Alternative Performance Capturing

Michael Bußler, Simon Spielmann, Nicole Rothermel, Volker Helzle*
Filmakademie Baden-Wuerttemberg - Institute of Animation, Visual Effects and Digital Postproduction**

Abstract In this talk we present a combination of well-known and proven HCI techniques to create a markerless performance capturing system based on low-cost consumer hardware for live performances with virtual characters. The use case for our approach was a theatrical play on the occasion of the celebration of the 60th anniversary[1] of the federal state of Baden-Wuerttemberg, Germany, in which a virtual alter ego of the current prime minister interacted directly with the stage actors and musicians.

Introduction Marker-based motion capture systems have become the de facto standard for creating realistic and natural movement in computer animation. Such systems often require long setup times and are unaffordable for budget limited productions. Other methods, like the one proposed by [Stoll et al. 2011] use a set of low-cost, unsynchronized cameras to capture motion data without markers, but also require calibration and have high demands on computing power and memory to achieve almost real-time processing. The development of low-cost depth-sensing cameras like the Microsoft Kinect has made spatial capturing available to a large audience of consumers. Such cameras allow capturing of movement in space without markers by design: An actor can start performing immediately without need for markers to be placed. As shown by [Ye et al. 2012], it is sufficient to use three handheld Kinect cameras to gather highly detailed motion capturing data, even though not with real-time performance. Our approach combines two Kinect cameras for the capturing of body and facial performances of an actor, as well as two Nintendo Wiimotes, to control a virtual character live and in real-time. Operating on off-the-shelf hardware, the system is cost and time saving and therefore affordable for small budget productions.

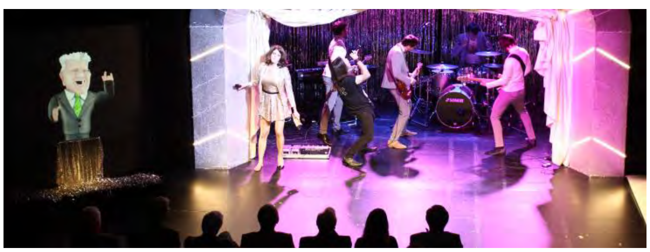

Figure 1. *The virtual prime minister performing live on stage.*

Facial and Body Capturing While body capturing alone gives a virtual character only basic movement like a puppet, the combination with facial capturing, head movement and gaze tracking yields a very natural look and feel. We implemented the system with Frapper, our open-source application framework, which is able to load and steer animated characters with pre-defined animation clips and to apply motion capture data received via network. The body capturing was implemented with the Microsoft Kinect SDK that allows defining orientation constrains for the bones and to stream skeleton data over the network. Relying solely on bone orientations, an actor can play different virtual characters independent of their scale and appearance. Unfortunately, the Kinect SDK is not able to capture hand gestures, so that we also used two Wiimotes to give the actor direct control over a set of pre-animated hand and body gestures, like pointing, waving, clapping and dancing or entering and leaving the stage. The facial performance capturing was done using FaceShift[2], a markerless capturing solution that streams FACS-based animation parameters as well as the head orientation and gaze direction using a Kinect camera.

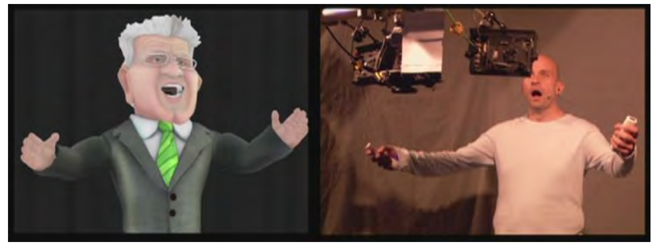

Figure 2. *Direct control of the virtual character.*

Discussion After collecting some experience with the arrangement of the two cameras, our alternative performance capturing system could be set up much faster than a marker-based capturing system, which requires time-consuming calibration steps. Although it took the actor some time to learn the 21 pre-built animations, he was absolutely convinced about our capturing solution, as he could just step on a predefined position and starts acting immediately. Some problems arose from the interference of the two Kinect cameras, especially in the region of the shoulders, which we could avoid by tweaking the solid angle of the infrared output pattern of the facial Kinect to light only a small region around the head. The actors' face had to stay within this area for full facial capturing, which was an acceptable limitation for our use case and could easily be handled by fixing the camera to the actors head. As we use only a single Kinect for the body capturing, the actor had to stand front-facing to the camera all the time. Using multiple Kinects would result in more freedom for the actor and more stable capturing results. Overall, our approach serves well the described use case of projecting an actor's performance directly on a virtual character, and could in principle be used as cost effective solution in other productions that involve direct interaction with a naturally behaving virtual character.

References

STOLL, C., HASLER, N., GALL, J., SEIDEL, H.-P., THEOBALD, C. 2011. Fast Articulated Motion Tracking using a Sums of Gaussians Body Model. In *ICCV,* 2011

YE, G., LIU, Y., HASLER, N., JI, X., DAI, Q., THEOBALD, C. 2012. Performance Capture of Interacting Characters with Handheld Kinects. In *Proceedings ECCV12*

[1] `http://research.animationsinstitut.de/171.0.html`
[2] `http://www.faceshift.com`

*e-mail: research@filmakademie.de
**web: http://www.animationsinstitut.de

SIGGRAPH 2013, July 21 – 25, 2013, Anaheim, California.

ACM 978-1-4503-2261-4/13/07