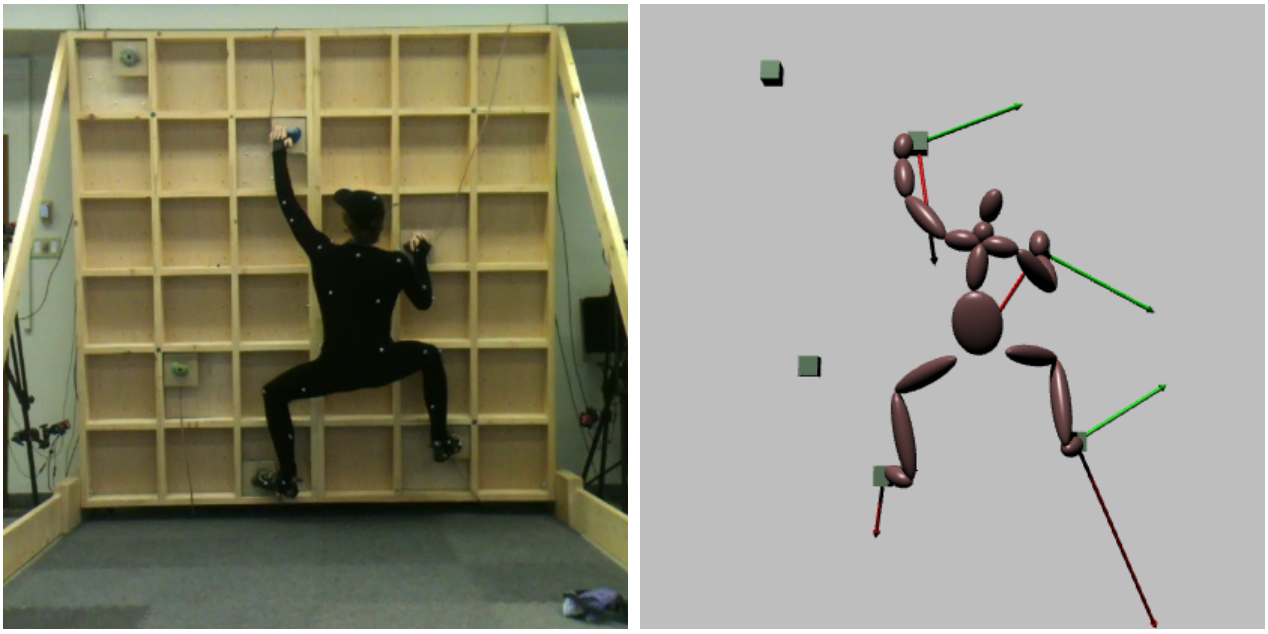
are distributed. Typically, more torque can be generated on the hand holds because we use holds that protrude to allow better grasping, as shown in Figure 3.2. Given that we have three different types of sensors, and taking into consideration the impact of the type of holds and movement on measured forces, choices for gains and excitation voltages vary for each sensor. Table 3.2 shows settings for an example capture session where sensors 1 and 2 are more sensitive models and undergo important torques as they are equipped with larger hand holds. Notice that when there is ambiguity in choosing a gain and excitation voltage for a channel, priority is always given to a higher excitation voltage to minimize noise.

In sum, the pre-processing to perform a successful capture session first consists of placing the cameras to avoid occlusions, calibrating the motion capture setup, and finding the origins and orientations of the sensor frames in the newly defined world frame using one of the two space calibration methods. Then, marker positions and sensor sensitivities are carefully chosen to have quality data by trial and error. Those test trials are run for new skeleton configurations as well as new jumper pin locations for gain and excitation voltage at each channel. During the capture session, the subject uses a rigid object firmly held with a tracked hand to strike a sensor from each of the two sensor groups prior to climbing, and

Sensor	vf_x	vf_y	vf_z	$v\tau_x$	$v\tau_y$	$v\tau_z$	gf_x	gf_y	gf_z	$g\tau_x$	$g\tau_y$	$g\tau_z$
1	10	10	10	2.5	2.5	2.5	1000	1000	1000	1000	1000	1000
2	10	10	10	2.5	2.5	2.5	1000	1000	1000	1000	1000	1000
3	10	10	10	5	2.5	5	2000	2000	2000	1000	1000	1000
4	10	10	10	10	10	10	4000	4000	4000	1000	1000	1000
5	10	10	10	10	5	10	4000	4000	4000	1000	1000	1000
6	10	10	10	5	5	10	4000	4000	4000	1000	1000	1000

Table 3.2: Excitation voltages (in V) and gains for an example trial.

then performs the desired task. The first step of post-processing consists of using the knocks to synchronize the two force data streams with the motion stream. In order to perform this synchronization, we use the sharp peaks in the early force samples for each group, and the frame where the tracked hand abruptly changes momentum in the vicinity of the appropriate sensor. Then, we extract voltages from the unsigned integer output and use the calibration matrices to convert the recorded voltages to force measurements. Finally, we use the sensor frames estimated with the space calibration procedure to express the measured forces in the motion capture world frame. Figure 3.8 depicts the results of a successful capture, where green squares represent sensor locations, red arrows represent forces measured by the sensor, and green arrows represent the torques. Note that while the arrows offer useful information about relative magnitudes, forces and torques are scaled independently and arbitrarily to fit reasonably within the image.



(a) (b)

Figure 3.8: Climber during a trial (a) and obtained visualization (b).

Chapter 4

Reconstruction of Static Poses

As a first step in understanding how motion and forces are produced during climbing, we examine the relationship between posture and contact forces. Note that in this work, we use the terms "contact forces" and "forces measured by the sensors" interchangeably. This is because forces (as opposed to torques) are not affected by the distance between contact and sensor. To study the relationship between forces and motion, several capture trials with nine male and female participants of various skill levels were recorded. During each session, climbers were asked to perform dynamic and static motions. For each participant, close to twenty trials lasting from thirty seconds to a minute were performed. Finding a relationship between measured forces and observed pose is difficult, and we distinguish between two forms of ambiguity: similar poses can generate different forces, as shown in Figure 4.1, and similar forces can be generated by different poses. To reduce ambiguity, we restrict our focus to simpler problems where data consists of static poses and quasi-static transitions. For the trials that we consider in this chapter, participants were asked to stay in contact with a given set of holds and move slowly while performing natural and controlled movements to ensure as little dynamic movements as possible. In order to have a wide variety in poses and forces, climbers were asked to move their centre of mass as far as possible toward and away from the wall, as close and as far as possible from the ground, and as far as possible to their left and right. Controlled transitions, which offer a wide variety of additional poses, are also part of

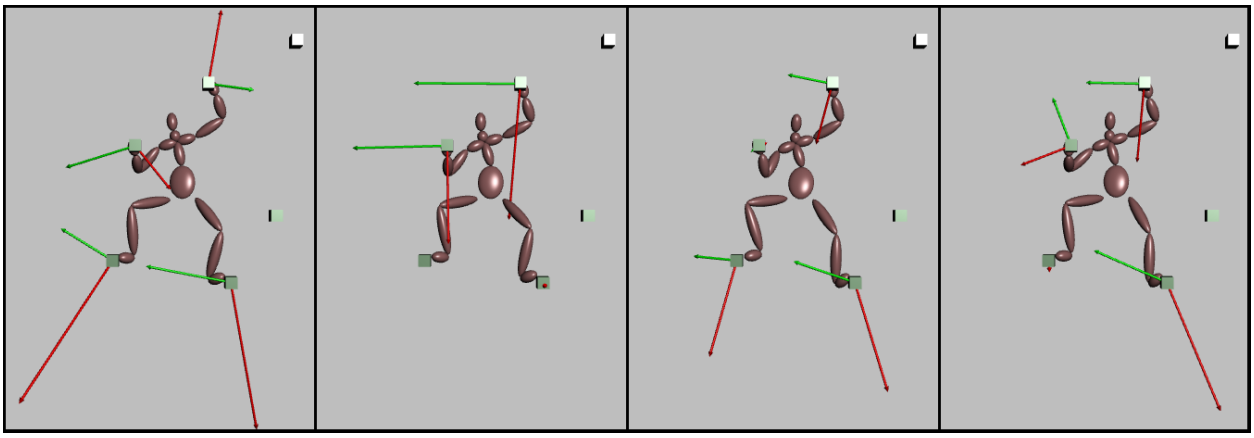


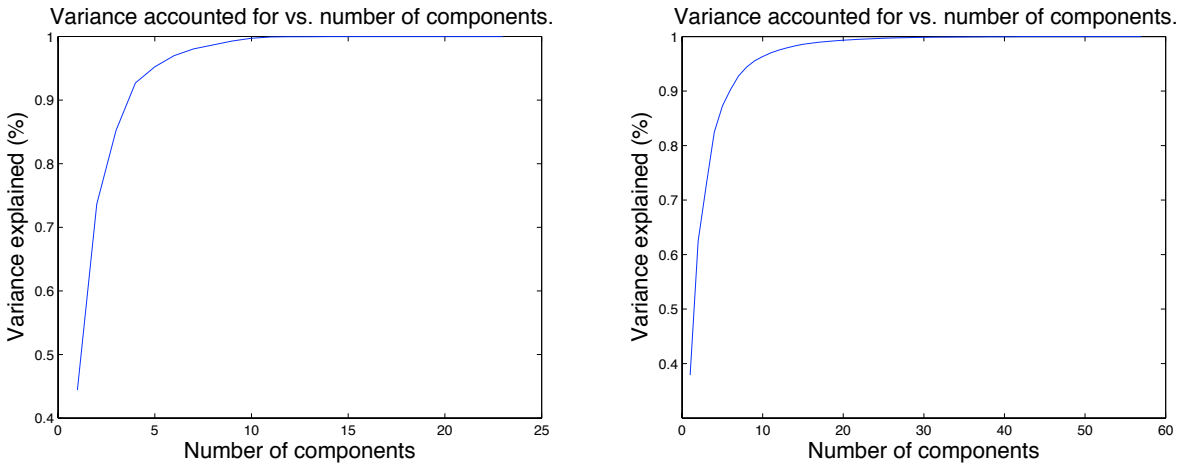
Figure 4.1: Subject generating arbitrary forces while maintaining a pose.

the considered data. In the static case, the pose determines the contact forces, but the forces do not necessarily determine a pose. However, we believe that there is correlation between forces at the hold and corresponding pose due to the way people select natural postures for different situations. Below, we focus on reconstructing the climber's posture from a given set of hold positions and forces measured at the sensors.

4.1 Statistical Analysis

We take a preliminary look at the data to see how much of the posture and force spaces is exploited by various climbers. The skeleton pose for a frame is given as the three coordinates of the root (or hips) in the world frame, followed by the three Euler angles describing the orientation of the root in the world, and the three Euler angles describing the joint angle for each of the 17 joints in the frame of the parent body in the Biovision Hierarchy[1] (BvH). The corresponding set of forces is given by the forces and torques in world coordinates measured by the sensors at the four holds. A principal component analysis [Jol02] of the two data sets reveals that only a few components are necessary to capture most of the variation in the data in both cases, as shown in Figure 4.2. The lower dimensionality of both data sets can be explained by the highly constrained nature of the problem: hands and feet are fixed, and

[1]BvH is a standard format for recording skeleton poses with motion capture [bvh].



(a) (b)

Figure 4.2: Variance explained against number of components for forces (a) and poses (b).

forces generated at the holds in the case of slow, controlled, and natural movement can be explained by gravity.

