Introduction

The sensory capabilities of our devices are reaching out into the third dimension. iPhones with camera capabilities—which seemed like a novelty just a little over decade ago—are now commonplace, and instead of just capturing images of our surroundings we are becoming increasingly interested in sensing, modeling, and making sense of the 3-dimensional spaces we inhabit.

We wish to solve a small subset of this problem: finding the precise location of an object in 3-dimensional space, indoors (e.g., in a setting not accessible by GPS). Additionally, we would like to solve this problem without the need for additional expensive hardware, and we would like to be able to locate things passively, (i.e. without the need to mount sensors/transmitters on those things which we wish to locate).

To solve this problem, we propose an innovative solution using 3D acoustic multilateration using a network of distributed iOS devices. By making use of conventional and readily-available hardware, we can lower the barrier of entry for 3D location systems, and since sound is generated by a vast number of potential objects of interest, acoustic multilateration provides a flexible method of localization that can be applied to a number of use cases.

Objectives

The final objective for our project was a distributed network of iOS devices (latest generation only) where each device was placed in different locations throughout a room. By streaming the audio between each of these devices, and applying some signal processing, we can calculate the time-difference-of-arrival for the waveform of the noise that we wish to localize. From this information, we can construct a matrix equation representing our physical system; we then solve this equation to find the actual location of a sound source in 3D space. For solving the multilateration equation in 3D space, 5 receivers are required; in 2D space, only 4 receivers are required. The approximate software architecture of our system is below.

Given a set of audio signals from all the iPhones in the network, we must calculate the relative time delay between each of them. To do this, we use a cross-correlation method. We first filter the signals with a linear-phase FIR bandpass filter; we then filter the signal in reverse to achieve twice the filter order and to make the overall transfer function zero-phase. At this stage, we also downsample the audio signals to reduce the computational cost of the cross-correlation. Once the signals have been filtered and downsampled, they are cross-correlated and the maximum value in the resulting signal corresponds to our relative delay.

Conclusion

The biggest challenge in implementing this project was getting accurate time synchronization; the time synchronization proved to be the limiting factor in the accuracy of our system. An example of the delay-compensated audio signals captured by our network are presented below; after much optimization of the network infrastructure and latency, we were able to get audio windows to arrive in the correct order, and within 10ms of their correct arrival time over both uchiigo-secure, and a private WiFi network.

However, while this accuracy was impressive considering the asynchronous multi-threaded nature of the iOS networking stack, and the short time we had for development, it was not accurate enough to achieve realizable accuracy for multilateration.

To test our implementation of acoustic multilateration, we built a servo-powered motorized mount for a GoPro camera; this was controlled via PAM from a Raspberry Pi that was also connected to our audio network. The Raspberry Pi used the results of the multilateration to track the sound source in real-time (as we were doing approximately 10 multilateration calculations per second) . We colloquially named this device, shown to the right, as Casper.

Due to this latency uncertainty, we had trouble achieving the accuracy that we desired; in a large, quiet room (50ft x 30ft x 60ft), we were able to get our GoPro mount to successfully pan left, center, or right with roughly 70% accuracy. This accuracy decreased significantly in a smaller room where we did not have the accuracy gains afforded by having each receiver separated by large distances. In a smaller room, our accuracy decreased to below 50%; this made the technology unreliable for regular use.

While the idea behind this project is still technically feasible, a significant challenge still lies in improving the synchronization of the audio network; without highly accurate synchronization, any attempt at multilateration will result in calculated locations with extremely large uncertainties. With more development time, more sophisticated latency-compensation techniques could be designed that might increase the accuracy of our system such that it could become feasible for a wide number of applications.

References


"Motion Capture Systems | VICON." 7 Feb. 2016 <http://www.vicon.com/>


Acknowledgements

This project was completed as part of CS 234/334 Mobile Computing (Winter 2016), taught by Prof. Andrew A Chien, with TA support by Yun Li and Yan Liu.