ESE 650 Final Project: Acoustic SLAM

Mitch Fogelson, Tyler Altenhofen MFOG@SEAS.UPENN.EDU, TYLERALT@SEAS.UPENN.EDU

Abstract

For this project we investigated localization and mapping of a microphone using acoustic signals. In our experiments we had four speakers arranged in an arbitrary configuration as well as our microphone, which had line of sight to each speaker. We were able to get both the mapping and localization to work relatively well, but unable to integrate the simultaneous localization and mapping together due to high noise in our measurements. In this report we will talk about our setup, methods, results and further discussion.

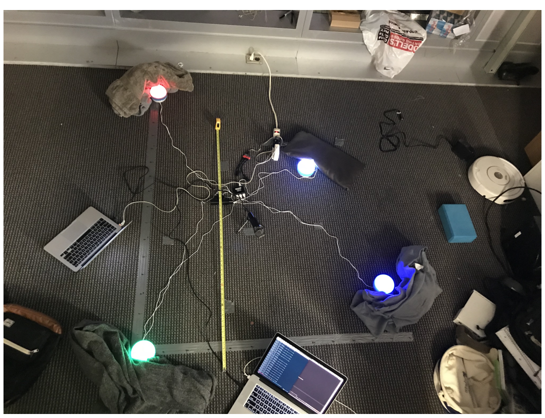

Figure 1. The Speaker Set-Up in Our Apartment

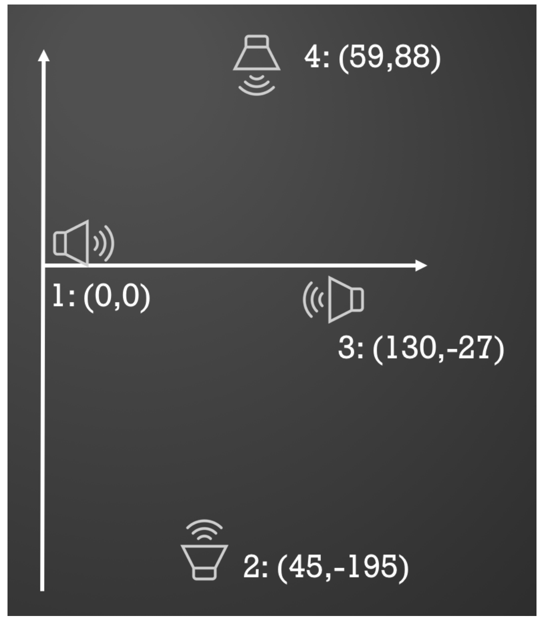

Figure 2. A Graphical Representation With the Positions in cm

1. Introduction

1.1. Objectives

In this project we set out to implement a particle filter for acoustic SLAM. SLAM is simultaneous localization and mapping. The particle filter is a Monte Carlo based algorithm that spreads an array of possible positions based on a motion update with added noise. We then compare the particles based on the correlation with the projected map. Our map for such a project would include the four speaker locations. Therefore, the particle with the maximum correlation with the speaker locations will be the updated position and we will update the map using this particle.

University of Pennsylvania, ESE 650 Course Project.

1.2. Setup

We created our testing setup within our apartment. We laid out duct-tape to create a grid on our floor. This acted as our ground truth in measuring the accuracy of our system. We used 4 speakers bought from Amazon. These were low quality stereo output speakers. Each speaker had a unique chirp, a frequency output that is inaudible to humans but unique to each speaker and could be picked up by the microphone. We used an omni-directional microphone that could sample up to 48,000 hz. This microphone acted as our robot in our experiments and could be added on to any differential drive or other robotic platform. We used blankets and other clothing to help reduce the echoing in the room. (See Figure 1 for Setup and Figure 2 for our Ground Truth)

2. Methods

2.1. Mapping

The map of the environment is the locations of each of the speakers. As discussed before, our speakers will each output a uniquely identifiable chirp at the same time. This allows us to collect the arrival time of each these chirps with our microphone. Because we don't know the emission time of these chirps we can't immediately identify our distance from each of the speakers. What we can identify is the relative distance of each of the speakers (e.g. speaker 2 is 1 meter farther away than speaker 2). More specifically, if we have speaker arrival times we can say:

$$\Delta d_{i,j} = 343 * (a_i - a_j)$$

Where Δd_{ij} is the difference in distance from of speakers i, and j, 343 is the speed of sound in m/s and a_i is the arrival time of the chirp from speaker i.

If we know our location relative to a single speaker (that is one of our starting assumptions), we can identify the distance of any speaker by adding our distance from the known speaker to the two speakers' relative distances.

$$d_j = d_i + \Delta d_{i,j}$$

With this protocol we have measurements for the distance of all speakers provided we are given our location relative to a single speaker and a set of received chirps. We can now begin generating a map of possible locations for each of the other speakers.