In static or quasi-static poses, forces contain a lot of information on the location of the centre of mass: without momentum, forces at the holds are highly dependent on the distribution of the climber's weight, as shown in Figure 4.3. We investigate the correspondence between forces and static postures by performing linear regression [SL12] and canonical correlation analysis [Tho00]. We use the kernel-trick [BA01] to obtain polynomial regression models from linear regression. For simplicity, and because we do not know the exact mass of the different limbs, we approximate the centre of mass by the location of the root of the skeleton (i.e., the hips). The location of the root can be predicted using first-order polynomial regression with a maximum error in the order of 10 cm, as shown in Figure 4.4.

Despite accurately predicting hip position, regression fails to predict angles from forces. Figure 4.5 shows the difference between predicted and actual x component of the right hip angle: using more expressive functions for the regression such as second-order polynomials

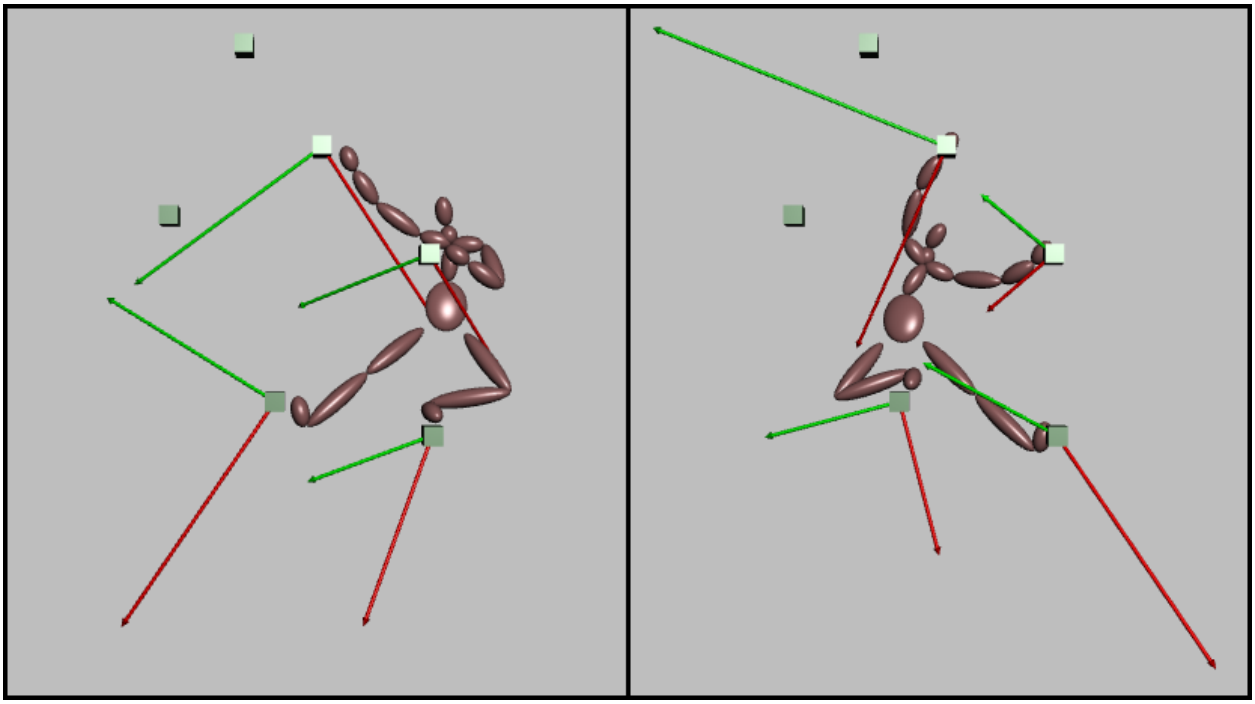


Figure 4.3: Correlation between forces and centre of mass position for static poses.

does not help. An explanation is that multiple linear regression is not suited for Euler angles (which is the joint angle representation provided by the motion capture software) as they are not independent, as they combine in a non-linear way, and as they are not a unique representation of an orientation. As such, canonical covariance analysis is not helpful in explaining joint angles from forces at the holds, as illustrated in Figure 4.6. Using once again differences between predicted and actual angles, the grey-scale colour-map illustrates that the majority of samples for different joint angles offers unacceptable error. In fact, rotations are not a linear space and linear analysis tools such as regression or principal component analysis are not appropriate in this case, regardless of the format in which orientations are expressed. While a better statistical method would be principal geodesics analysis [FJLP04], we choose to take a physics-based approach to tackle the problem, in the interest of finding a physical relationship to interpret and model a hypothetical correlation.

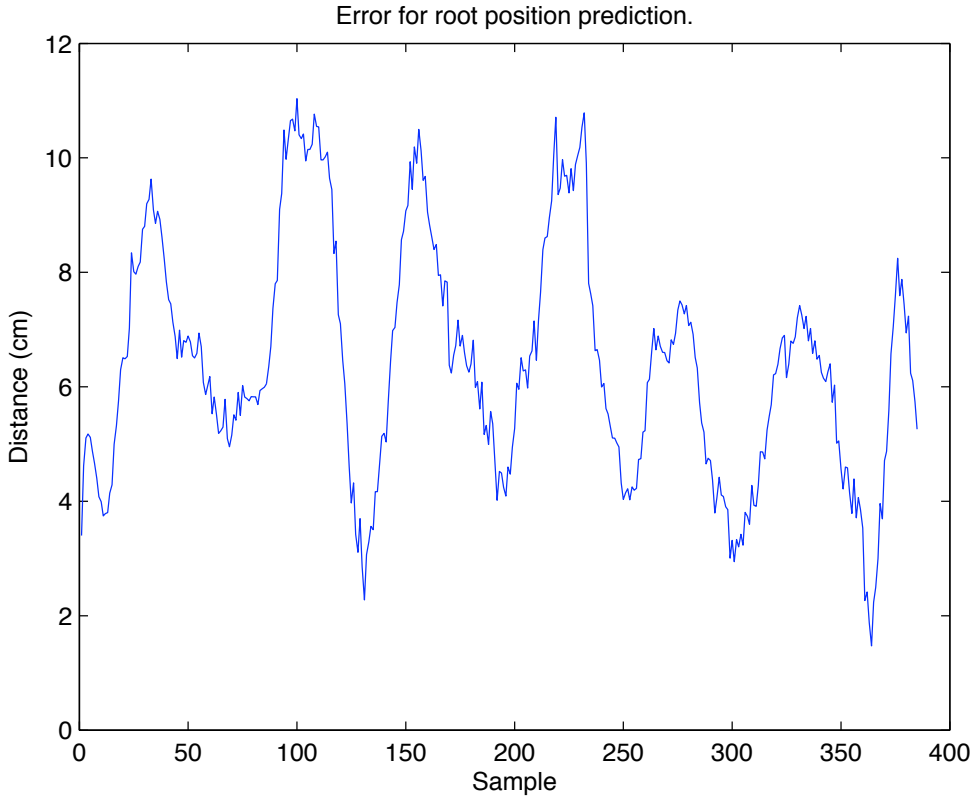
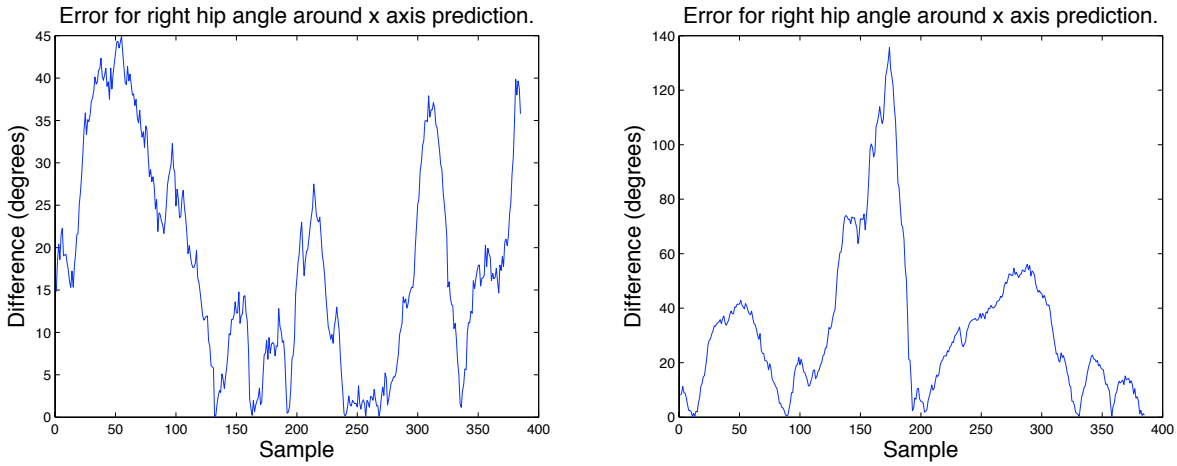


Figure 4.4: Root position regression error using first-order polynomials.



(a) (b)

Figure 4.5: Joint angle regression error with first- (a) and second-order polynomials (b).

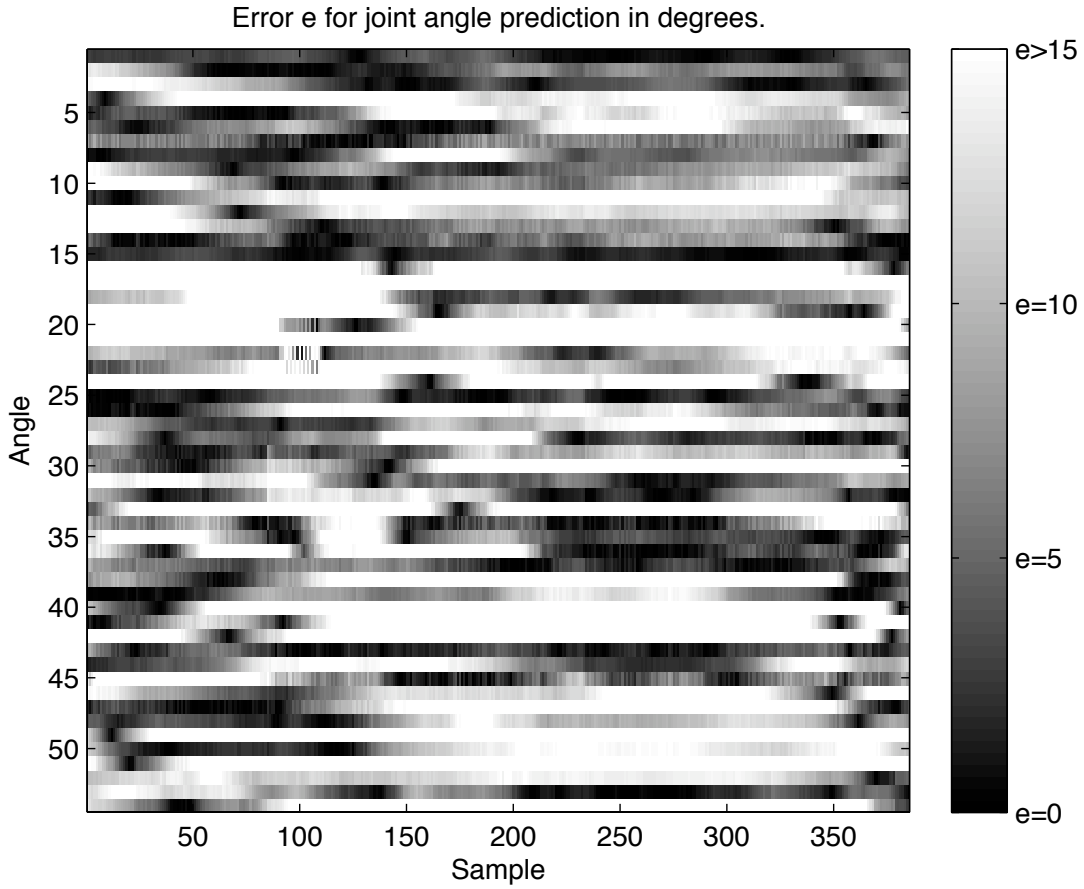


Figure 4.6: Joint angle prediction error using canonical covariance analysis.

