Static Pose Reconstruction with an Instrumented Bouldering Wall

Rami Aladdin
Computer Graphics and Animation Lab, School
of Computer Science, McGill University
rami.aladdin@cs.mcgill.ca

Paul G. Kry
Computer Graphics and Animation Lab, School
of Computer Science, McGill University
kry@cs.mcgill.ca

ABSTRACT

This paper describes the design and construction of an instrumented bouldering wall, and a technique for estimating poses by optimizing an objective function involving contact forces. We describe the design and calibration of the wall, which can capture the contact forces and torques during climbing while motion capture (MoCap) records the climber pose, and present a solution for identifying static poses for a given set of holds and forces. We show results of our calibration process and static poses estimated for different measured forces. To estimate poses from forces, we use optimization and start with an inexpensive objective to guide the solver toward the optimal solution. When good candidates are encountered, the full objective function is evaluated with a physics-based simulation to determine physical plausibility while meeting additional constraints. Comparison between our reconstructed poses and MoCap show that our objective function is a good model for human posture.

Categories and Subject Descriptors

I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism—*Virtual reality*

Keywords

Motion Capture; Force Capture; Physics-Based Simulation; Optimization

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 center of mass (COM) to a location above the feet. Another way to generate plausible postures for virtual humans is to simply use MoCap and record the poses of a real person. MoCap can be challenging to modify, however, and we must ultimately come up with a computational model that lets us edit the motion for new purposes.

In this paper, we build on the idea of bringing together physics, optimization, and MoCap, which is a popular approach in research on virtual humans. However, we also recognize that MoCap alone is only part of the picture, and that contact forces are critical for building up a clear understanding of how posture and motion are produced: contact forces let us resolve the ambiguity in determining the torques applied at different joints. Contacts and contact forces are often critical, whether in the context of standing balance, object manipulation, or 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 have designed an instrumented climbing wall that allows forces and torques to be measured at the holds. In this investigation, we focus on static and slow moving, near-static postures in order to simplify the problem. 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. We use optimization to find a pose that meets the contact constraints while being valid for the measured forces and minimizing 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. We use Covariance Matrix Adaptation Evolution Strategy (CMA-ES) to compute the solution as it alleviates the need for a derivative in the objective function. 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 neighborhood of the optimal solution quickly, and reduces the number of computationally expensive evaluations of our full objective function. Our posture optimization method is beneficial because MoCap of climbing can be challenging due to occlusions, and we can instead produce posture estimates for a climb without requiring the capture subject to wear a MoCap suit.

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. Our climbing wall design permits forces and torques to be measured at the holds, and allows a variety of different experiments to be conducted. The pitch of the wall is adjustable, and hold positions are easily reconfigurable in pockets of the torsion box. Also, the wall is easily disassembled for transport or long term storage. The calibration process for force sensors mounted in arbitrary locations is presented, in addition to a convenient calibration based on known mounting locations. Finally, our estimates of posture from recorded forces combine physics, capture, and optimization to produce plausible poses which we validate with ground truth MoCap measurements.

2. RELATED WORK

We first review related work in biomechanics involving climbing and the use of instrumented climbing holds. Then, we focus on the field of animation, on works that combine motion and forces, and on works revolving around the use of optimization for physics-based simulations.

Climbing has been studied by various research groups in the field of biomechanics using instrumented holds. Rougier et al. [15] describe the first instrumented wall designed to measure the amplitude of forces at the holds in 1991. Testa et al. [17] use 3-Degrees of Freedom (3DoF) sensors. Quaine et al. [12] use two video cameras to record two 2-D 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. [4], where eight holds are equipped with 6DoF sensors. Works using these setups mainly focus on analyzing the performance of a climb and the climbers' level of expertise [4, 3, 15, 16], and on studying the effects of specific posture changes or constraints on forces at the holds [13, 12, 17]. Our setup is different in that it incorporates MoCap 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 are looking at the amount of information contained in forces to explain a pose.

