Documenting character performance using modern technology

C. Ye[1], S. Maddock[1], F. Babbage[2]
[1] Department of Computer Science, The University of Sheffield
[2] School of English Literature, Language and Linguistics, The University of Sheffield

1 Introduction

The documentation of performance can take many forms, from the mechanical to the emotional, from the objective to the subjective, from space-specific to the abstract, and from the personal to the collective. We are currently investigating how to document the mechanical aspects of a performance using a range of modern technologies. There are now many technologies that can be used to record a character's performance. The simplest is to film a performance producing a video record as two-dimensional (2D) data. More advanced three-dimensional (3D) motion capture technologies capture depth in a scene using multiple cameras. However, many of these systems are expensive and rely on a controlled environment, where the performance goes to the environment. In contrast, filming can easily be done in situ and multiple video cameras can be used, although there is then the problem of matching the data from each camera. More recently, cheap 3D data capture devices have appeared, such as Microsoft's Kinect[1] . This is not as accurate as the expensive systems and has capture area limitations, but offers the possibility of cheap, in situ 3D capture of performance. Our initial experiments are focussing on the movement of a single person captured using both standard two-dimensional (2D) video and also three-dimensional (3D) depth information from Microsoft's Kinect. The work is part of the wider, interdisciplinary RECITE research network[2] , which is investigating how theatre spaces and performance can be digitally mapped, modelled and documented. The following sections describe the work on video annotation and using 3D data capture, and are followed by a short summary.

2 Video annotation

This approach is based on the idea of rotoscoping, where an animator creates an animation sequence by tracing each frame of captured video. The resulting animation mimics, with artistic license to emphasize, the real captured movement. This is a technique that dates back to the work of Max Fleisher on Koko the Clown[3] and was used extensively in Disney films[2]. Our work instead matches a 3D skeleton with rotatable joints to each frame of video, as illustrated in Figure 1. Each joint of the skeleton is manipulated by the user to match the character's gesture in the particular frame of animation. The software is based

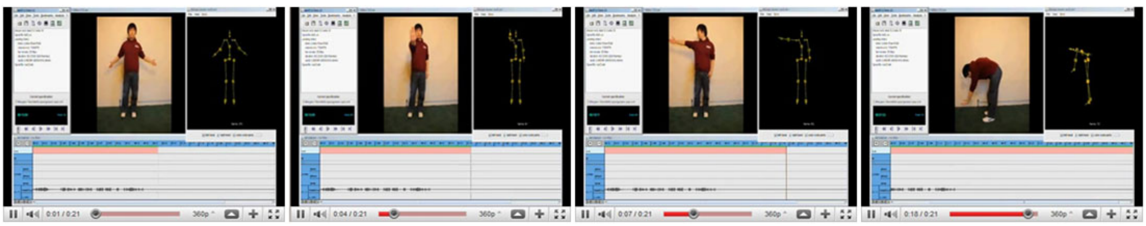

Figure 1: Annotation playback using Anvil

on Anvil[1], which uses a 3D virtual puppet to annotate hand gesture in video, and also documents

[1] http://www.xbox.com/en-GB/
[2] http://www.shef.ac.uk/recite/
[3] http://en.wikipedia.org/wiki/Koko_the_Clown

other media such as sound. Instead of the hand gesture tool, we use QAvimator[4] to support annotation of the whole figure. We do not label every frame of the video. Instead we label key poses and then use QAvimator to interpolate the motion between poses. This is similar to the production of cartoon animation where key frames are drawn and then inbetween frames are created using a process called interpolation. The current process is laborious and inflexible. The next stage of the work will attempt to automate the fitting of the skeleton to the video frames.

3 Three-dimensional capture

Microsoft's Kinect can capture both video and 3D data in a scene. We are focussing on the ability to capture the depth of object surfaces from the camera in a scene. From this data, the human figure is extracted and a skeleton is fitted, as shown in Figure 2. This work is based on OpenNI demonstration

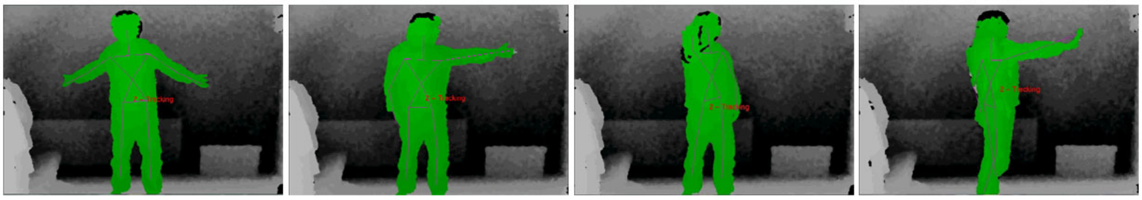

Figure 2: Tracking a person in 3D space

software[5]. We are currently working on extracting this data in real-time and integrating it with Autodesk MotionBuilder to drive a synthetic character in a 3D virtual world. This opens the possibility of not only documenting the movement but also allowing a performer to watch and interact with their own performance, producing performance-based animation.

4 Summary

Our initial experiments have shown that modern technology can be used to document the mechanical aspects of a human's performance, although the accuracy of the resulting motion needs further work. Our aim is to build in constraints, based on knowledge of likely human movements, into the capture process. This will be based on machine learning approaches and will speed-up the rotoscoping work and improve the accuracy of using Kinect.

However, it remains a moot point as to how accurate the capture process needs to be. It is not clear how much of the mechanical aspects of a performance are required to enable a description of the artistic impression of the movement. Whilst a beginner may learn mechanical movements, an experienced performer is registering movement at a higher level of description. Our future work needs to address this wider question.

References

[1] Nguyen, Q. and Kipp, M., Annotation of Human Gesture using 3D Skeleton Controls, In: Proceedings of 7th International Conference on Language Resources and Evaluation (LREC), ELDA, 2010.

[2] Thomas, F. and O. Johnston, Disney Animation: The Illusion of Life, Abbeyville Publishers, NY, 1981.

[4] http://www.qavimator.org/
[5] http://www.openni.org/