For simplicities sake we will assume that the speaker we know our location relative to is at the origin. After receiving a chirp we calculate the distance of each of the other speakers. We then create a map of possible speaker locations. This appears as a circle centered on the location of the microphone with radius equal to the calculated speaker distance. After calculating this we add Gaussian blur to handle the noisiness of our reading. A picture of this can be seen in figure 3.

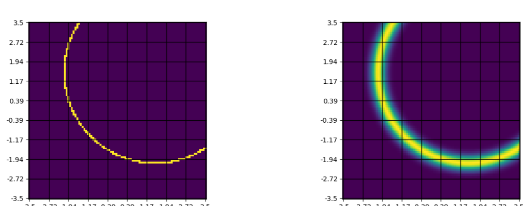

Figure 3. probability map of speaker 1 given a microphone location and distance. Centered on $(1.5, 1, 5)$ with radius ≈ 3

We use this circle estimation to continuously update our probability map. As we move around we update the map by adding in more rings of speaker locations. Figure 4 shows how we progress from a ring of possible locations to a more definitive estimate of where the speaker is located.

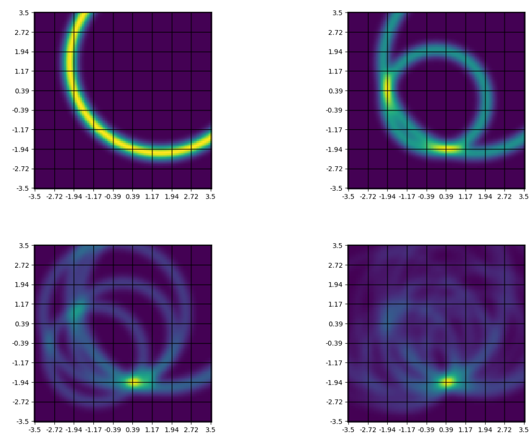

Figure 4. Probability map of speaker location given 1, 2, 5, and 20 samples respectively.

2.2. Localization: Particle Filter

Our state for our "robot" is a simplified unicycle model, where we can ignore the orientation since our microphone is omni directional. Our update model is movement in x and y with noise.

The other algorithm we implemented for our acoustic setup was a particle filter. To start out we are assuming that we have a set of particles with x and y positions, and a probability associated with each particle. We then apply a motion update which moves each of the particles in the correct direction and adds some Gaussian noise. This can be seen in figure 5.

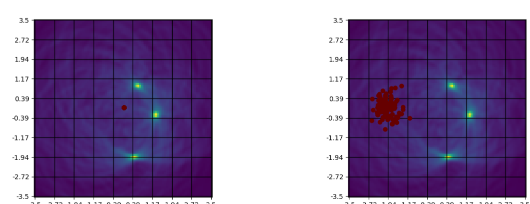

Figure 5. Applying an odometry and adding noise. Red dots are the possible positions.

After the odometry update and added noise, we update the probabilities of each of the particles. To do this we get another set of mic chirps and calculate the relative distances given each of the particles. We can then reconstruct the possible locations of each speaker. Figure 6 shows this ring of possible locations of speaker 1 overlaid on the associated existing microphone map.

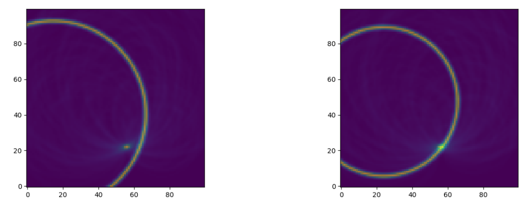

Figure 6. Particle testing out the fit with map of speaker 1. This is a map of possible speaker locations given the particle position and the receive chirps(circle part) overlaid on the known map of speaker 1. First image shows a badly fitting particle as the ring doesn't overlap with where we think the speaker is. The second shows a better fit that overlaps well with speaker 1.

With the map of known speaker locations and the map created using the particle and received set of chirps we can create a good heuristic for the probability of the particle position. This is done by running a correlation on the two maps and summing over the values. If the ring of possible locations overlaps with the know map location of the speaker we will get a very high value and if it doesn't we will get a low value. We calculate this heuristic for each speaker and multiply our original particle probabilities by the correlation. We follow this up with a regularization step. An update can be seen in figure 7.

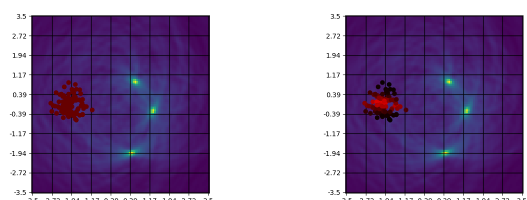

Figure 7. Updating particle probabilities given a received set of chirps. Brighter red correlated to higher probabilities.

We were only able to do our particle filter localization in simulation due to noise in our mapping model.