Pfeil et al. [11] provide a tool for quickly designing and testing different hold configurations on a virtual climbing wall. While the tool provides a good means of visualizing possible routes, our approach differs as it relies on a physically simulated model of the climber.

A wide range of contributions to capture and synthesis of posture or force data is found in computer animation. Similarly to Brubaker et al. [1], who look at the converse problem of estimating forces acting on the system from given motion, we use a simulated character to model the subject and constraints. Rosenhahn et al. [14] also use simulation to enforce constraints found through analytics in MoCap of special interactions.

MoCap and force sensing is used by Kry et al. [9] 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 to look at the relationship between static poses and forces at the holds.

Closer to our work, Yin et al. [18] and Kry et al. [8] look at how full body poses and grasp configurations can be estimated using force measurements from respectively a foot-ground pressure sensor pad and a graspable device. These approaches use a database consisting of pairs of MoCap and force data to find the pose for the closest set of measured forces. We, on the other hand, learn reduced features from recorded data and use a physics-based approach to find a pose that best explains the forces.

A similar approach is used by Ha et al. [5] to reconstruct human motion from consumer-grade force sensors, and while they also use a form of MoCap to obtain hand positions, one can argue that because of the constrained nature of climbing, end effector positions are given by the holds whereas only feet locations can be obtained in the case of standing on a platform. A fundamental difference is that our work focuses on statics and looks at 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.

Another body of relevant research in animation concerns the use of optimization for physics-based simulations. de Lasa et al. [2] describe a means of expressing locomotion using a small number of features. Jain et al. [7] also look at constrained optimization for virtual characters to design controllers by formulating high-level objectives, and use those constraints to generate, among other controllers, a climbing controller. Using CMA-ES [6] with simple penalties around the main objective is also a method we share with Nunes et al. [10]. However, these authors look at generating plausible motion through the optimization of selected objectives, while we are interested in optimizing the physical quality of a static pose reconstruction for a set of forces.

3. WALL DESIGN

The instrumented bouldering wall's main structure is an 8'×8' torsion box. The wall doesn't need to be fixed to a second supporting 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, as shown in Figure 1.

Each climbing hold is mounted onto a front plate, 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" shown in Figure 3a. The torsion box design will ultimately allow us to hide sensors and wires inside and behind the pockets, leaving sensitive material out of harm's way and giving the climber the same freedom they would have on a regular wall.

The sensor sandwiches can be inserted into each pocket in four different orientations, providing good freedom 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, or unscrewing the hold from the sensor.

The sensors are multi-component transducers with 6 channels for linear and angular forces expressed in a local coordinate frame. Each sensor is connected to a 6-channel strain gage amplifier with adjustable excitation voltage and gain. 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. The sensors are
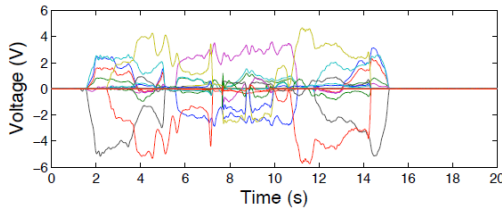
Figure 1: The instrumented bouldering wall.

Figure 2: Voltages recorded for an example session.

sensitive enough to capture small tremors of muscle fatigue, as well as contact forces as light as 0.05 N.

The 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 2.

4. CALIBRATION AND CAPTURE

The main challenge in capturing data from both the Mo-Cap and force sensing equipment is to find the relationship between the two independently recorded data sets, both in time and space. In this section, we describe the calibration process, present the major steps involved in capturing synchronized data, and show a sample visualization.

4.1 Calibration

Each sensor possesses its own local frame in which the force measurements are expressed as voltages, while skeleton reconstruction is expressed in the MoCap world frame. Below, we describe two techniques for calibrating the position and orientation of the sensors, which allow the force and torque measurements to be expressed in the MoCap world reference frame. We also describe a simple technique for synchronizing the two data sets in time.

