A Simple Direction Of Arrival Test (Sound Source Localization) (5/12/2012)

Introduction

As WIV processing inherits many of the characteristics of the human cochlear, it is not surprising that it can be used to determine the direction of arrival of sounds. To do so, it is necessary to measure the arrival times for a sound to reach two separate microphone placed in a manner to represent two ears.

With an accurate time measurement, the difference between the arrival times at each microphone can be used to accurately determine the source placement in a two-dimensional plane. Correspondingly, the placement of three microphones can localize a source in a three dimensional space.

The Setup

Two Shure unidirectional recording mikes were placed level to each other on stands approximately five feet high and separated by one foot. The stands were placed in the corner of a 30ft by 30ft room and pointed to the center of the room. No effort was made to reduce sound reflection by the room's walls.

Each mike was connected to a software-recording studio; A/D sampling was set at 96,000 samples per second. To insure the accuracy of the difference measurement, both A/D channels were latched simultaneously before sampling.

As this was a speaking voice and to keep extraneous high-frequency sounds out of the system, a digital 1KHz low-passed filter was applied to the channels during recording. The 1 KHz value was selected based on the speed of sound, the one foot separation, and the simplicity of the algorithm being used.

The sound source was a male staffer standing fifteen feet from the microphones and 40 degrees to the left. The staffer remained 15 feet from the microphone during the experiment and moved within the room in three increments. (40 degrees to the left of the microphone; center to the microphone, 40 degrees to the right of the microphone.)

Starting on the left, the staffer spoke "left", <pause>, "left", <pause>, "left" in a normal speaking voice. He then moved to the center and similarly was recorded speaking "center", <pause>, "center", <pause>, "center". And lastly, he moved to the right of the microphones and spoke "right", <pause>, "right", <pause>, "right".

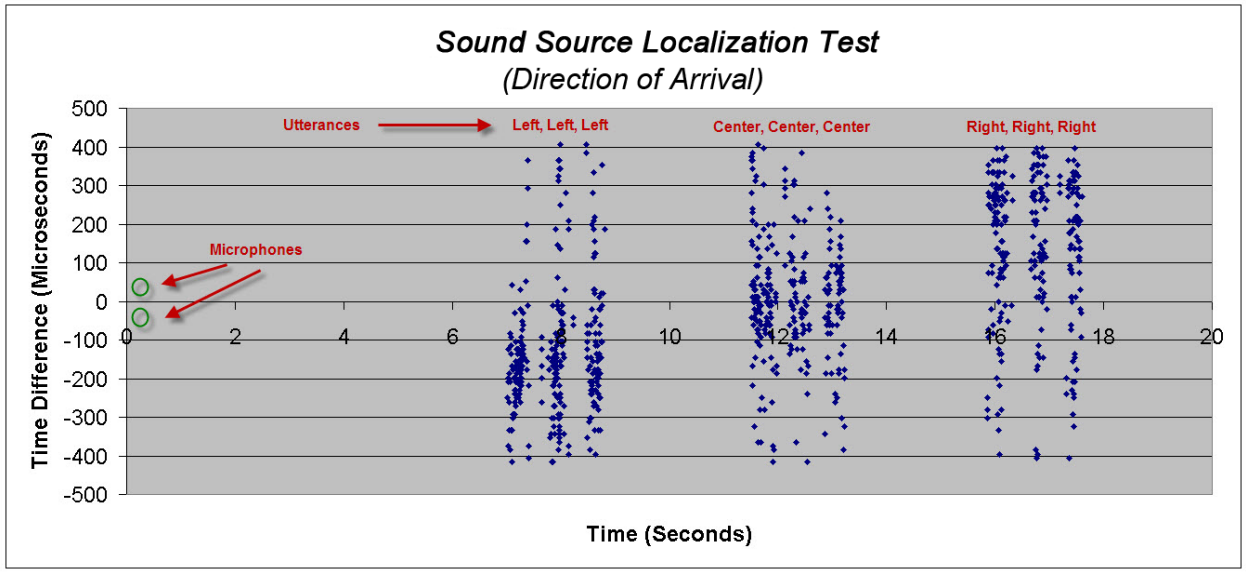
Classroom Lab Notes

Processing

The recording was stored in a file and imported into Audacity – an audio editing package. Using Audacity, the audio was amplified so that peaks were about -2 Db and the quiet space was reduced to -40 Db. The edited data was again written to disk.

The WIV processing was performed using a custom C++ program that imported the edited audio file, the audio data was converted to WIVs and direction of arrival data to another file. The data from this file was then imported into Microsoft Excel for plotting.

In Excel, low amplitude WIVs were ignored as a way of reducing noise. Using a simple formula, the time differences between each channel were used to calculate the location within the room. The plot shows the calculated time difference between the two channels against the linear timeline of the experiment. The results are seen below.



This was a simple test but it can be seen that the grouping of the WIVs correlates well with the speaker's position in the room. Localization could have improved by using a more robust mathematical algorithm and by reducing sound reflections within the room.