4.2 Physics-Based Approach

In order to take advantage of the climbing-specific constraints and to account for physical plausibility, we use a physics-based simulation – where physical plausibility of a pose is a measure of the likeliness of that pose given the physical constraints. Specifically, hands and feet are constrained to locations given by the four sensors at which forces are non-zero, the climber must face the wall, he cannot penetrate it, has joint angle limits, and is subjected to gravity.

4.2.1 Physical Simulation and CMA-ES

In this section, we describe the simulation that captures the physical constraints, the optimization method used to reconstruct a pose from forces at the holds, and the projection step required to interface the optimization with the simulation.

Simulation

The virtual character used to simulate the climber has its dimensions defined by the captured climber (stored in the BvH format). The said character is modelled using kinematic chains: it is a set of rigid bodies attached using different types of joint constraints, shown in Figure 4.7. Joint angle limits for each articulation are chosen arbitrarily to prevent unnatural poses (e.g., poses where a knee is bent backwards), while ensuring that the skeleton can reach all the recorded motion capture poses. The mass of the skeleton is proportional to that of the climber, and the mass distribution among the different limbs is chosen using anthropometric data from a study by Kreighbaum et al. [KB96].

To simulate contacts between the hands and feet and their corresponding holds, we use joint constraints and constrain the end effectors to fixed locations. Those locations are determined by assuming the climber is facing the wall, is upright, and is exerting force on the holds (i.e, each end-effector is assigned to one of the four holds where force is measured, e.g., the right foot rests on the lowest, rightmost of these holds). Finally, contacts between the character and a simulated climbing wall are used to prevent invalid poses.

The simulation is built using the Open Dynamics Engine[2] (ODE). Note that solving an optimization involving such rigid constraints can be done using linear or quadratic programming methods in the case of static poses. We choose a physical simulation instead with the intent of exploring dynamics in the future. In the case of dynamic movement, several space-time constraints as well as energy in the system need to be taken into account (e.g., a

[2]ODE is an open-source, high-performance library for simulating rigid body dynamics [Smi].

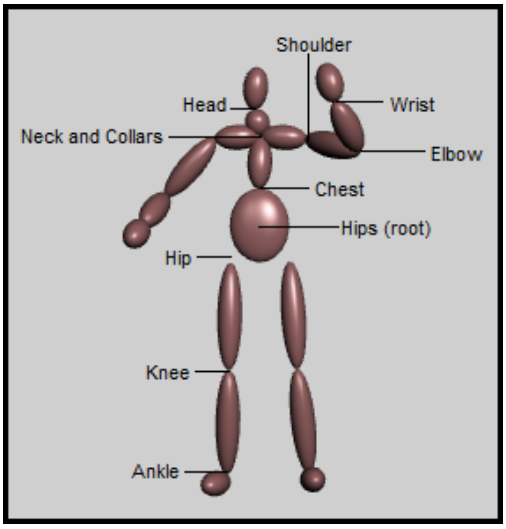


Figure 4.7: Skeleton model used for simulation.

hand can attach to different holds over the course of a dynamic movement).

CMA-ES

We want to find a pose that best explains an arbitrary function – which we have yet to define – of forces at the holds. Treating this as an optimization, we want the pose that minimizes our objective while taking into account the constraints enforced by the simulation. In our particular case, a pose is described only using angles at each joint. Root position and orientation are not needed anymore since end effectors are constrained to the hold positions, and therefore joint angles suffice to define a unique pose. In other words, we are looking for the set of joint angles

$$\underset{\phi' \in \Phi^*}{\operatorname{argmin}} f(\phi', F_S), \tag{4.1}$$

where Φ^* is the manifold of joint angles that describe poses respecting the constraints, f is the objective function, and F_s is the set of linear forces at the four holds, measured by the sensors. We choose CMA-ES to perform this optimization because we evaluate our objective function via simulation. CMA-ES is an iterative, derivative-free, sample-based optimization method where the fitness of samples is used to update the sampling multivariate normal distribution. At each iteration, a number of samples is generated using the current sampling

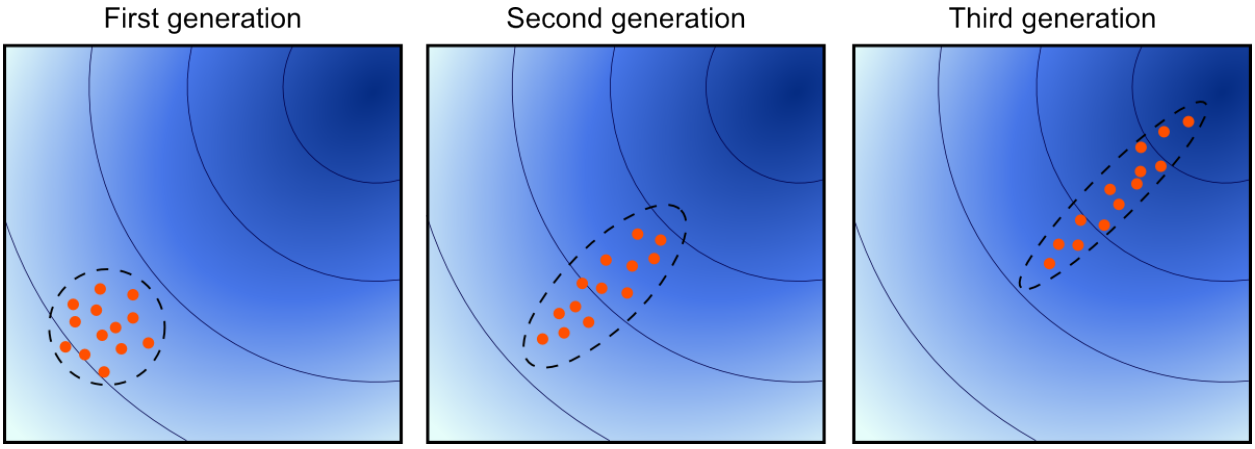


Figure 4.8: Illustration of the behaviour of CMA-ES (used with permission) [cma].

distribution. The fitness of these samples is computed. The mean as well as the covariance matrix of the sampling distribution is then updated for the next iteration such as to maximize the likelihood of fit samples. As shown in Figure 4.8, the multivariate normal distribution adapts the shape and centre of the sampling ellipsoid to move toward the optimum while directing the search along the path to the optimum, located in the top right corner.

Projection Step

We use CMA-ES to find the set of joint angles describing a pose that respects the physical constraints and optimizes our objective. CMA-ES produces sets of arbitrary Euler angles ϕ using a multivariate normal distribution. However, we are interested in joint angles that describe poses satisfying the constraints enforced by the simulation. Therefore, we devise a projection step Γ to find, for a given arbitrary set of joint angles ϕ, the closest set of angles $\phi' = \Gamma(\phi)$ that describes a valid pose on the constrained pose manifold Φ^*.

This projection is achieved using a proportional-derivative controller, by starting with a valid pose and treating the produced angles ϕ as desired angles at each joint. In other words, we start from a pose satisfying the constraints (which can be taken from motion capture data), and we apply torques at each joint individually to try and match the angles

produced by CMA-ES, using differences in orientations, as well as stiffness and damping parameters that offer a trade-off between speed and stability. After a small amount of time, the simulation stabilizes into a new pose with joint angles ϕ' that is treated as the projection of the pose described by ϕ onto the constraint manifold.

The fitness of the original sample ϕ is computed using the projected pose given by ϕ', obtained after performing the projection step Γ. Note that a minimum coordinate representation could alternatively be used as a search space, but with four fixed end effectors, we find that simulation is an easier way of searching the space of valid poses (i.e., the space of poses that satisfy the physical constraints).

4.2.2 Physical Plausibility as an Objective

In order to find the pose that best explains forces at the holds, our objective function evaluates the fitness of a CMA-ES sample ϕ by computing the physical plausibility of the pose given by $\phi' = \Gamma(\phi)$, with the additional physical constraints that the pose is static and generates the contact forces measured by the sensor. The plausibility for that pose is computed as

$$\rho(\phi', F_S) = \|\tau_r\|^2 \tag{4.2}$$

where τ_r is the torque generated at the root (or hips), where the contact forces F_S at the sensors and gravity are acting on the system, and where the system is static. In other words, we look at the constraint torque generated to enforce statics for the pose given by ϕ' when the forces acting on the system are gravity and the forces measured by the sensors. We first explain how we compute τ_r and then justify why it is a good measure of physical plausibility.

We find the plausibility ρ using the ODE simulation. Rigid constraints are set on skeleton joint angles to enforce the pose given by ϕ', while another fixed constraint is applied to the root position and orientation in the world to enforce statics. In order to simulate the forces generated by the holds on the hands and feet, the opposite of the forces measured by

the sensors are applied to the appropriate end effectors. Gravity is also taken into account. The constraint torques and forces that are necessary at the various joints and at the root to maintain a static pose are computed using ODE, and their values give indications as to the plausibility of the pose for the recorded forces. We first examine the information that can be drawn from the wrench applied at the root to enforce statics, and we then focus on the torques at the inner joints to maintain the pose.

The wrench that needs to be applied to the root to enforce statics contains information on the physical plausibility of the pose for F_S. Typically, for a static pose, resultant forces and torques at the root are zero as gravity and the sum of the forces generated by the holds cancel out. A large torque means that there is a rotational acceleration on the root and that the system would not be static were it not for the constraints. Assuming forces at the holds are explained by the position of the centre of mass, as illustrated by Figure 4.3, we can always hope to find a pose that minimizes torque at the root by distributing the skeleton's weight appropriately. Linear force at the root however is the difference between the chosen weight of the skeleton and the sum of forces at the holds and is therefore pose-invariant.

Torques at the inner joints contain critical information as they describe the necessary effort to maintain the pose described by ϕ'. A small total inner torque is in general a good measure of high human plausibility as it signifies that it is more natural to choose a pose that requires less effort – where human plausibility of a pose is a measure of the likeliness that a human produces that pose. However, in the case of our simulation, minimum overall inner torque does not indicate that a simulated pose is a more natural choice. This can be explain by the simplicity of our simulated model, which does not capture properties crucial in evaluating effort for climbing. For instance, the shape of the holds is crucial to account for the difficulty of grasping and therefore of applying force along a given limb (e.g., it is easier to pull with the arm when the hand is grasping a hold that offers a comfortable grip). Also, different subjects use distinct strategies to distribute effort over the muscle groups

(e.g., experienced climbers use their legs more). Therefore, using inner joint torque as a measure of the likeliness that a human produces a pose requires more information on the force that a subject can generate with the different muscles, on the strategies used to distribute effort, as well as on the shapes of the holds and their repercussions on those strategies.