4.1.1 Converting the raw data to forces

As shown in Figure 2, sensors provide voltages for each of the 6 channels. Using the inverse sensitivity matrices that are provided with each sensor, we compute the forces and torques from these voltages while taking into consideration possible cross-talk. The forces and torques are given by

$$F = SV, \qquad (1)$$

where F is the vector containing the forces and torques, S the inverse cross-talk sensitivity matrix for the sensor and

V the vector containing normalized voltages for the corresponding channels. Normalized voltages for a channel are found by dividing measured values by the appropriate gain and excitation voltage, which are configured on the amplifier boards for each individual channel, for each sensor.

4.1.2 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 transform

$$_s^w E = \begin{bmatrix} _s^w R & {}^w p_s \\ \mathbf{0} & 1 \end{bmatrix} \qquad (2)$$

that converts coordinates expressed in the local frame of the sensor to coordinates in the world MoCap frame. Here s denotes the sensor frame, w the world frame, ${}^w p_s$ the origin of the sensor frame in world coordinates, and $_s^w R$ is the rotation matrix that aligns the axis in the sensor frame with the corresponding axis in the world frame. In order to find ${}^w p_s$ and $_s^w R$, we look at the influences of force and contact location on sensor measurements. We will show how precise contact locations are found, and how these positions are used along with recorded forces to compute the sensor frames.

To obtain contact locations, a calibration tool consisting of a rigid body with a sharp tip is tracked with MoCap as it is rotated around the contact point, as shown in Figure 3b. For a recorded set of orientations of our rigid body, the contact point is found using a least-squares approach by finding the pair $({}^b p_c, {}^w p_c)$ that minimizes

$$\sum_{j=1}^{n} \left\| {}_b^w E_j \; {}^b p_c \; - \; {}^w p_c \right\|^2, \qquad (3)$$

where b is the body frame, n the number of collected samples, ${}^w p_c$ is the position of the contact point in world coordinates, and $_b^w E_j$ is the calibration tool coordinate frame for sample j in the recorded MoCap trajectory. In other words, we are looking for the point p_c that is invariant for $_b^w E$ over the data. That is the point that best describes a center of rotation for all individual frames, which is the location of the contact between the tip and surface. Using the position of the tip in body coordinates, we can thereafter find any contact location in world coordinates from the captured calibration tool frame.

We can now look at how a set of linear forces (with no torque) applied at different contact points are measured by the sensor. More specifically, and assuming that the contact coordinate frame is aligned with the world coordinate frame, the measured torque can be expressed as

$$^c\tau = {}_s^w R \; {}^s\tau + ({}^w p_s - {}^w p_c) \times {}_s^w R \; {}^s f , \qquad (4)$$

where $^c\tau$ is the torque in the contact frame, and $^s f$ and $^s\tau$ are respectively the forces and torques in the sensor frame. 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 measured forces becomes an optimization problem where we are trying to find the pair $({}_s^w R, {}^w p_s)$ that minimizes

$$\sum_{i=1}^{k}\sum_{j=1}^{m} \left\| {}_s^w R \; {}^s\tau_{ij} + ({}^w p_s - {}^w p_{ci}) \times {}_s^w R \; {}^s f_{ij} \right\|^2, \qquad (5)$$
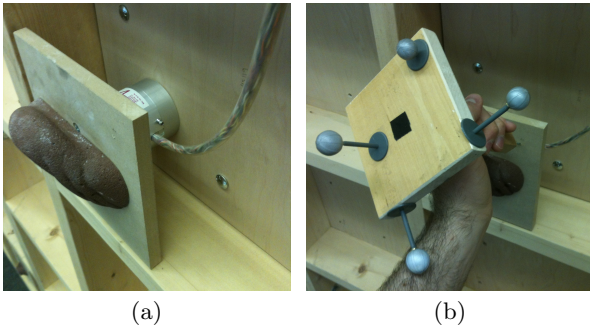
(a)   (b)

