Edward Ly

CS 488 Annotated Bibliography

Graham-Knight, K., & Tzanetakis, G. (2015). Adaptive Music Technology Using the Kinect. *PETRA '15*. 32. doi:10.1145/2769493.2769583

Graham-Knight and Tzanetakis from the University of Victoria use the Kinect for the purpose of creating touchless musical instruments for people with disabilities. The positional data from the Kinect is then sent to the visual programming language Max/MSP through the Open Sound Control (OSC) Protocol for analysis and playback. The biggest limiting factor of this system, however, is the 857 ms average latency, which is too large to be practical for live performances, but it can be instructive to see if there is a similar implementation that has a latency that is small enough for people to not be able to hear.

Han, S., Kim, J., & Kim, J.D.K. (2012). Follow-Me!: Conducting a Virtual Concert. *UIST '12*. 65-66. doi:10.1145/2380296.2380324

Han et al. from the Advanced Media Lab at the Samsung Advanced Institute of Technology in South Korea developed their own specialized system for gesture recognition by using ultrasound to gather 3D positional data while also using that data to control tempo, volume, as well as instrumental emphasis at the same time. Their approach is similar to what I originally desired to accomplish earlier this year, except that now, I would also like to experiment with whether or not it is possible to do the same on the Kinect. This paper and its list of references are relatively short, however, but they explain the relevant parts of the implementation well enough to be able to understand the mechanisms behind the device.

Lee, E., et al. (2006). *iSymphony*: An Adaptive Interactive Orchestral Conducting System for Digital Audio and Video Streams. *CHI EA '06*. 259-262. doi:10.1145/1125451.1125507

Lee et al. of the Media Computing Group at the RWTH Aachen University in Germany created a system using modified Buchla Lightning II batons that uses gestural recognition to not only control the tempo, volume, and instrumental emphasis of an orchestral audio recording, but also control the playback speed of the accompanying orchestral video recording as well. The gesture recognition itself uses a framework called Conducting Gesture Analysis (CONGA) to detect and track beats, while a variation of a phase vocoder algorithm with multiresolution peak-picking is used to render real-time audio playback. While video control is an interesting and natural extension of the audio control, it will likely be outside the scope of my project, although I would imagine that the implementation of controlling both audio and video would not be too different. Like the previous paper, I would also like to experiment with whether or not it is possible to do the same using the Kinect instead.

Pellegrini, T., et al. (2014). Towards soundpainting gesture recognition. *AM '14*. 18. doi:10.1145/2636879.2636899

Pellegrini et al. created their own gesture recognition system using RGB/depth cameras for the specific purpose of soundpainting, composing music through new gestures rather than controlling a set orchestra, in either live or studio environments. Here, Hidden Markov Models are used to detect a custom set of gestures for variety of different trigger events such as “Play” and “Whole Group.” This paper presents an alternative application of gestural recognition as a custom virtual instrument, although if there is a desire for the music to be at a certain tempo, there is no mention of how the computer can adapt to this aspect of the music. Tempo recognition is also important when trying to detect beats in preset music.

Sarasúa, Á. & Guaus, E. (2014). Beat Tracking from Conducting Gestural Data: a Multi-Subject Study. *MOCO ’14.* 118-123. doi:10.1145/2617995.2618016

Sarasúa and Guaus from the Escola Superior de Musica de Catalunya in Spain have developed and tested a computer’s beat-detection capabilities using the Kinect as well as other modules and frameworks, both open and proprietary, to analyze the input data. Human participants were also involved to contribute to the computer’s learning capabilities, even though human error and time deviations had to be taken into consideration. This paper is especially useful given that sections 2 and 3 describe the approach used in developing the software for this kind of problem. My goal is to add to this software by having the music react to the change in tempo given by the live placement of beats.