For the above reasons, we only use the rotational part of the wrench that needs to be applied to the root to enforce statics as a measure of the quality of the pose given by ϕ' for explaining F_S. The closer that torque is to zero, the more plausible the pose. Consequently, finding a unique pose with physical plausibility as an objective is under-constrained: the set of centre of mass positions that minimize the said torque lie on a line parallel to the linear part of the wrench applied to the root to enforce statics. To incorporate the missing information on human plausibility and guide our optimizer toward a smaller set of good solutions, we use statistical information gathered on the position of the root for a given set of forces. We believe that this statistical information captures constraints associated with the shape of the holds as well as with biomechanical properties of the climber.

4.2.3 Hint objective for improved optimization

While entire poses could not be estimated from forces using linear models even in the static case, the aforementioned missing information concerning holds and human morphology can be partly extracted from the data to help the optimization converge. As explained earlier, linear regression can be used to accurately estimate root position with a maximum error in the order of 10 cm.

The model to predict the position of the hips for a given set of forces is input into CMA-ES and used to help reduce ambiguity. During initialization, r', the expected root position, is computed with the regression model given the forces measured by the sensors. The distance between the current location of the hips for a candidate pose and the estimated root position for the sensor forces is used as a hint to guide the optimization toward a good solution. Since

the regression error is in the order of 10 cm, we use r' to guide CMA-ES toward the set of poses with a root position inside a 10 cm sphere from the estimate.

Conceptually, we guide our pose toward a configuration where the centre of mass, which we approximate as root position, lies in a sphere, and use physical plausibility to project that centre of mass onto the line that minimizes constraint torque applied to the root. In practice, the optimal solution is described as the most physically plausible pose where the hips are within a 10 cm radius of the expected hips position.

In sum, the optimization procedure starts with an arbitrary initial pose that satisfies the constraints. Linear regression is used to find r', the predicted root position for forces F_S. For a new CMA-ES sample ϕ, the projection $\phi' = \Gamma(\phi)$ on the constrained joint angles manifold is found using a proportional-derivative controller. The actual position of the root r is given by the pose described with ϕ'. The fitness – which we want to minimize – of the sample ϕ for explaining F_S is defined as

$$\zeta(\phi) = \begin{cases} k \left\| r - r' \right\|^2 + G(\phi') & \text{if } \left\| r - r' \right\| > t \\ \rho(\phi', F_S) + G(\phi') & \text{otherwise} \end{cases}, \tag{4.3}$$

where t is an estimate of the maximum regression error, used to determine the significance of the information provided by r', $G(\phi')$ is a set of linear penalties applied when the skeleton is not facing the wall, and k is a large weight that ensures CMA-ES discards samples where the hips lie outside the sphere given by the regression. To penalize solutions where the skeleton does not face the wall, $G(\phi')$ grows linearly with the orientation of the root about the y-axis. Since the distance between actual and expected root positions r and r' is inexpensive to compute for a given pose, the physical plausibility ρ is only evaluated when the solution is promising in terms of human plausibility.

A limitation of this fitness function is its discontinuity. However, and because of the weight k, poses with a hip position within the sphere given by the model are always fitter.

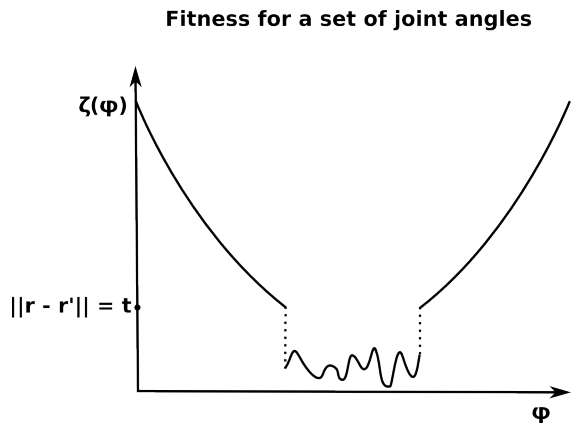


Figure 4.9: Illustration of the shape of the fitness function.

Therefore, there is no risk that CMA-ES oscillates back and forth between both sides of the discontinuities. Figure 4.9 gives a sketch of the behaviour of the objective function.

CMA-ES is guided toward the promising set of poses with reasonable human plausibility using an inexpensive term. When in the subset of promising poses, CMA-ES uses physical plausibility to find an optimal solution. Figure 4.10 illustrates the optimization process for an example optimization using CMA-ES with 10 samples per iteration and 100 total iterations. The figure illustrates the exploration of pose space, starting from an arbitrary pose and converging toward the optimal pose, where each picture contains the population cloud for a given iteration (indicated in the top-right corner). Our projection method allows an easy exploration of the set of valid poses, and the optimization, after exploring those poses in the initial iterations, converges when in the vicinity of the optimal solution by progressively lowering the variance for each of the sampling variables.

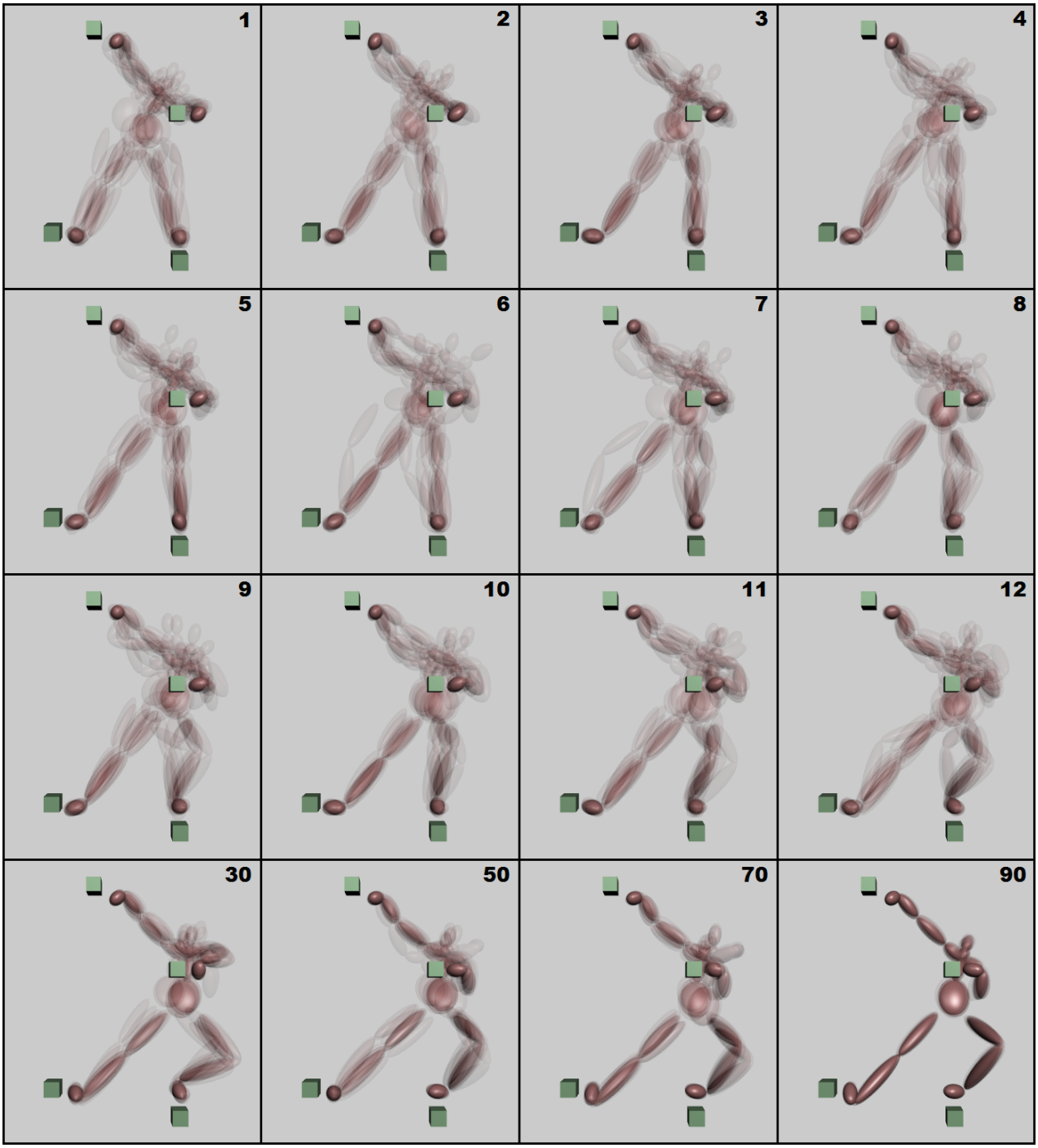


Figure 4.10: Visualization of the optimization process.

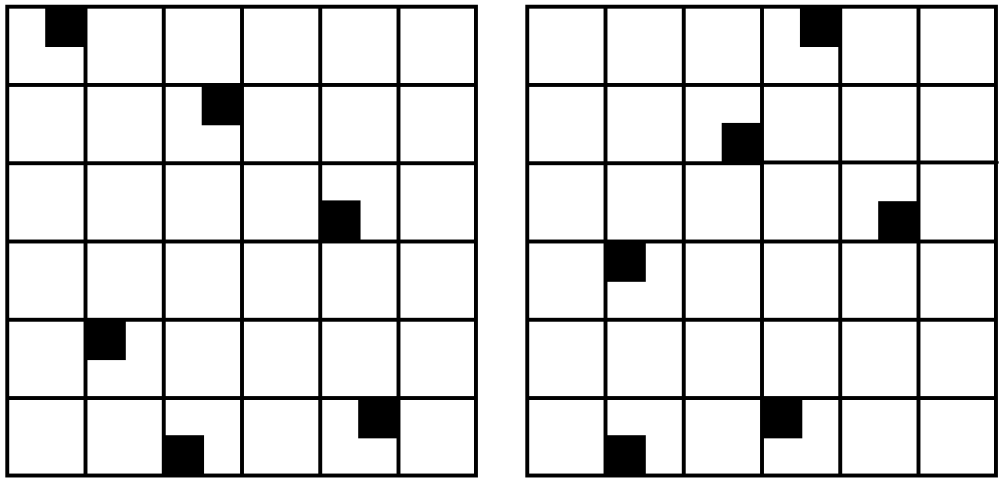
Chapter 5

Results and Discussion

In this chapter, we present the results and an evaluation of their quality for the calibration of the wall and for the reconstruction of static poses. We present error measurements for expressing the forces in the motion capture world frame, where we first analyze the performance of our calibration tool, and then discuss the error for sensor position and orientation estimates. Then, we discuss static pose reconstruction, where we first expand on the quality of the regression for predicting hips position, then detail our choices for the parameters of CMA-ES, and finally compare the optimal poses to ground truth motion capture measurements. Timings are given for a dual-core 3GHz Intel i3 CPU with 4GB RAM.

5.1 Calibration

Our calibration procedures were used to capture nine different subjects and collect close to an hour of climbing data. Each subject was asked to repeatedly perform several tasks with different objectives such as shifting centre of mass position and completing routes following different progressions. Two different hold patterns (shown in Figure 5.1) were designed by an expert climber, and different types of holds were used in each position following the level of expertise of the subject. Our setup supports the expedient calibration procedure: in practice, most of the capture sessions were recorded using the pocket-based procedure. Three



(a) (b)

Figure 5.1: First (a) and second (b) bouldering puzzles.

markers are placed in corners of the wall to define the coordinate frame in which sensor frames are expressed as a function of pocket placement.