Figure 3: Close-up on a "sensor sandwich" (a) and the calibration tool in action (b).

where k is the number of contacts and m is the number of recorded force samples per contact. Put simply, we are looking for the sensor frame that can best explain force and torque measurements for every given contact point. We find $^w p_c$ using the calibration tool and MoCap, and $^s \tau$ and $^s f$ are given by the sensors. In order to perform this calibration, we then need a wide set of forces applied at various contact points to ensure that the problem is not under-constrained.

We use MoCap to track the calibration tool while applying linear forces at a contact point, with the sensors recording forces and torques. For a set of contact points, and the forces and torques recorded during the contact, we use Equation 5 to solve for $^w_s R$ and $^w p_s$. The equation is nonlinear due to rotations, and once linearized (as we describe later), it is quartic. For that reason, we use an iterative and alternating least squares approach. We set $^w_s R$ and $^w p_s$ to initial approximations and then solve for one of the unknowns while fixing the other. Then, we alternate and repeat until we converge to an acceptable error. In practice, three different contact points are enough to find good solutions.

With $^w_s R$ fixed, $^w p_s$ can be easily computed as the least squares solution of Equation 5. Solving for the best rotation is trickier. We choose to 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

$$^w_s R_{l+1} = {}^w_s R_l \ R \ . \tag{6}$$

We parameterize R in exponential coordinates, $R = \exp([\omega])$, where $[\omega]$ is the skew symmetric matrix that performs the cross product operation $\omega \times$, and we linearize the rotation with a first order approximation of the exponential,

$$^w_s R_{l+1} \approx {}^w_s R_l \ (I + [\omega]). \tag{7}$$

When solving for $^w_s R$ with a fixed $^w p_s$, we consequently solve for ω that minimizes Equation 5, where $^w_s R$ becomes $^w_s R_l (I + [\omega])$ and $^w_s R_l$ is the most recent estimate. With the least squares solution ω, we use Equation 6 to update the estimate for $^w_s R$ using the exponential map $R = \exp([\omega])$ computed with Rodrigues' formula. To limit the error induced by the linearization, we bound the norm of ω so that the update is no more than 0.1 radians.

In sum, the calibration process consists of calibrating the calibration tool to have accurate contact positions, recording motion and forces as the tool is used to apply linear forces at different contact locations rigidly attached to the sensor and computing a least-squares estimate for the frame using an iterative method and first-order approximation.

4.1.3   Pocket-based space calibration

Although time consuming as it requires several MoCap and force recordings for each sensor to be calibrated, the last approach is accurate and indispensable when sensors are not in specific locations. In cases where sensors are attached in pockets in known orientations, a speedier calibration method is to place markers on known locations on the wall and define a wall coordinate frame in which a sensor position and orientation can be expressed. The origin of each sensor frame can be computed from its location in a given corner of the pocket where it is mounted, while the orientation can be set to a combination of 90-degree rotations about the axis of the wall frame. The positions and orientations are then easily expressed in the world frame using the position of the wall, determined by the markers attached to it and MoCap.

4.1.4   Measurement synchronization

Since our equipment does not easily permit a synchronous recording of motion and forces, 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 with an object tracked with MoCap. 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 MoCap data. For simplicity, we capture forces and motion at the same rate, even though forces can be sampled at a much higher frequency.

4.2   Capture

As in any system involving MoCap, one of the main challenges when it comes to collecting quality data is camera placement to avoid occlusion. 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. Sometimes marker positions are only chosen after several trials, especially when recording new dynamic motions. Our MoCap setup consists of 24 cameras, 6 of which have wide-angle lenses; these 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.

