6.111 Final Project Abstract

Auditory Localization

Francis Wang and Keshav Gupta

In this project we aim to demonstrate a system that is able to discern the direction of incoming sound through an array of four microphones arranged in a cross. At the center of the cross will be a camera, and the position of the inferred sound source will be overlaid on the scene and displayed on a monitor. We will use the Adafruit I2S MEMS Microphone Breakout Board based off the SPH0645LM4H IC for the microphone array. The microphone is digital, giving much better noise tolerance during data transmission than an analog equivalent.

We will infer the direction of the sound source by the relative delay of the waveforms measured by the two orthogonal pairs of microphones. The key challenge of this project would be the digital signal processing required to extract the delay from the waveforms. This is a well understood problem and the standard approach in the literature is the so-called generalized cross-correlation phase transform (GCC-PHAT) method. The algorithm requires taking the discrete Fourier transform of the signal and normalizing the amplitudes by computing some inverse square roots, both of which are computationally non-trivial on the FPGA hardware. We believe we can get it to work with Xilinx's IP cores, although as a fallback plan we can just compute the straightforward time-domain cross correlation to get an estimate of the relative delay.

The signal processing required for auditory localization is highly parallelizable and is well suited for implementation in hardware. It has applications in robotics and other interactive systems such as smart home devices.