However, several small-scale tests were performed for the optimization-based calibration, as well as a full wall calibration. For each sensor, three contact point locations are considered. For each contact location, 20 seconds of motion capture and 20 seconds of force measurements are recorded. Using Matlab, calibrating the calibration tool from 2500 frames of motion capture takes less than a second, while estimating the position and orientation of a sensor with the iterative and alternating method from 3 contact locations and 2000 motion capture and force samples per point takes less than a minute. We present results for contact point as well as sensor frame estimation.

5.1.1 Contact Points

After calibrating the tool, we use the least-squares estimate ${}^{b}p_c^*$ (i.e., the position of the sharp tip in body coordinates) for estimating the sensor frames. Motion capture and forces are recorded as the tool is used to apply forces at the holds. The position of the contact in

world coordinates ${}^{w}p_c^*$ is then estimated from the motion capture and ${}^{b}p_c^*$ using the matrix ${}^{w}_{b}E_j$ (provided by the software at each frame) that describes the orientation and position of the calibration tool in the world. We define the error as

$$\epsilon_c = \frac{1}{n}\sum_{j=1}^{n} \left\| {}^{w}_{b}E_j \; {}^{b}p_c^* - {}^{w}p_c^* \right\|, \tag{5.1}$$

where

$${}^{w}p_c^* = \frac{1}{n}\sum_{j=1}^{n} {}^{w}_{b}E_j \; {}^{b}p_c^*. \tag{5.2}$$

This error is zero when ${}^{b}p_c^*$ is the exact location of the tip of the calibration tool in the body frame, when ${}^{w}_{b}E_j$ gives the exact orientation and position of the tool in the world, and when the contact location is perfectly constant throughout the trial. In other words, it is zero when ${}^{w}_{b}E_j \; {}^{b}p_c^*$ is frame-invariant.

The error ϵ_c can be seen as the average radius of the point cloud ${}^{w}_{b}E_j \; {}^{b}p_c^*$ where the estimated contact points lie in world coordinates. A 2D view of an example cloud is shown in Figure 5.2, where the individual contact point location estimates in world coordinates ${}^{w}_{b}E_j \; {}^{b}p_c^*$ for each frame are in blue, and their average ${}^{b}p_c^*$ is in red. Table 5.1 shows the averaged estimation error ϵ_c as well as averaged standard deviation σ_c for the point cloud for the three contact points at each hold during a full calibration procedure.

Several factors can explain the magnitude of these error and standard deviation. First, the matrix ${}^{w}_{b}E_j$, that describes the orientation and position of the tool at each frame, is estimated by the software from the position of four markers placed on a flat surface of the tool, as shown in Figure 3.6. The reported position of each of these markers is subjected to noise in the order of 0.3 mm from the motion capture setup. Therefore, the matrices that we use to calibrate the calibration tool as well as to compute contact estimates are noisy. This explains the high estimation error, but also the large standard variation for the point clouds. Also, the calibration procedure is done by hand and there is room for slight displacements of the contact locations when the tool is rotated around for calibration as well as when it is

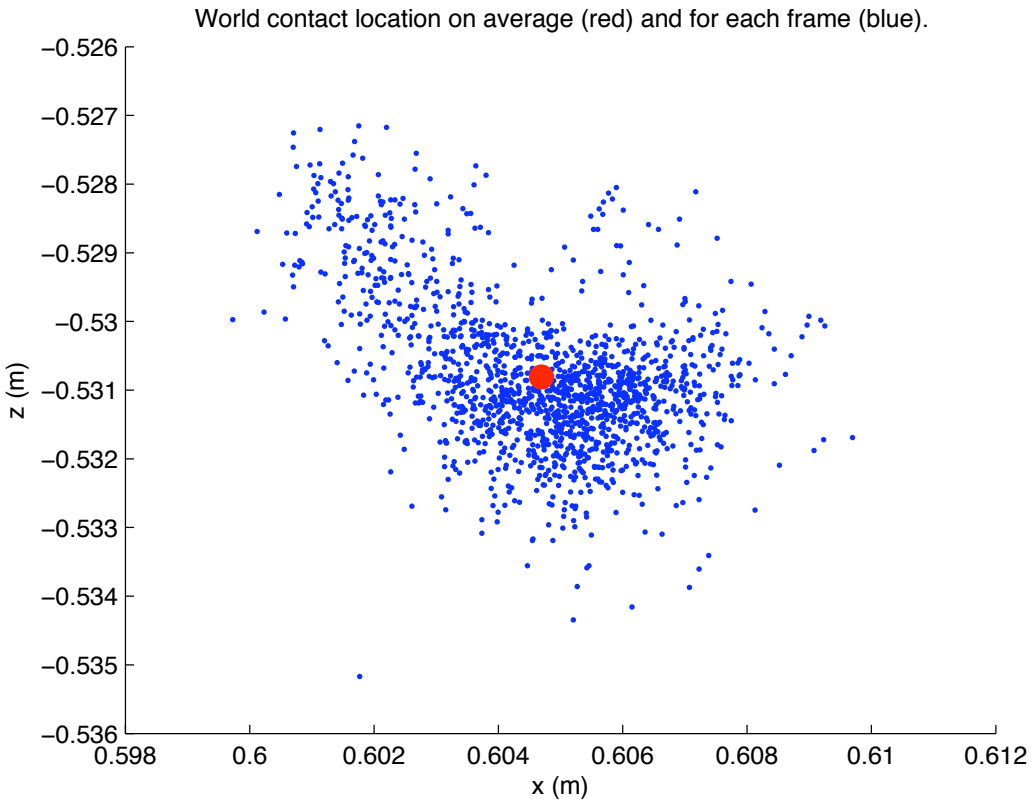


Figure 5.2: Point cloud illustrating contact point estimation error.

Sensor	**Average ϵ_c (mm)**	**Average σ_c (mm)**
1	4.3	7.7
2	3.3	6.1
3	4.6	5.2
4	3.4	5.7
5	4.4	8.6
6	4.1	9.4

Table 5.1: Averaged errors and standard deviations for three contact point estimates.

used to apply the forces in different directions. Finally, the tip of the tool is not perfectly sharp, and has a diameter close to 0.5 mm.

5.1.2 Sensor Frames

Sensor frame estimate error is expressed as the mean length of the torque vector at each world contact location ${}^{w}p^{*}_{ci}$, estimated from each motion capture trial, using the least-squares estimates ${}^{w}_{s}R^{*}$ and ${}^{w}p^{*}_{s}$, which describe the sensor frame, and the forces ${}^{s}f_{ij}$ and torques ${}^{s}\tau_{ij}$, measured for each corresponding contact at each frame of the force capture trial. For a given sensor, frame estimation error is given by

$$\epsilon_s = \frac{1}{kn}\sum_{i=1}^{k}\sum_{j=1}^{n} \| {}^{w}_{s}R^{*}\ {}^{s}\tau_{ij} + ({}^{w}p^{*}_{s} - {}^{w}p^{*}_{ci}) \times {}^{w}_{s}R^{*}\ {}^{s}f_{ij} \| . \tag{5.3}$$

This error is zero when each contact location is perfectly constant throughout the trial and is perfectly described by ${}^{w}p^{*}_{ci}$, when ${}^{s}f_{ij}$ and ${}^{s}\tau_{ij}$ are the exact forces and torques applied in the sensor frame for the corresponding contact point, when ${}^{w}_{s}R^{*}$ and ${}^{w}p^{*}_{s}$ give the exact frame of the sensor, and when there is indeed zero torque applied at the contact.

Table 5.2 shows the sensor frame estimation error ϵ_s for each of the six sensors during the full wall calibration. The iterative solver stops when the change in error is less than 0.01 Nmm per iteration. Since a torque is a force applied at a distance, and given that in our setup, contacts with holds occur approximately 15 cm from sensors, an error of 75 Nmm at the sensor represents a weight discrepancy of a mere 50 grams at the hold. Since forces applied at the holds during the calibration and also during the capture trials are two to three orders of magnitude larger, this is a very satisfying result. Note that no synchronization step is performed for that calibration as we average contact point location over all the recorded motion capture samples. Motion and forces are however recorded approximately at the same time, and we make sure that forces are applied at a single contact point over the whole duration of the two streams.

Sensor	ϵ_s (Nmm)
1	75
2	55
3	92
4	87
5	102
6	96

Table 5.2: Errors for sensor frame estimates.

Sensor frame estimation error can be explained by the contact point estimate error, which is in the order of 4 mm, the sensor noise, which is in the order of 0.1 N for forces and 1 Nmm for torques, and human error, as minor torques are sometimes involuntarily applied at the contacts due to friction and slight slips at the contact point.

A simple validation technique is to compare the climber's weight to the weight derived from measured forces expressed in the world frame for static poses. While this does not provide an accurate measure of the calibration precision, it is a good means of detecting errors in recorded gains, excitation voltages and sensor frame estimates. For the pocket-based calibration, this step is critical as it is its main measure of quality. In practice, the weight computed using the sensor calibration was always accurate with respect to the reported weight of the climber.

Note that similar to what is done for pocket-based calibration and in the case where sensors are attached to a fixed surface, markers can be used to avoid performing sensor frame calibration every time a new motion capture calibration is required. Markers can indeed be placed in arbitrary locations on the fixed surface to define a coordinate frame in which the sensor frames can be saved. Since the markers are fixed to the surface, the location of the sensors in surface frame can be re-used as long as the sensors are not moved. After

a new motion capture calibration, the position of the markers can be used to retrieve the sensor frames in the new world coordinates.

5.2 Reconstruction of Static Poses

Pose reconstruction was run on several, varied static poses. The reconstruction is compared to ground truth for 5 of these poses in Figure 5.3. We first present the results for estimating the root position using regression. Then, we focus on the details of the CMA-ES setup by looking at the chosen parameters. Finally, we review results of the full optimization by quantifying the reconstruction error.

5.2.1 Regression

Using a data set containing a wide variety of static poses and over 4000 frames of synchronized motion capture and force data, we look at how well first-order polynomial regression on the position of the root generalizes. We train the model on 90% of the samples and test it on the remainder. Our error metric for the regression is simply the distance between the motion capture ground truth and the predicted position of the hips for the corresponding forces at the sensors. Table 5.3 shows the best, average, and worst reconstruction error, using our error metric, for six different 90/10 partitions. As expected, we noticed that dynamic motions and higher errors in root location estimation coincide, confirming our hypothesis that the correlation is more pronounced in the static case.