For the force sensors, the only concern is to choose gain and excitation parameters that will allow for high sensitivity while preventing sensor overload. This is also done by trial and error but is far less time consuming. Typically, more torque is exerted by hands as the holds need to allow for grasping and tend to protrude more, as shown in Figure 3a.

The pre-processing to perform a successful capture session consists of calibrating the cameras to avoid occlusions, finding the origins and orientations of the sensor frames using one of the two space calibration methods, and carefully choosing marker positions and sensor sensitivities to have quality data. During the capture session, the subject uses a rigid object to hit sensors prior to climbing and performs the desired task. The post-processing consists of using the hit to synchronize the two data streams in time, using the calibration matrices to convert the recorded voltages to force measurements, and using the estimated sensor frames to express the measured forces in the MoCap world frame. Figure 4 depicts the results of a successful capture, where squares represent the sensors, and red and green arrows represent forces and torques measured by each sensor.
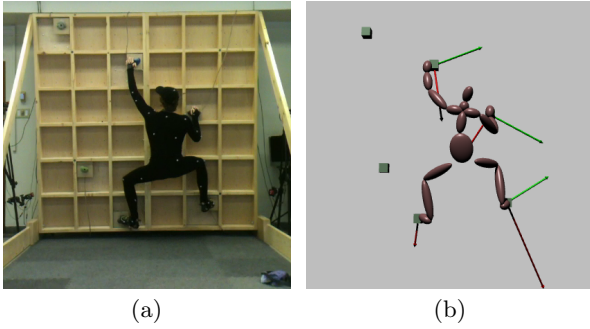
(a)   (b)

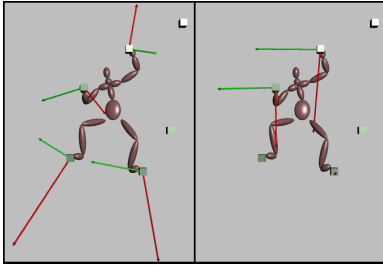Figure 4: Climber during a trial (a) and visualization (b).

Figure 5: Subject generating a range of arbitrary forces at the sensors while maintaining the same pose.

5. RECONSTRUCTION OF STATIC POSES

In order to look at the relationship between motion and forces, several capture trials with male and female participants of various skill levels were recorded. 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 5, and similar forces can be generated by different poses. To limit ambiguity, we restrict our focus to simpler problems where data consists of static poses and quasi-static transitions. To ensure as little dynamic movements as possible, participants were asked to move slowly while performing natural and controlled movements.

5.1 Statistical analysis

We took a preliminary look at the data to see how much of the posture and force spaces was exploited by various climbers. We investigated the correspondence between forces and static poses by performing linear regression and canonical correlation analysis. In static or quasi-static poses, the forces offer a lot of information on the location of the COM: without momentum, forces at the holds are highly dependent on the distribution of the climber's weight, as shown in Figure 6. For simplicity, and because we do not know the exact mass of the different links, the position of the COM is approximated by the root of the skeleton (the hips). The location of the root can be predicted using linear regression with a maximum error around 10 cm. However, the models failed to predict entire poses from forces. A likely explanation is strong non-linearities in the correlation due to joint angles. A principal component analysis 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
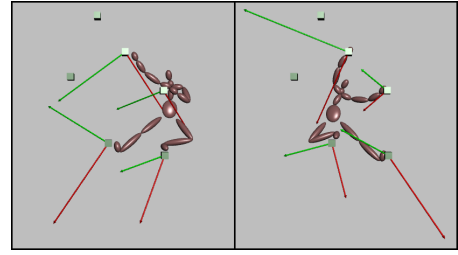

Figure 6: Correlation between forces at the sensors and COM position for static poses.
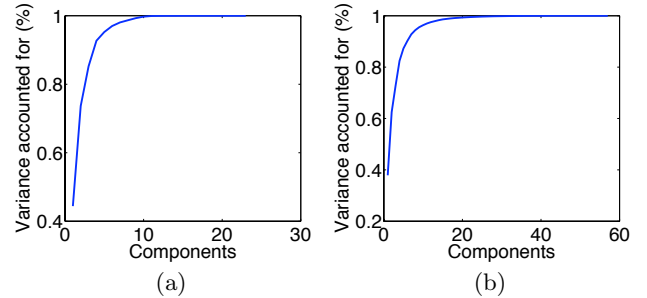

(a)   (b)

Figure 7: Variance explained against number of components for (a) forces and torques and (b) MoCap joint angles.

