Zero Gesture: Exploring Full-Body Gestural Interfaces using Motion-Capture in Simulated Micro-Gravity

Connor Dickie[1]
Human Media Lab, Queen's University[1]
Kingston, Ontario. Canada.
connor@cs.queensu.ca

Paul Strohmeier[1,2]
University College Maastricht[2]
Maastricht, the Netherlands.
paul@cs.queensu.ca

ABSTRACT

Full-body motion capture systems like Vicon and Kinect have enabled user interface researchers to create a rich gestural language that allows users to interact with computers in ways that are difficult or impossible with traditional input methods. This paper describes a novel underwater motion capture system that enables users to operate in a simulated micro-gravity environment. By removing the gravitational constraint that has been a constraining factor for previous research in this space we show that there is opportunity to extend the existing gestural vocabulary to support a new class of interactions.

ACM Classification: H5.2 [Information interfaces and presentation]: User Interfaces. - Gestural User Interfaces.

General terms: Design, Human Factors, Performance.

Keywords: Gestural Interface, Micro-Gravity, Motion-Capture.

INTRODUCTION

The study of realtime full-body gestural interfaces using motion-capture began in the late 70's with simple systems that used video cameras and early computer vision tools used to track a collection of markers attached to a performer's body [1]. Despite the primitive affordances of such systems they represented a significant gain over previous manually post-processed techniques employed by researchers dating back to the late 1800's [2].

Today researchers have access to a number of compelling technologies that enable automated, realtime full-body motion-capture ranging from professional tools such as Vicon [3] which provides very precise spatio-temporal resolution, to the Microsoft Kinect [4] a consumer-grade tetherless solution for full-body motion-capture. Both the Vicon and Kinect have similar operating constraints in that they require an uncluttered stage for the actor to perform their gestures which may include simple pointing using only a single hand, or simulated running, jumping, twisting, bending and turning just to name a few of the possibilities.

MOTIVATION

Despite the wide array of previously identified body configurations that we currently exploit in gestural systems, actors are unable to achieve a large set of poses that are physiologically possible due to the fact that gravitational constraints severely limit what can be achieved. To address this shortcoming we developed a novel motion-capture system based on Kinect hardware to explore the possibilities of gestural interaction unencumbered by gravitational constraints.

SYSTEM DESIGN

Without access to a true micro-gravity environment like those encountered on the International Space Station or in a parabolic flight, our remaining option was to simulate a micro-gravity environment via submersion of the actor in water. Although the underwater environment is not a true micro-gravity, it is sufficient to test our preliminary designs and is an accepted technique that has been used since the 1960's to train spaceflight participants [5].

Implementing a full-body motion-capture system that operates underwater introduced a number of problems that had to be solved before useful tracking data was achieved. For reasons of cost and accessibility we employed a Kinect which uses structured-light to extract a realtime 3D model of the imaged scene. Specifically, a 940nm Near InfraRed (NIR) LASER coupled with a MEMS mirror array produces a known pattern that is projected into the environment which is then imaged by an associated NIR sensitive camera. The deformation of this known dot pattern caused by the scene is then used to produce the final 3D model.

The use of NIR and a fine dot pattern to create a 3D model was of particular concern during the conceptual phase of system design as water is known to not only attenuate NIR, but also has a 33% greater index of refraction than air [6], both of which could serve to introduce noise into the signal rendering the 3D model useless.

After a series of tests it was determined that not only is water sufficiently transparent at 940nm, but due to the non-compressibility of water the 33% refraction transformation was a constant that we could easily subtract from the values that were used to generate the 3D model.

Before we could determine the previously mentioned results we needed to waterproof the Kinect. This was achieved by removing the Kinect from the plastic housing and removing the fan used to vent heated air and submerging the entire unit in silicone gel.

UIST'11, October 16–19, 2011, Santa Barbara, CA, USA.

After preliminary experiments on disposable web-cameras we elected to keep the lenses of both the NIR camera and NIR LASER emitter exposed directly to the water in an effort to keep NIR signal attenuation and refraction to a minimum. The construction of the Kinect camera lens and LASER window is sufficiently waterproof to withstand immersion without additional considerations.

Despite these lucky engineering breaks we did encounter a number of issues that needed to be addressed. The simplest of these problems was the interference from the dot pattern reflecting off the interface between the surface of the water and the air. This was alleviated simply by making sure that the projected LASER did not intersect with the water/air interface. We achieved this by pointing the Kinect at a sufficient angle such that we no longer experienced interference.

Surprisingly, the most serious issue that we encountered was software-based. Like many computer vision systems, complexity of the feature-finding algorithm is lowered by automatically rejecting candidates that are predetermined to be uninteresting via template matching [7]. In the case of the human skeleton solvers employed by both the OpenNI drivers [8] and the official Kinect SDK from Microsoft [9] candidates that are tilted more than 80 degrees away from vertical (if considering the floor the actor is standing on as the normal from which measurements are taken) are automatically discarded. The interim solution that we adopted was to circumvent the limitations of the human skeleton solvers by dynamically re-orienting the Kinect such that it remained in sync with the actors orientation.

EARLY FINDINGS

In the short period of time that we have been using the system we have identified a number of compelling extensions to the gamut of gestures that we typically employ in full-body motion-capture systems. Immediately we can see that without the constraint of gravity we are able to engage in myriad "flying" and "hovering" type poses [Fig. 1]. Considering that many virtual environments employ flying as a means to quickly travel or survey the environment, being able to transition naturally into this mode maintaining a one-to-one correspondence with an avatar helps users understand how the system operates. Additionally many candidates for gestural interaction that were previously considered too difficult for the average user to accomplish are now more available to this audience simply because risk of injury is mitigated.

CONCLUSIONS & FUTURE CONSIDERATIONS

We have just scratched the surface of full-body gestural interfaces in a micro-gravity environment. Future development includes extending the human skeleton solvers to include the greater expressional capabilities and the refinement of our waterproofed motion-capture solution.

Figure 1: An image depicting the “flying” pose that one is able to achieve in a micro-gravity environment.

ACKNOWLEDGMENTS

We would like to thank Dr. Sarah Jane Pell and the Aquabatics Research Team at the University of Western Australia for their guidance and advice.

REFERENCES

1. Tom DeWitt, Pantomation, videotape, 10 min., in SIGGRAPH Video Review, No. 2, Section 9. 1981.
2. Eadweard Muybridge, Sallie Gardner at a Gallop, Zoopraxiscope, Palo Alto, California. June 19, 1878.
3. Vicon Motion Systems Inc. http://www.vicon.com
4. Microsoft Inc. http://kinect.xbox.com
5. NASA Neutral Buoyancy Lab. http://dx12.jsc.nasa.gov/history/index.shtml
6. Optics, Fourth Edition. Pearson Higher Education. ISBN 9780321188786. March, 18th. 2003.
7. R. Brunelli and T.Poggio, Face Recognition: Features Versus Templates., in IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 15, Issue 10. October, 1993.
8. OpenNI. http://www.openni.org
9. Microsoft Research, Microsoft Inc. http://research.microsoft.com/en-us/um/redmond/projects/kinectsdk