5.2.2 CMA-ES

The sampling variables that we use for CMA-ES are the desired Euler angles for the chest, shoulders, elbows, wrists, hips, knees and ankles (See Figure 4.7 for an illustration of the different joints). We choose to ignore collars as well as neck and head joints, and we ignore Euler angles that are irrelevant for a given joint. Collars are ignored because of their limited range of motion and because their orientation is directly affected by the disposition of the

Test	Best (cm)	Average (cm)	Worst (cm)
1	0.95	5.99	10.49
2	0.72	3.37	7.94
3	2.03	5.44	11.07
4	2.59	7.11	12.60
5	1.31	5.72	9.39
6	0.53	5.42	13.64

Table 5.3: Root position regression error.

body. Neck and head joints do have an impact on centre of mass location, but they are left out because climbers do not use their heads to shift their weight in practice, instead keeping it in a natural posture. For these joints, the desired joint angles are the joint angles for the rest pose. In other words, the collars are encouraged to remain in a natural pose and the head and neck are kept straight, which we find is a good strategy to reconstruct human poses. Finally, for some articulations, only one or two Euler angles are needed to explore the set of possible configurations. The knee for example is simulated using a hinge joint, and only one Euler angle is used to describe the desired angle between the upper and lower leg around the axis of the joint. In total, a CMA-ES sample is 24-dimensional. This number is small compared to the 57 values used to describe all joint Euler angles as well as root orientation and position in the constraint-free case, but it is still a considerable number of variables for an optimization.

The parameters we choose for CMA-ES are 10 samples per iterations and a total of 100 iterations. Having a small population size causes the covariance matrix to be updated very regularly and provides a more local search at each iteration, as illustrated by Figure 4.10. The choice of parameters is very problem-specific. In contexts where local minima are spread over a wide portion of the space of possible solutions, it is important to have a very large covariance and number of samples per iteration to properly explore the solution space.

However, the minima for our distance metric (that captures human plausibility) belong to a set of poses that share the same root position. As for our physical plausibility metric, the local minima belong to a set of poses which hips positions lie along a line L. Indeed, recall that we are trying to minimize the torque generated by the constraint force at the root: the said torque is invariant over any line parallel to that force because the two quantities are related by a cross product. Conceptually, the minima of our combined fitness metric therefore belong to the set of poses with root positions lying on the portion of the aforementioned line L that is within the 10 cm maximum error sphere of our distance metric. Given the physical constraints (such as hold positions and joint angle limits), these minima are very similar in practice, and are concentrated in a small portion of the pose manifold (or solution space). In sum, our context-specific choice of parameters enables subtle search near the mean of the sampling distribution without compromising convergence toward the optimal solution. Finally, and in order to explore a varied set of poses, we choose 30 degrees as the initial standard deviation for the generated Euler angles. The CMA-ES optimization process for the third example from Figure 5.3 is visualized in Figure 4.10. Recall that while there are other means of solving this optimization, an important motivation for the current framework that solves optimization problems using physical simulation and CMA-ES is that it can be extended to solve complex dynamics problems with space and time constraints.

5.2.3 Reconstruction

Reconstruction error is computed as the average distance between geometric centres of corresponding skeleton bodies for the optimal solution and the motion capture ground truth. In other words, reconstruction error is computed as

$$\epsilon_r = \frac{1}{m} \sum_{i=1}^{m} \left\| {}^{w}b_i - {}^{w}b_i^* \right\|, \tag{5.4}$$

where m is the total number of bodies in the skeleton (in our case, 18), ${}^{w}b_i$ is the position of the geometric centre of body i in world coordinates in the simulation, and ${}^{w}b_i^*$ is the position of the corresponding body in the ground truth motion capture data. We choose to compute

reconstruction error using body positions because it is simple in terms of computation as well as interpretation. Also, this distance metric allows us to compare average reconstruction error with the threshold on the regression term, and assess the performance of the plausibility term in our objective function.

Optimizations are run using ODE and Java. The projection step lasts approximately one second. After the projection, root position is given instantaneously while computing the physical plausibility takes close to a second. The weight k, used to define the promising region, is set to 1000 and the threshold t, used to apply penalties pertaining to the distance between actual and predicted root position, is 10 cm. Average reconstruction error is near that threshold for all bodies, indicating that our method for computing the optimality of the pose for a given set of forces using physical plausibility provides an accurate model for reconstructing the remaining degrees of freedom.

Durations as well as error measurements for the examples shown in Figure 5.3 are given in Table 5.4. In practice, a sample evaluation with the full objective function is twice as long as a sample evaluation with the hint term, as the extra step of evaluating physical plausibility takes time to set up and stabilize. Without the hint term, the duration of an optimization with the same settings is close to an hour.

Recall that the shape of the hold is not taken into account by our simplistic model. Instead, contacts with the holds are simulated using ball joint constraints. Therefore, and since the orientation of the end effectors does not have an important impact on the weight distribution of the body, hands and feet do not necessarily match well in the provided examples (such as Example 3). Ball joint constraints remain however preferable to fixed constraints, as they provide more freedom in the exploration of the space of possible poses. Another limitation that comes from the simplistic nature of the simulation is the choice of joint angle limits. This is most visible with Example 1, where the left arm is stuck in a

Example	**Time (min)**	ϵ_r **(cm)**
1	24.1	11.19
2	30.0	7.77
3	26.5	8.17
4	26.0	8.88
5	33.1	8.67

Table 5.4: Durations and reconstruction errors for optimization examples.

configuration where the elbow and wrist are locked. Such configurations can be avoided if all the articulations are simulated using ball joint constraints with no joint limits. However, joint limits are crucial in delimiting the set of plausible poses. In practice, they are chosen arbitrarily through trial and error to provide an adequate trade-off between maximizing the set of valid human poses that the skeleton can produce and minimizing the set of unnatural poses that it can reach.

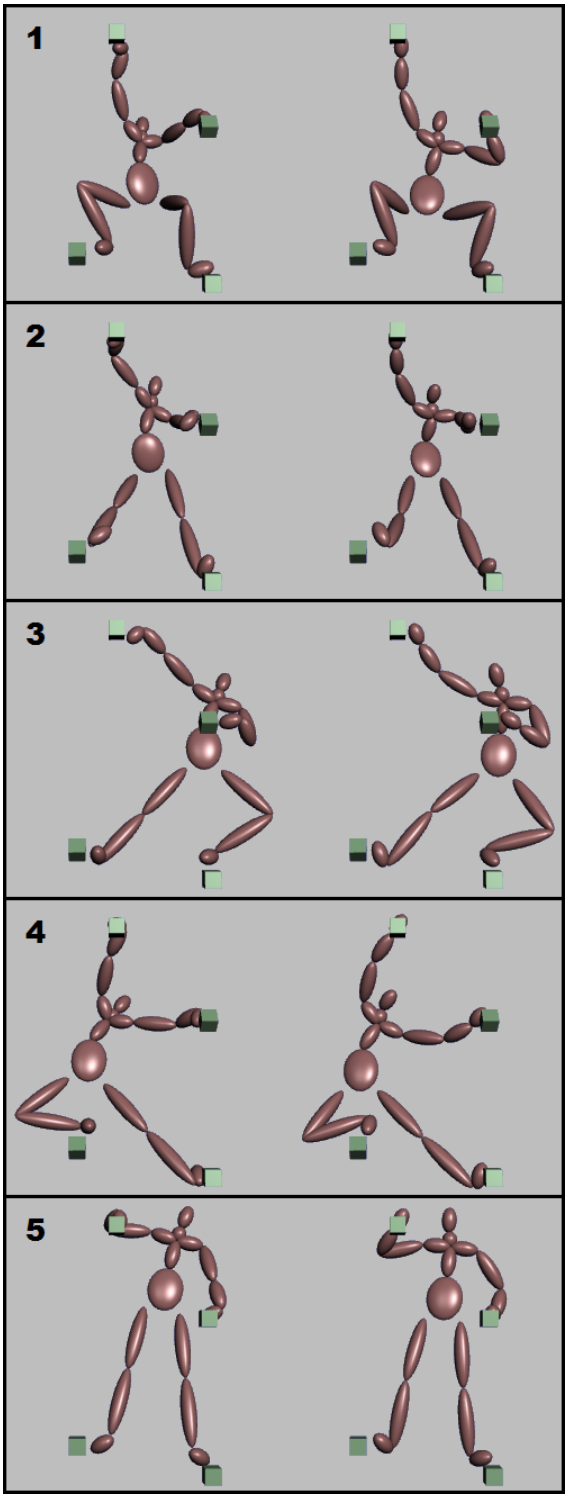


Figure 5.3: Reconstruction (left) and ground truth (right) for varied examples.

Chapter 6

Conclusions

Climbing and bouldering are difficult skills to learn because the actions and postures involved depend not only on body position but also on how forces are applied. Building up a clear understanding of how posture and motion are produced during climbing is useful for learning, as well as to produce physically plausible computer animation of virtual humans.

6.1 Summary and Contributions

In the interest of building up a clear understanding of how posture and motion are produced during climbing, we present the design of an instrumented climbing wall that capture forces at the holds and offer calibration solutions to couple force sensing with a motion capture setup. This innovative setup allows the capture of all contact forces and torques and full 3D postural information, and offers a flexible and portable solution to the study of climbing. Our optimization-based calibration method is also easily generalizable to any arbitrary setup consisting of force sensing and motion capture machinery, and enables the expression of measured forces in the pose capture frame.

In an effort to understand the relationship between posture and contact forces, and because occlusions render motion capture difficult, we propose an objective function to describe

what constitutes a natural pose for a given set of forces. As a first contribution, we focus on the simpler case of static and quasi-static poses. We present new ideas in computing posture from contacts alone, using optimization to find a physically valid pose while meeting other important plausibility constraints. While linear regression cannot directly provide estimates of posture from forces, it is useful in identifying the location of the hips, which we use as an inexpensive hint on human plausibility to compensate for the simplicity of our model. We use this hint to guide our optimization toward a promising subset of poses and compute a measure of the physical plausibility of a pose to choose the optimal solution. Comparison between our reconstructed poses and motion capture shows that our objective function is a good model for human posture.

The three main contributions of this thesis are: an instrumented climbing wall design for the study of climbing, a calibration process for arbitrary systems composed of force sensing and motion capture devices, and a new optimization-based method for estimating climbing posture from contact forces in the static case. Our secondary contributions are: a fast and convenient space calibration in the case of known sensor mounting locations, and a simple synchronization method for systems with asynchronous recording of motion capture and force data.

6.2 Limitations and Future Work

Our calibration and capture procedures, as well as our pose reconstruction method, offer satisfying results. We can identify a number of improvements that could lead to faster, more accurate results, as well as exciting new possibilities based on the current work.

Several improvements could increase calibration precision and speed. Using a calibration tool with a sharper tip, but also calibrating the calibration tool using a smaller motion capture setup would provide better contact point estimates. To improve sensor frame estimates, forces could be sampled at a higher rate and re-sampled using a smoothing kernel to try

and limit the impact of the noise. We also noticed that many of the cameras placed on the ceiling behind the climber offer similar perspectives. Reducing their number will speed up the motion capture system calibration, which is currently the bottleneck of the full capture procedure. Another lengthy process is time synchronization, as we examine each of the data streams manually. The search for the appropriate frames could be automated, but a simpler solution would be to use the hardware to initiate the recordings at the same moment. Finally, while our intuition is that the pocket-based calibration is less accurate than its optimization-based counterpart, it would be interesting to use the error metric given in Equation 5.3 to empirically compare the accuracy of both methods.

While the goal of our pose reconstruction approach is to offer a means of characterizing the relationship between forces and posture, the optimization process is slow. However, the problem of evaluating sample fitness for a CMA-ES iteration is embarrassingly parallel and pose reconstruction could be accelerated tenfold with the current settings by distributing the evaluations. Another concern is that the regression is not likely to generalize well across subjects in the current state. While we could devise a better model that takes into account the mass and size of the climber, the real problem lies in the absence of tangible information on human strength limits as well as how they distribute force among the different muscles. If we had such information, we could determine the human plausibility of a pose using only the inner joint torques and the shapes of the holds.