Figure 7. The lower dimensionality of both data sets can be explained by the highly constrained nature of the problem.

5.2 Physical simulation and CMA-ES

In order to take advantage of these constraints and to account for physical plausibility, we use a physics-based simulation. 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, and has joint angle limits. The virtual character used to simulate the climber has its dimensions defined by the captured climber (stored in the Biovision Hierarchy format). The skeleton is a set of rigid bodies attached using different types of joints. Joints are also used to constrain hands and feet to proper hold locations, determined by assuming the climber is facing the wall, is upright and is exerting force on the holds. Joint angle limits and weight distribution are chosen to match that of the climber, while contacts with the simulated wall prevent unnatural poses. The simulation is built using the Open Dynamics Engine (ODE).

We want to find a pose that best explains an arbitrary function – which we define in Sections 5.3 and 5.4 – of forces at the holds. Treating this as an optimization, we want the pose that minimizes our objective while taking into account the hard constraints enforced by the simulation. In other words, we are looking for the set of joint angles

$$\underset{\phi' \in \Phi^*}{\operatorname{argmin}} f(\phi', F_S), \qquad (8)$$

where Φ^* is the manifold of joint angles that describe poses respecting the constraints, f is the objective function, and F_S is the measured set of forces. We choose CMA-ES, which uses sample fitness to update the sampling distribution, because we evaluate our objective functions via simulation.

CMA-ES produces sets of arbitrary joint angles ϕ using a multi-variate normal distribution, and we devise a projection step $\Gamma(\phi)$ 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 (from MoCap data) and treating the produced angles ϕ as desired angles. 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 angles ϕ onto the constraint manifold. The fitness of the original sample ϕ is computed using the projected pose given by ϕ'.

5.3 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 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 sensors. The plausibility is computed as

$$\rho(\phi', F_S) = \|\tau_r\|^2 \qquad (9)$$

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 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 wrench applied to the root to enforce statics contains information on the physical plausibility of the pose for F_S. 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. Assuming forces at the holds are explained by the position of the COM, as illustrated by Figure 6, 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 weight of the simulated skeleton and the sum of forces at the holds and is therefore pose-invariant.

Torques at the inner joints describe the necessary effort to maintain the pose given by ϕ'. A small total inner torque is in general a good measure of high human plausibility. However, the simplicity of our simulated model makes it difficult to interpret inner torques, as it does not capture the shape of the holds or model the strategies used by different climbers to distribute effort among different muscle groups.

Therefore, we only use the torque 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 the 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 COM 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.

5.4 Hint Objective for improved optimization

As mentionned 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 COM, which we approximate as root position, lies in a sphere, and use physical plausibility to project that COM 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 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

$$\zeta(\phi) = \begin{cases} k\|r - r'\|^2 + G(\phi') & \text{if } \|r - r'\| > t, \\ \rho(\phi', F_S) + G(\phi') & \text{otherwise} \end{cases}, \qquad (10)$$

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 that grow as the skeleton faces away from 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. 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. Figure 8 illustrates the exploration of pose space by CMA-ES, where each picture contains the population cloud for the indicated iteration. Our projection method enables an easy exploration of the set of valid poses, and the optimization converges when in the vicinity of the optimal solution by progressively lowering the variance for the CMA-ES sampling variables (the desired joint angles ϕ).

