Real-Time Motion Capture for Interactive Entertainment

Richard Marks, Rishi Deshpande, Care Michaud-Wideman, Vangelis Kokkevis, Stewart Sargaison, Eric Larsen
Sony Computer Entertainment America R&D*

Abstract

Real-time motion capture technology will allow users to directly control the motion of virtual characters. This technology can be implemented by processing live video and by analyzing such information as movement, color, and depth. We have implemented several game ideas based upon this control methodology.

1 Introduction

Rapid advances in graphics and animation capabilities have greatly improved the user experience for modern interactive entertainment systems. Advances in user interface technologies will complement this by allowing users to input data more intuitively and interactively, thereby providing a richer, more immersive user experience. Our work examines the user of real-time optical motion capture as a user input mechanism.

Currently, optical motion capture is a widely used technique for obtaining animation data for movies and video games. The movement of actors and athletes is recorded very precisely in motion capture studios using expensive cameras and computing equipment. We perform motion capture in real time to allow users to interactively control virtual characters through related movements of their own body.

2 Exposition

We have investigated the use of motion, color, and depth as cues for motion capture. Depth is particularly robust as a segmentation mechanism, and can provide pixel-accurate distinction of foreground (user) from background (environment). Color is useful for labeling body parts, and motion is an excellent attentional mechanism.

In our exhibit, we demonstrate two different motion capture systems. The first uses color and motion data from EyeToy, a USB webcam peripheral for PlayStation2. Video (60Hz 320x240 YUV420) from EyeToy is processed by a PlayStation2 to infer the pose of a reference skeleton model. An assumption is made that the user is wearing a solid-colored, short-sleeved shirt. The color of the arms and torso are segmented and analyzed. The second motion capture system (Figure 1) uses a prototype RGBZ camera provided by 3DVsystems, Inc. The camera provides RGBZ for every pixel (320x120 at 60Hz). A PlayStation2 development system processes this output to infer the 3D pose of a reference skeleton model. This is accomplished by analyzing the foreground depth data and determining the likelihood a given pixel corresponds to a particular user-skeleton limb. The labeled image is fit with a pre-calibrated skeleton model of the user, and the skeleton pose and joint angles are the output of the vision system. For both motion capture systems, the raw data is fed into a physics-based animation system to provide a smoother result.

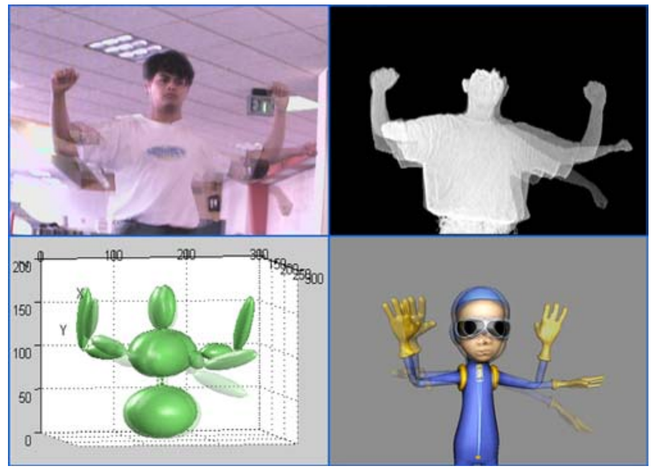

Figure 1. Live RGB and Z video streams of the player are used to compute the motion of a reference model designed for tracking. The computed motion is then mapped in real time to the virtual character *Seymour the Spaceboy*.

Several demonstration applications have been created that highlight the benefits of this user interface technology. The applications split into two categories: those that mix video and graphics (enhanced reality), and those that are all graphics. The all-graphics applications focus on an alternative form of control for animated video game characters. The enhanced reality applications focus on allowing the user to directly interact with computer-generated graphics, thereby blurring even more the distinction between what is real and what is virtual.

----

*website: http://research.scea.com