3. Results

3.1. Mapping

Our mapping results appeared promising, but performed poorly. In figure 8 you can see the errors of the speaker location to the ground truth varied from 34cm to 77cm. Given

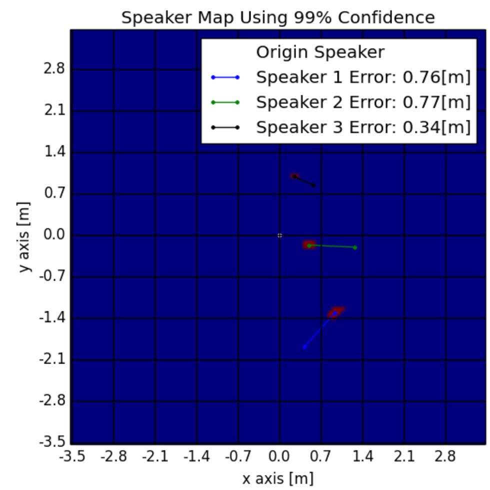

Figure 8. Mapping Results with Ground Truth Errors

that we were working in a space 2m by 4m, these results are less then ideal. In figures 9-11 you can see the heat map corresponding to each of the speakers individually with the ground truth denoted by a plus sign. To construct our combined map and determine accuracy, we converted each of the heat maps into a binary map. We do this by taking all points that were within 99 percent of the max value. We then took the centroid of the mask as our speaker location. We used the Euclidean distance as our error metric.

4. Discussion

In this section we will discuss several challenging aspects of implementing our project and enhancements to our project that could have gotten us better results.

4.1. Syncing Speaker Clocks

Syncing the clocks of our speaker was one of the most critical features of this project. If our clocks were off by 0.001 seconds, that translates in 34 cm error in our position estimate. We spent a lot of time figuring out how to prioritize sound output from our computers to reduce lag. In the end we came up with an interesting way to output our speaker chirps within 0.0001 seconds of our desired output time. We did this by starting our output thread and then programmatically supplying it with the correct signal so that the speaker would chirp at the right time.

4.2. Space

We decided to set up our experiments within our apartment since it was convenient location for us to work. However, our living room is quite small and has various walls and surfaces that can cause echos. Due to this we covered most

of our speakers and surfaces with blankets, in attempt to reduce this noise. In future experiments we would like to move to a larger room or outdoors where we would be less prone to such a situation.
We were also limited by the cord length of our speakers and extenders and given more space between the speakers, our errors would have been less significant to the scale of our map.

4.3. In Home Robot

We thought it would be an interesting idea for a home robot to have the ability to localized based on the sounds it hears as it navigates through a home. For example, if there is a way to determine a unique chirp for the appliances in the house, we may be able to implement our technique to localize and map out a home with a robot using this acoustic data. If we could configure the robot to localize based on these unique noises, this project might have some more practical applications.

5. Conclusion

In conclusion, we successfully implemented acoustic SLAM in simulation and had promising results for mapping in our experiments. We were able to adapt methods used in class with a unique and challenging set of sensor data. We developed a unique setup in our apartment to test our algorithms and were able to configure the hardware for our application. We learned a lot about the different methods to achieve SLAM and the limitations and extents of using sonar as a sensor source. Tyler did more on the mapping and particle filter framework and Mitch worked more on the communication protocol for the speakers and microphone. However, we worked closely together on each aspect of this project so it would be hard to separate contributions specifically to an individual.

6. Appendix

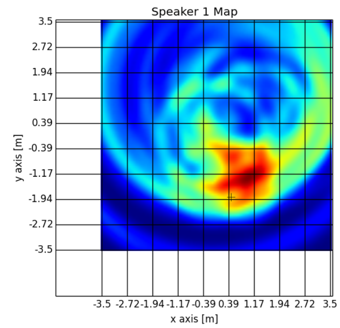

Figure 9. Heat Map of Speaker 1 Results with Ground Truth [+]

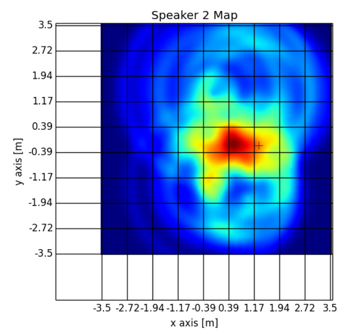

Figure 10. Heat Map of Speaker 2 Results with Ground Truth [+]

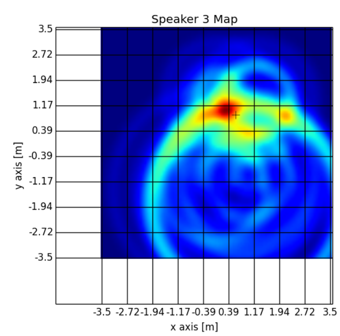

Figure 11. Heat Map of Speaker 3 Results with Ground Truth [+]