6. RESULTS AND DISCUSSION

Our calibration methods were used to capture nine subjects and collect close to an hour of climbing data. Two routes were designed and each subject was asked to perform
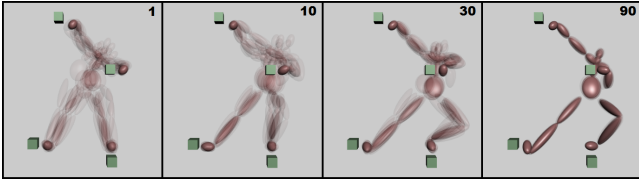
Figure 8: Vizualisation of the optimization process.

Sensor	Average ϵ_c (mm)	ϵ_s (Nmm)
1	4.3	75
2	3.3	55
3	4.6	92
4	3.4	87
5	4.4	102
6	4.1	96

Table 1: Average error accross the three contact point estimates and error for sensor frame estimate for each sensor.

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 2: Root position regression error.

several takes with objectives such as shifting COM position and completing routes following different progressions.

6.1 Calibration

Our setup supports the expedient calibration procedure: in practice, most of the capture sessions were recorded using the pocket-based procedure. 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, a number of tests as well as a full wall calibration were performed for the optimization-based calibration. For each sensor, three contact point locations were considered. For each contact location, 20 seconds of MoCap and 20 seconds of force measurements were recorded. Using Matlab, calibrating the calibration tool from 2500 frames of MoCap took less than a second, while estimating the position and orientation of a sensor with the iterative and alternating method from 2000 samples per point and 3 contact locations took less than a minute.

Contact point estimate error is expressed as the mean error for the least-squares estimate ${}^b p_c^*$. For each frame, ${}^w_b E_j$ is used to compute the contact location in world coordinates. Based on Equation 3, we describe the error as

$$\epsilon_c = \frac{1}{n} \sum_{j=1}^{n} \left\| {}^w_b E_j \ {}^b p_c^* - {}^w p_c^* \right\|, \qquad (11)$$

where ${}^w p_c^*$ is the average of ${}^w_b E_j \ {}^b p_c^*$ over all samples. This quantity can be seen as the size of the point cloud where the actual contact point lies in world coordinates. Table 1 shows average errors for contact point estimates at each hold for a full calibration procedure. The high errors can be explained by the use of the matrices ${}^w_b E_j$, which are estimated by the software from the position of four markers, each of which has an individual reconstruction error in the order of 0.3 mm.

Sensor frame estimate error ϵ_s is expressed as

$$\frac{1}{km} \sum_{i=1}^{k} \sum_{j=1}^{m} \left\| {}^w_s R^* \ {}^s \tau_{ij} + ({}^w p_s^* - {}^w p_{ci}^*) \times {}^w_s R^* \ {}^s f_{ij} \right\|, \quad (12)$$

where ${}^w_s R^*$ and ${}^w p_s^*$ are the least-squares estimates for the sensor frames. Table 1 shows average errors with a contact point estimation error in the order of 4 mm, and sensor noise in the order of 0.1 N for forces and 1 Nmm for torques. Since a torque is a force applied at a distance, and given that in our setup contacts with holds occur approximately 15 cm from the origin of the sensors, an error of 75 Nmm represents a weight discrepancy of a mere 50 grams at the sensor. The iterative solver stops when the change in error is less than 0.01 Nmm per iteration.

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 give any indication of the calibration precision, it is a good means of detecting errors in recorded gains, excitation voltages and sensor orientations with respect to the wall.

6.2 Reconstruction of static poses

Using a data set containing a wide variety of static poses and over 4000 frames of synchronized MoCap and force data, we looked at how well linear regression on the position of the root generalizes. We trained the model on 90% of the samples and tested it on the remainder. Table 2 shows the distance between predicted and actual root position for 6 such 90/10 partitions. As expected, dynamic motions and higher errors coincide, confirming our hypothesis that the correlation is more pronounced in the static case.