An important direction for future work is pose reconstruction from forces in the case of dynamic movements. The current approach is not directly applicable to reconstruct poses from forces in the dynamic case, as we do not look at how forces evolve over a period of time. In the dynamic case, the weight of the climber and the forces at the holds might not cancel out since the climber might carry momentum (e.g., in the case of upward momentum, forces at the holds have a far smaller magnitude that the weight of the climber). However, since we use a physical simulation to evaluate the fitness of poses, the current framework can be

expanded to study dynamics. We could use a static sample to reconstruct an initial pose for the climber. Then, optimization could be used to determine the set of torques generated at the climber's joints that best explains measured changes in forces at the holds over small time windows. That set of torques could then be used to update the pose in a coherent way.

Finally, we believe that our methods for estimating posture from forces will be useful for augmented reality applications. For instance, and taking inspiration from Todorov et al. [TJ02], who show that augmented feedback is beneficial for learning a task , data projectors could be used to provide real-time visual feedback directly on the wall as a person is climbing, for learning or rehabilitation.

Appendix A

Space Calibration Optimization Detail

In this appendix, we detail the solving of the least-squares optimizations described in Section 3.2. We describe the solution for the contact point estimation, and then look at the sensor frame and describe the solution for frame origin and frame orientation estimation.

A.1 Contact Point

Recall that in order to find the contact locations, we choose to solve for

$$\underset{{}^{b}p_c, {}^{w}p_c}{\operatorname{argmin}} \sum_{j=1}^{n} \left\| {}^{w}_{b}E_j \; {}^{b}p_c - {}^{w}p_c \right\|^2, \tag{A.1}$$

where ${}^{b}p_c$ and ${}^{w}p_c$ are the position of the contact in the body and world frame, and where

$${}^{w}_{b}E_j = \begin{bmatrix} {}^{w}_{s}R_j & {}^{w}p_{bj} \\ 0 & 1 \end{bmatrix}, \tag{A.2}$$

is the transform that converts coordinates in the body frame to coordinates in the world frame (given by the motion capture software for each recorded frame j). We then set up the problem of finding the estimates for ${}^{b}p_c$ and ${}^{w}p_c$ as a classical least-squares system

$$\underset{x}{\operatorname{argmin}} \left\| A\, x - b \right\|^2, \tag{A.3}$$

where

$$x = \begin{bmatrix} {}^{b}p_c \\ {}^{w}p_c \end{bmatrix}, \tag{A.4}$$

A is a block matrix containing

$$A = \begin{bmatrix} \vdots \\ {}^{w}_{s}R_j - I \\ \vdots \end{bmatrix}, \tag{A.5}$$

and b is a block vector containing

$$b = \begin{bmatrix} \vdots \\ -{}^{w}p_{bj} \\ \vdots \end{bmatrix}. \tag{A.6}$$

For n frames, A is composed of n 3-by-6 blocks and b is composed of n 3-by-1 blocks. In the case of our least squares problems, where there are more constraining equations than there are free variables, the solution to Equation A.3 is found using A^+, the pseudo-inverse [Alb72] of A:

$$x = A^+ \, b \, . \tag{A.7}$$

A.2 Sensor Frame

Recall that in order to estimate the sensor frames, we choose to solve for

$$\underset{{}^{w}_{s}R, {}^{w}p_s}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{j=1}^{m} \| {}^{w}_{s}R \; {}^{s}\tau_{ij} + ({}^{w}p_s - {}^{w}p_{ci}) \times {}^{w}_{s}R \; {}^{s}f_{ij} \|^2, \tag{A.8}$$

where k is the number of contacts, m is the number of recorded force samples per contact, ${}^{s}f$ and ${}^{s}\tau$ are the force and torque in the sensor frame, ${}^{w}p_{ci}$ is the position of each contact in the world, and ${}^{w}_{s}R$ and ${}^{w}p_s$ are the orientation and origin of the sensor frame in world coordinates. We use an iterative and alternating approach: we set ${}^{w}_{s}R$ and ${}^{w}p_s$ to an initial approximation and solve for one of the unknowns while fixing the other. Then, we alternate and repeat until we converge to an acceptable error. Below, we show details for the simpler case of ${}^{w}p_s$, and review the different steps in solving for ${}^{w}_{s}R$.

A.2.1 Origin of the Frame

Using Equation A.8, we set up the problem of finding the estimate for ${}^{w}p_s$ as a classical least-squares system

$$ {}^{w}p_s = A_p^{+}\ b_p\ , \tag{A.9} $$

where

$$ A_p = \begin{bmatrix} & \vdots & \\ 0 & {}^{w}_{s}r_3 \cdot {}^{s}f_{ij} & -\,{}^{w}_{s}r_2 \cdot {}^{s}f_{ij} \\ -\,{}^{w}_{s}r_3 \cdot {}^{s}f_{ij} & 0 & {}^{w}_{s}r_1 \cdot {}^{s}f_{ij} \\ {}^{w}_{s}r_2 \cdot {}^{s}f_{ij} & -\,{}^{w}_{s}r_1 \cdot {}^{s}f_{ij} & 0 \\ & \vdots & \end{bmatrix}, \tag{A.10} $$

$$ b_p = \begin{bmatrix} \vdots \\ -\,{}^{w}_{s}r_1 \cdot {}^{s}\tau_{ij} - {}^{w}p_{ciz}\,({}^{w}_{s}r_2 \cdot {}^{s}f_{ij}) + {}^{w}p_{ciy}\,({}^{w}_{s}r_3 \cdot {}^{s}f_{ij}) \\ -\,{}^{w}_{s}r_2 \cdot {}^{s}\tau_{ij} + {}^{w}p_{ciz}\,({}^{w}_{s}r_1 \cdot {}^{s}f_{ij}) - {}^{w}p_{cix}\,({}^{w}_{s}r_3 \cdot {}^{s}f_{ij}) \\ -\,{}^{w}_{s}r_3 \cdot {}^{s}\tau_{ij} - {}^{w}p_{ciy}\,({}^{w}_{s}r_1 \cdot {}^{s}f_{ij}) + {}^{w}p_{cix}\,({}^{w}_{s}r_2 \cdot {}^{s}f_{ij}) \\ \vdots \end{bmatrix}, \tag{A.11} $$

and where ${}^{w}_{s}r_1$, ${}^{w}_{s}r_2$ and ${}^{w}_{s}r_3$ denote respectively the first, second and third rows of the rotation matrix ${}^{w}_{s}R$, ${}^{w}p_{cix}$, ${}^{w}p_{ciy}$ and ${}^{w}p_{ciz}$ are respectively the x, y and z coordinate of the ith contact point ${}^{w}p_{ci}$, and ${}^{s}f_{ij}$ and ${}^{s}\tau_{ij}$ are the jth force and torque measurements for that i^{th} contact point. For k contact points and m force measurements per contact point, A_p and b_p are respectively composed of km 3-by-3 blocks and km 3-by-1 blocks.

A.2.2 Orientation of the Frame

Recall that when solving for ${}^{w}_{s}R$ with a fixed ${}^{w}p_s$, we choose to solve for

$$ \underset{\omega}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{j=1}^{m} \| {}^{w}_{s}R_l (I + [\omega])\ {}^{s}\tau_{ij} + ({}^{w}p_s - {}^{w}p_{ci}) \times {}^{w}_{s}R_l (I + [\omega])\ {}^{s}f_{ij} \|^2 , \tag{A.12} $$

where ${}^{w}_{s}R_l$ is the most recent estimate for ${}^{w}_{s}R$. We set up the problem of finding the estimate for ω as a classical least-squares system

$$ \omega = A_r^{+}\ b_r\ , \tag{A.13} $$

where

$$A_r = \begin{bmatrix} & \vdots & \\ A_{ij11} & A_{ij12} & A_{ij13} \\ A_{ij21} & A_{ij22} & A_{ij23} \\ A_{ij31} & A_{ij32} & A_{ij33} \\ & \vdots & \end{bmatrix}, \tag{A.14}$$

$$b_r = \begin{bmatrix} & \vdots & \\ -\,{}^{w}_{s}R_l\ {}^{s}\tau_{ij} - {}^{c}p_{si} \times {}^{w}_{s}R_l\ {}^{s}f_{ij} \\ & \vdots & \end{bmatrix}, \tag{A.15}$$

and where

$$\begin{aligned} A_{ij11} &= -\,{}^{w}_{s}r_{12}\ \tau_{ijz} + {}^{w}_{s}r_{13}\ \tau_{ijy} \\ &\quad -(-\,{}^{c}p_{siz}\ {}^{w}_{s}r_{22} + {}^{c}p_{siy}\ {}^{w}_{s}r_{32})\ {}^{s}f_{ijz} \\ &\quad +(-\,{}^{c}p_{siz}\ {}^{w}_{s}r_{23} + {}^{c}p_{siy}\ {}^{w}_{s}r_{33})\ {}^{s}f_{ijy}\ , \end{aligned} \tag{A.16}$$

$$\begin{aligned} A_{ij12} &= {}^{w}_{s}r_{11}\ \tau_{ijz} - {}^{w}_{s}r_{13}\ \tau_{ijx} \\ &\quad +(-\,{}^{c}p_{siz}\ {}^{w}_{s}r_{21} + {}^{c}p_{siy}\ {}^{w}_{s}r_{31})\ {}^{s}f_{ijz} \\ &\quad -(-\,{}^{c}p_{siz}\ {}^{w}_{s}r_{23} + {}^{c}p_{siy}\ {}^{w}_{s}r_{33})\ {}^{s}f_{ijx}\ , \end{aligned} \tag{A.17}$$

$$\begin{aligned} A_{ij13} &= -\,{}^{w}_{s}r_{11}\ \tau_{ijy} + {}^{w}_{s}r_{12}\ \tau_{ijx} \\ &\quad -(-\,{}^{c}p_{siz}\ {}^{w}_{s}r_{21} + {}^{c}p_{siy}\ {}^{w}_{s}r_{31})\ {}^{s}f_{ijy} \\ &\quad +(-\,{}^{c}p_{siz}\ {}^{w}_{s}r_{22} + {}^{c}p_{siy}\ {}^{w}_{s}r_{32})\ {}^{s}f_{ijx}\ , \end{aligned} \tag{A.18}$$

$$\begin{aligned} A_{ij21} &= -\,{}^{w}_{s}r_{22}\ \tau_{ijz} + {}^{w}_{s}r_{23}\ \tau_{ijy} \\ &\quad -({}^{c}p_{siz}\ {}^{w}_{s}r_{12} - {}^{c}p_{six}\ {}^{w}_{s}r_{32})\ {}^{s}f_{ijz} \\ &\quad +({}^{c}p_{siz}\ {}^{w}_{s}r_{13} - {}^{c}p_{six}\ {}^{w}_{s}r_{33})\ {}^{s}f_{ijy}\ , \end{aligned} \tag{A.19}$$

$$
\begin{aligned}
A_{ij22} &= {}^{w}_{s}r_{21}\ \tau_{ijz} - {}^{w}_{s}r_{23}\ \tau_{ijx} \\
&\quad +({}^{c}p_{siz}\ {}^{w}_{s}r_{11} - {}^{c}p_{six}\ {}^{w}_{s}r_{31})\ {}^{s}f_{ijz} \\
&\quad -({}^{c}p_{siz}\ {}^{w}_{s}r_{13} - {}^{c}p_{six}\ {}^{w}_{s}r_{33})\ {}^{s}f_{ijx}\ ,
\end{aligned} \tag{A.20}
$$

