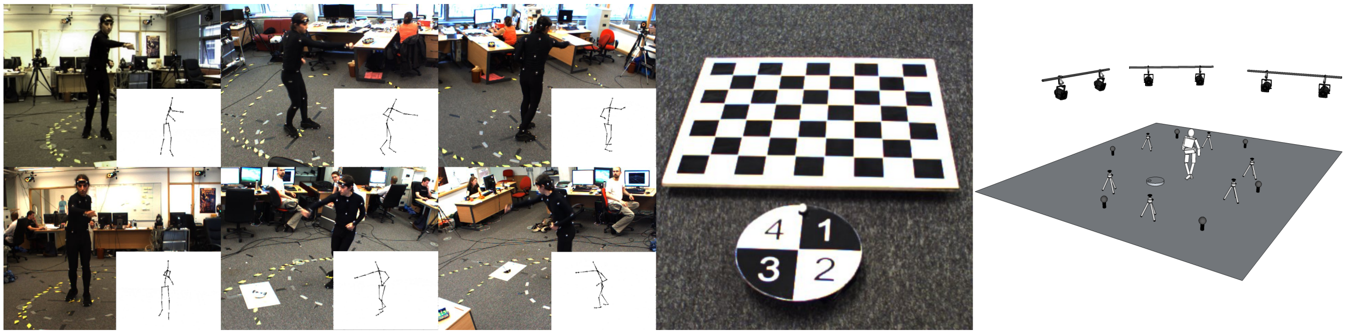
Synchronized Real-time Multi-Sensor Motion Capture System

Jonathan Ruttle* Michael Manzke† Martin Prazak‡ Rozenn Dahyot§
Trinity College Dublin



1 Introduction

This work addresses the challenge of synchronizing multiple sources of visible and audible information from a variety of devices, while capturing human motion in realtime. Video and audio data will be used to augment and enrich a motion capture database that will be released to the research community. While other such augmented motion capture databases exist [Black and Sigal 2006], the goal of this work is to build on these previous works. Critical areas of improvement are in the synchronization between cameras and synchronization between devices. Adding an array of audio recording devices to the setup will greatly expand the research potential of the database, and the positioning of the cameras will be varied to give greater flexibility. The augmented database will facilitate the testing and validation of human pose estimation and motion tracking techniques, among other applications. This sketch briefly describes some of the interesting challenges faced in setting up the pipeline for capturing the synchronized data and the novel approaches proposed to solve them.

2 System Overview

Video Capture: Six cameras (Point Grey Research Flea 2 FL2–03S2C with Fujinon Lens–15F3–60C), each 24-bit color, are capable of running at 80 frames per second (fps) with a 640×480 resolution. An important consideration with the cameras is the volume of data being produced. Six cameras running at 60 fps produce 2.67 Gbps of data which has to be saved to hard disk. Each camera requires calibration to determine its intrinsic and extrinsic parameters. This is being done using the standard checker-board calibration grid and the Camera Calibration Toolbox for Matlab. A unique feature of this acquisition system is that the positions of the cameras will be varied, from equally spaced 360 degrees around the subject, to one sided 180 degrees around the subject, to stereoscopic pairs of cameras. This allows a greater range of techniques to be tested on the repository.

Audio Capture: If the subject is making a noise, this will be picked up by the microphones. Using the time difference between when the noise reaches each microphone it is possible to calculate the position of the noise (subject) in the room, potentially a good first guess for the motion tracking.

Motion-Capture: Our marker-based motion-capture system (Vicon iQ 2.5) has the capacity of capturing data from 13 infrared cameras running at up to 200 fps. The motion-capture system requires the user to wear reflective markers at specific points tight on the body. This is best achieved by wearing a skin tight black Velcro suit, but ordinary clothes can be worn if they are skin tight and the markers are attachable. The Vicon's proprietary software is able to give a good estimation of the subject's skeletal motion from the relative position of the markers. However, when baggy clothes are worn the relative position of the marker to the bone is lost and the software cannot estimate the skeleton. In this case other methods will be investigated to provide ground-truth data and all raw data will be provided in the database.

Synchronization: Experiments are being carried out to determine the best method to synchronize all the cameras and different devices. All the cameras can be synchronized via hardware to achieve best results. To synchronize the cameras to the audio and motion-capture system a device is being built, as seen in figure above, consisting of a rotating disk which can be seen by at least one of the video cameras. The disk contains a reflective marker which can be seen by the motion-capture system and the device makes an audible beep on each rotation. This allows accurate synchronization of each device by aligning the three sources together in a post-process.

3 Results and Discussion

Early capture sessions are showing signs of good synchronization and high calibration accuracy. More work is being undertaken to produce a finished acquisition system. The resulting system should have potential for evaluating research methods in many different applications and scenarios, such as pose estimation and markerless motion tracking methods.

References

BLACK, M. J., AND SIGAL, L., 2006. HumanEva: Synchronized video and motion capture dataset for evaluation of articulated human motion, technical report CS–06–08. http://vision.cs.brown.edu/humaneva/index.html.

*e-mail:ruttlej@cs.tcd.ie
†e-mail:Michael.Manzke@cs.tcd.ie
‡e-mail:prazakm@cs.tcd.ie
§e-mail:Rozenn.Dahyot@scss.tcd.ie

SIGGRAPH Asia 2009, Yokohama, Japan, December 16–19, 2009.
ISBN 978-1-60558-858-2/09/0012