Pose reconstruction was run on several, varied static poses. The reconstruction is compared to ground truth for 5 of these poses in Figure 9, and timings as well as error measurements for those runs are given in Table 3. Reconstruction error is computed as the average distance between geometric centers of two corresponding bodies from the optimal pose and the ground truth. Examples were run using ODE and Java, on a dual-core 3 GHz Intel i3 CPU with 4 GB RAM. CMA-ES runs for 100 iterations with 10 samples per iteration, and the parameters for evaluating the fitness of a pose are the desired angles for hips, knees, ankles, chest, shoulders, elbows and wrists. The threshold to stop applying 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 provides an accurate model for reconstructing the remaining degrees of freedom.

7. CONCLUSIONS

We present an instrumented climbing wall, its design, and calibration procedures. We also present new ideas in computing postures during climbing from forces alone, using optimization to find a physically valid pose while meeting other important plausibility constraints. While linear regression cannot directly provide good estimates of posture from forces, it is useful for identifying the location of the
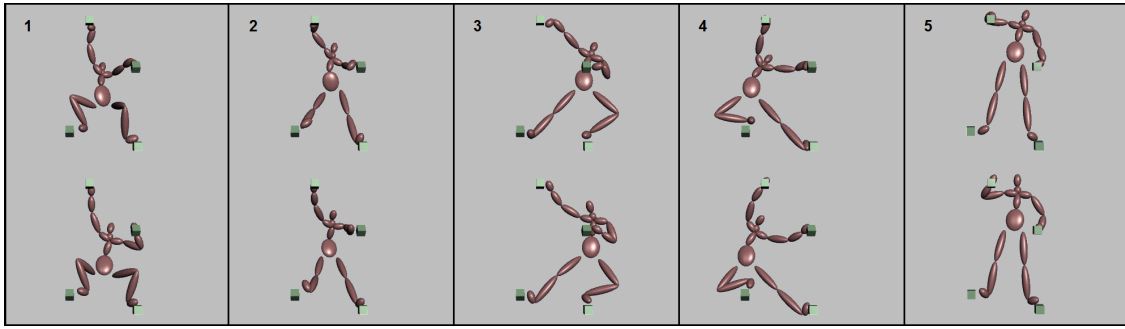
Figure 9: Reconstruction examples. At the bottom, the ground truth. At the top, the result of the optimization.

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 3: Durations and reconstruction errors for the pose reconstruction examples.

hips, which we use as an inexpensive hint to guide our optimization towards the optimal solution. Comparison between our reconstructed poses and MoCap show that our objective function is a good model for human posture.

In the future, several avenues can be explored, such as automating the calibration procedure and improving the calibration tool, as well as expanding the framework to study dynamics using the physical simulation.

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. We believe that our methods for estimating posture from forces will be useful for augmented reality applications. For instance, in learning or rehabilitation, data projectors could be used to provide real-time visual feedback directly on the wall as a person is climbing.

8. REFERENCES

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

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

[3] 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.

[4] 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.

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

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

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

[8] P. G. Kry and D. K. Pai. Grasp Recognition and Manipulation with the Tango. In *ISER*, volume 10. Springer, 2006.

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

[10] 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.

[11] J. Pfeil, J. Mitani, and T. Igarashi. Interactive Climbing Route Design Using a Simulated Virtual Climber. In *SIGGRAPH Asia Sketches*, pages 2:1–2:2, 2011.

[12] 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. *Human Movement Science*, 16(2–3):337–346, 1997.

[13] 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.

[14] 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.

[15] P. Rougier, R. Billat, R. Merlin, and J.-P. Blanchi. Conception d'un Système pour Étudier la Relation Posturo-Cinétique dans un Plan Vertical. *ITBM*, 12:568 – 580, 1991.

[16] 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.

[17] 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.

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