$$
\begin{aligned}
A_{ij23} &= -\ {}^{w}_{s}r_{21}\ \tau_{ijy} + {}^{w}_{s}r_{22}\ \tau_{ijx} \\
&\quad +({}^{c}p_{siz}\ {}^{w}_{s}r_{12} - {}^{c}p_{six}\ {}^{w}_{s}r_{32})\ {}^{s}f_{ijx} \\
&\quad -({}^{c}p_{siz}\ {}^{w}_{s}r_{11} - {}^{c}p_{six}\ {}^{w}_{s}r_{31})\ {}^{s}f_{ijy}\ ,
\end{aligned} \tag{A.21}
$$

$$
\begin{aligned}
A_{ij31} &= -\ {}^{w}_{s}r_{32}\ \tau_{ijz} + {}^{w}_{s}r_{33}\ \tau_{ijy} \\
&\quad -(-\ {}^{c}p_{siy}\ {}^{w}_{s}r_{12} + {}^{c}p_{six}\ {}^{w}_{s}r_{22})\ {}^{s}f_{ijz} \\
&\quad +(-\ {}^{c}p_{siy}\ {}^{w}_{s}r_{13} + {}^{c}p_{six}\ {}^{w}_{s}r_{23})\ {}^{s}f_{ijy}\ ,
\end{aligned} \tag{A.22}
$$

$$
\begin{aligned}
A_{ij32} &= {}^{w}_{s}r_{31}\ \tau_{ijz} - {}^{w}_{s}r_{33}\ \tau_{ijx} \\
&\quad +(-\ {}^{c}p_{siy}\ {}^{w}_{s}r_{11} + {}^{c}p_{six}\ {}^{w}_{s}r_{21})\ {}^{s}f_{ijz} \\
&\quad -(-\ {}^{c}p_{siy}\ {}^{w}_{s}r_{13} + {}^{c}p_{six}\ {}^{w}_{s}r_{23})\ {}^{s}f_{ijx}\ ,
\end{aligned} \tag{A.23}
$$

$$
\begin{aligned}
A_{ij33} &= -\ {}^{w}_{s}r_{31}\ \tau_{ijy} + {}^{w}_{s}r_{32}\ \tau_{ijx} \\
&\quad -(-\ {}^{c}p_{siy}\ {}^{w}_{s}r_{11} + {}^{c}p_{six}\ {}^{w}_{s}r_{21})\ {}^{s}f_{ijy} \\
&\quad +(-\ {}^{c}p_{siy}\ {}^{w}_{s}r_{12} + {}^{c}p_{six}\ {}^{w}_{s}r_{22})\ {}^{s}f_{ijx}\ ,
\end{aligned} \tag{A.24}
$$

where the second subscript index after rotation matrix row number designates the matrix column number, and where, for simplicity, we write ${}^{c}p_{si}$ to designate ${}^{w}p_{s} - {}^{w}p_{ci}$ (which is in fact the origin of the sensor frame expressed in the coordinate frame of contact i). Once again, for k contact points and m force measurements per contact point, A_r and b_r are respectively composed of km 3-by-3 blocks and km 3-by-1 blocks.

Bibliography

[Alb72] A. E. Albert. *Regression and the Moore-Penrose Pseudoinverse*. Academic Press Inc., 1972.

[AMT06] AMTI. *MSA-6 MiniAmp: Strain Gage Amplifier Instruction Manual*, 2006.

[BA01] G. Baudat and F. Anouar. Kernel-Based Methods and Function Approximations. In *International Joint Conference on Neural Networks*, 2001.

[BMP04] C. Bregler, J. Malik, and K. Pullen. Twist Based Acquisition and Tracking of Animal and Human Kinematics. *International Journal of Computer Vision*, 56(3):179–194, 2004.

[BSF09] M. A. Brubaker, L. Sigal, and D. J. Fleet. Estimating Contact Dynamics. In *International Conference on Computer Vision*, pages 2389–2396, 2009.

[bvh] BVH. `http://www.mindfiresolutions.com/BVH-biovision-hierarchy`.

[CDL+09] M. T. Ciocarlie, H. Dang, J. Lukos, M. Santello, and P. K. Allen. Functional Analysis of Finger Contact Locations During Grasping. In *WHC*, pages 401–405, 2009.

[cma] CMA-ES. `http://en.wikipedia.org/wiki/CMA-ES`.

[dLMH10] M. de Lasa, I. Mordatch, and A. Hertzmann. Feature-Based Locomotion Controllers. In *ACM SIGGRAPH*, 2010.

[FJLP04] P.T. Fletcher, S. Joshi, C. Lu, and S.M. Pizer. Principal Geodesic Analysis for the Study of Nonlinear Statistics of Shape. *IEEE Transactions on Medical Imaging*, 23(8):995 – 1005, 2004.

[FN06] F. K. Fuss and G. Niegl. Instrumented Climbing Holds and Dynamics of Sport Climbing. In *The Engineering of Sport 6*, pages 57–62. Springer New York, 2006.

[FN07] F. K. Fuss and G. Niegl. *The Impact of Technology on Sport II*, chapter The Fully Instrumented Climbing Wall: Performance Analysis, Route Grading and Vector Diagrams – a Preliminary Study, pages 677–682. Taylor and Francis, 2007.

[FN08] F. K. Fuss and G. Niegl. Instrumented Climbing Holds and Performance Analysis in Sport Climbing. *Sports Technology*, 1(6):301–313, 2008.

[Han06] N. Hansen. The CMA Evolution Strategy: A Comparing Review. In *Towards a New Evolutionary Computation*, volume 192, pages 75–102. Springer Berlin / Heidelberg, 2006.

[HBL11] S. Ha, Y. Bai, and K. C. Liu. Human Motion Reconstruction from Force Sensors. In *Proceedings of ACM SIGGRAPH/Eurographics SCA*, pages 129–138, 2011.

[HRT+09] N. Hasler, B. Rosenhahn, T. Thormählen, M. Wand, J. Gall, and H.-P. Seidel. Markerless Motion Capture with Unsynchronized Moving Cameras. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2009.

[Jol02] I. T. Jolliffe. *Principal Component Analysis, Second Edition*. Springer, 2002.

[JYL09] S. Jain, Y. Ye, and K. C. Liu. Optimization-Based Interactive Motion Synthesis. *ACM Transactions on Graphics*, 28(1):10:1–10:12, 2009.

[KB96] E. Kreighbaum and K. M. Barthels. *Biomechanics: A Qualitative Approach for Studying Human Movement*. Allyn and Bacon, 1996.

[KP06a] P. G. Kry and D. K. Pai. Grasp Recognition and Manipulation with the Tango. In *International Symposium on Experimental Robotics*, volume 10. Springer, 2006.

[KP06b] P. G. Kry and D. K. Pai. Interaction Capture and Synthesis. In *ACM SIGGRAPH*, pages 872–880, 2006.

[LCB91] B. L. Luk, A. A. Collie, and J. Billingsley. Robug II: An Intelligent Wall Climbing Robot. In *International Conference on Robotics and Automation*, 1991.

[MA99] A. T. Miller and P. K. Allen. Examples of 3D Grasp Quality Computations. In *1999 IEEE International Conference on Robotics and Automation*, pages 1240–1246, 1999.

[NCNV+12] R. F. Nunes, J. B. Cavalcante-Neto, C. A. Vidal, P. G. Kry, and V. B. Zordan. Using Natural Vibrations to Guide Control for Locomotion. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pages 87–94, 2012.

[QM99] F. Quaine and L. Martin. A Biomechanical Study of Equilibrium in Sport Rock Climbing. *Gait and Posture*, 10(3):233–239, 1999.

[QMB97a] F. Quaine, L. Martin, and J.-P. Blanchi. Effect of a Leg Movement on the Organisation of the Forces at the Holds in a Climbing Position: 3-D Kinetic Analysis. *Human Movement Science*, 16(2–3):337–346, 1997.

[QMB97b] F. Quaine, L. Martin, and J.-P. Blanchi. The Effect of Body Position and Number of Supports on Wall Reaction Forces in Rock Climbing. *Journal of Applied Biomechanics*, 13:14 – 23, 1997.

[RB91] P. Rougier and J.-P. Blanchi. Evaluation Objective de la Difficulté en Escalade Par la Relation Posturo-Cinétique. *Sciences Techniques pour la Recherche en Activités Physiques et Sportives*, 26:61 – 77, 1991.

[RB92] P. Rougier and J.-P. Blanchi. Mesure de la Force Maximale Volontaire à partir d'une Posture Quadrupodale en Escalade : Influence du Niveau d'Expertise. *Science and Sports*, 7(1):19–25, 1992.

[RBMB91] P. Rougier, R. Billat, R. Merlin, and J.-P. Blanchi. Conception d'un Système pour Étudier la Relation Posturo-Cinétique dans un Plan Vertical : Application sur une Population de Grimpeurs. *Innovation et Technologie en Biologie et Médecine*, 12:568 – 580, 1991.

[RSB+08] B. Rosenhahn, C. Schmaltz, T. Brox, J. Weickert, D. Cremers, and H.-P. Seidel. Markerless Motion Capture of Man-Machine Interaction. In *Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[RVC10] R. Rajalingham, Y. Visell, and J. R. Cooperstock. Probabilistic Tracking of Pedestrian Movements via In-Floor Force Sensing. In *Computer and Robot Vision*, pages 143 – 150, 2010.

[SHS+08] M. A. Spenko, G. C. Haynes, A. Saunders, A. A. Rizzi, M. Cutkosky, R. J. Full, and D. E. Koditschek. Biologically Inspired Climbing with a Hexapedal Robot. *Journal of Field Robotics*, 25(4 – 5):223 – 242, 2008.

[SL12] G. A. F. Seber and A. J. Lee. *Linear Regression Analysis, Second Edition*. John Wiley and Sons, 2012.

[Smi] Russell Smith. Open Dynamics Engine. `http://www.ode.org`.

[Tho00] B. Thompson. *Reading and Understanding More Multivariate Statistics*, volume 13, chapter Canonical Correlation Analysis, pages 285 – 316. American Psychological Association, 2000.

[TJ02] E. Todorov and M. I. Jordan. Optimal Feedback Control as a Theory of Motor Coordination. *Natural Neuroscience*, 5(11):1226–1235, 2002.

[TMD99] M. Testa, L. Martin, and B. Debû. Effects of the Type of Holds and Movement Amplitude on Postural Control associated with a Climbing Task. *Gait and Posture*, 9(1):57–64, 1999.

[TMD03] M. Testa, L. Martin, and B. Debû. 3D Analysis of Posturo-Kinetic Coordination associated with a Climbing Task in Children and Teenagers. *Neuroscience Letters*, 336(1):45–49, 2003.

[TWC+09] M. Tournier, X. Wu, N. Courty, E. Arnaud, and L. Revéret. Motion Compression using Principal Geodesics Analysis. *Computer Graphics Forum*, 28(2):355–364, 2009.

[YP03] K. Yin and D. K. Pai. FootSee: An Interactive Animation System. In *Proceedings of ACM SIGGRAPH/Eurographics SCA*, pages 329–338, 2003.