A Simplified Inverse Kinematic Approach for Embodied VR Applications

Daniel Roth\textsuperscript{*}  
University of Würzburg

Jean-Luc Lugrin\textsuperscript{†}  
University of Würzburg

Arnulph Fuhrmann\textsuperscript{¶}  
TH Köln

Julia Büser\textsuperscript{‡}  
University of Cologne

Gary Bente\textsuperscript{§}  
Michigan State University

Marc Erich Latoschik\textsuperscript{||}  
University of Würzburg

e-mail: marc.latoschik@uni-wuerzburg.de

e-mail: arnulph.fuhrmann@th-koeln.de

e-mail: bente@uni-koeln.de

e-mail: julia.bueser@smail.fh-koeln.de

e-mail: jean-luc.lugrin@uni-wuerzburg.de

e-mail: daniel.roth@uni-koeln.de

ABSTRACT
In this paper, we compare a full body marker set with a reduced rigid body marker set supported by inverse kinematics. We measured system latency, illusion of virtual body ownership, and task load in an applied scenario for reducing acrophobia. While not showing a significant change in body ownership or task performance, results do show that latency and task load are reduced when using the rigid body inverse kinematics solution. The approach therefore has the potential to improve virtual reality experiences.

Index Terms: H.5.1 [Information Systems]: Artificial—Augmented and Virtual Realities

1 INTRODUCTION

Full body motion capture for real-time embodied virtual reality (VR), is a powerful tool for body ownership (BO) illusions [11]. It was shown that full body motion can be tracked using single depth images [10]. However, single camera approaches often demonstrate high latencies for pose estimation [6]. In turn, optical tracking systems often require wearing a suit with markers that causes discomfort and extends preparation time. We developed an rigid body (RB) inverse kinematics (IK) solution and evaluated the system’s use for VR applications in an exemplary study using a pit adaptation for acrophobia induction. Our goal was to develop a simplified and replicable solution that reduces computational and network workload and thus lessens the overall system latency. We hypothesized that users are more comfortable with the less “invasive” rigid body (RB) marker set and thus would perceive a lower task load.

2 RELATED WORK

IK allows avatars to synthetically reproduce human motions, even with a limited number of sensor points [9]. Previous research showed that IK solutions can be accurate and are applicable for real-time virtual environments [7]. We investigated prominent IK-based solutions available for Unity3D, and Unreal Engine which are popular simulation engines. The “Inverse Kinematics” asset by Dogzer [2] and “UnrealMe” (used in [8]) require eight rigid bodies, which we considered to be too many and to increase the user’s discomfort. By comparison, we found that when using the “Simple IK” asset by Takohi Games [12] small movements had too extensive impact on the full body motion reproduction quality. Hence, as a trade-off between user’s discomfort and tracking quality, we developed a simplified IK-based solution that uses only five RBs but nevertheless focuses on exactly tracking representative limbs.

3 METHOD

Our approach uses five RB markers attached to both hands, both feet, and the head. The markers are tracked by an OptiTrack tracking system and results are sent via NatNet to the application. IK calculation is then performed in Unity3D by the Mecanim Animation Tool [13], and its functions for IK forehand and foot effectors.

We simplified the approximation of the torso position by centering the torso between both feet for the x- and z-axis. To allow for a torso flexion, the head position was accounted for with a factor of 0.2 to calculate the new root position. A linkage to the lower-positioned foot in world space determines the y value for the trunk. Our approximation for the body’s rotation accounts for both feet rotation and head rotation. Finally, we linked the head’s position and rotation directly to the head RB, attached to the head mounted display (HMD).

4 EXPERIMENT

We designed an experiment to compare the RB IK approach with a regular full marker set (FM) using a virtual pit environment (see Figure 1). We applied a between-group design (full marker set vs. reduced RB IK marker set). To account for eventual biases, we pretested acrophobia level using a simplified questionnaire (from a scale of one to three, how frightened are you: i) to go up steep stairs, ii) to go over a high bridge, iii) to look out of an airplane window high over the earth) and then used the scaling to evenly assign subjects to the groups. After a short VR acclimatization, we asked subjects to cross a plank three times between two skyscrapers in a height-inducing scenario. The plank was rebuilt in reality and had the same measurements and position as in the virtual scenario.

For motion tracking, we used an OptiTrack system with 12 Flex13 cameras, covering approximately 4x5m of active tracking area. We adapted an NatNet interface to stream RB and skeletal data to Unity3D and equipped users with a wireless backpack laptop and an Oculus DK2 HMD. Depending on the subject’s gender, we used male and female virtual characters.

The following measurements were taken in order to evaluate and compare the system:

Figure 1: FM markers (A), RB IK markers (B) and Virtual Pit (C).

\textsuperscript{*}e-mail: daniel.roth@uni-koeln.de
\textsuperscript{†}e-mail: jean-luc.lugrin@uni-wuerzburg.de
\textsuperscript{‡}e-mail: julia.bueser@smail.fh-koeln.de
\textsuperscript{¶}e-mail: bente@uni-koeln.de
\textsuperscript{§}e-mail: arnulph.fuhrmann@th-koeln.de
\textsuperscript{||}e-mail: marc.latoschik@uni-wuerzburg.de
- Latency: Video measurement with 240fps cameras similar to [4], but based on apexes of repeated motions.
- Body Ownership (BO): Adaptation of the scale for the Illusion of Virtual BO from [8].
- Task load A raw NASA Task load index [3] to analyze workload differences between the groups.
- Bias control: Pre- and post-acrophobia [1], immersive tendency [14], and task performance by video analysis.

5 Results

We performed frame counting measures for 35 repetitions with two raters and two camera types. As depicted in Figure 2, the analysis showed that the RB IK approach had a lower end-to-end latency ($M_{RB} = 104.76$, $SD_{RB} = 11.58$) compared to the FM approach ($M_{FM} = 119.82$, $SD_{FM} = 17.44$). Furthermore, visual analysis of the RB IK samples showed fewer outliers (Figure 3).

We excluded one participant from the analysis because of sickness and nausea, before and during the study. The final sample consisted of 27 (14 female, 13 male) participants ($M_{age} = 24.48$, $SD_{age} = 4.06$, range = 18 – 35), 18 were students. Only five participants had previous VR experience. The total performance time for the three trials, immersive tendency, or pre-acrophobia avoidance/anxiety was not significantly different between the groups (all $p > .05$). Subjective measures showed no significant differences between the FM and RB IK condition in the BO subscales or SSQ (all $p > .05$). The hypothesized reduction of task load however was significant ($t = 1.811$, $p = .04$; one-tailed).

6 Conclusion

Our results show that latency and the users’ task load can be reduced by using a simplified IK-based approach with a reduced set of markers for embodied immersive VR applications. While no significant differences in simulator sickness, performance, or BO arose, this approach provides a useful tradeoff between tracking quality and latency. We interpret the reduced task load to be a factor arising from the more comfortable experimental situation for the user, as our approach increases mobility and does not require users to wear a tight tracking suit. Based on our results a simplified RB-IK can be beneficial for single user (mirror-less) applications. Our interpretations are limited to a single apparatus and an IK model that may lack reproduction quality of extreme trunk and head movements and bending (e.g., the RB IK body animation broke when crawling over the plank or bending to large degrees). Therefore further measurements on different configurations need to be realized.

References