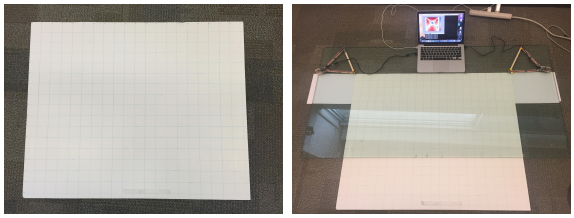
Paint with Music

XXX
XXX

XXX
XXX



(a) 1 meter by 1 meter grid (b) array placement

Fig. 1: Setup for localization accuracy testing

***Abstract*—The abstract goes here.**

I. Introduction

II. Experiment

To test localization accuracy, an one meter by one meter grid was set up and the arrays are placed at the top left and top right corners of the grid. Fig 1 shows a picture of the setup.

A total of 32 positions are chosen uniformly in this region where microphone data is recorded. To test how accuracy varies with window size, the algorithm is fed with recorded microphone data with different segment length. Fig 2 shows how accuracy changes with window length for three GCC algorithms. The error lowers as window size increases and plateaus after window size exceeds around 10 milli-second. The lowest error achieved is 2.53 centimeters. It is achieved when window size is set to 12 millisecond and GCC_PHAT is used to estimate difference of delay

Although accuracy improves with window length, the calculation time also increases with window length. The part of calculation that depends on window length is using cross correlation to estimate difference of arrival time. cross correlation can be calculated with FFT and the runtime is of order $O(N \log N)$. We measured how the computation time varies with window length and Figure 3 shows the result. The runtime increases approximately linearly in the window size region of interest.

We also calculated the localization error for each tested point in the region. Figure 4 shows a heatmap of the error distribution inside the grid. The error is below 3 cm for most areas inside the region. There is one error spike in the mid-left region and we contribute this to audio source placecment error becuase the error is fairly low and consistent around that spike region.

To test how well the arrays track movement, we mounted a rotating disk 40 centimeter in diameter onto the grid at

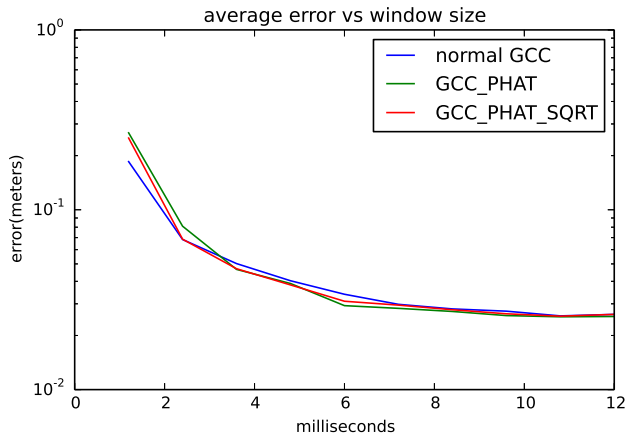


Fig. 2: accuracy versus window size

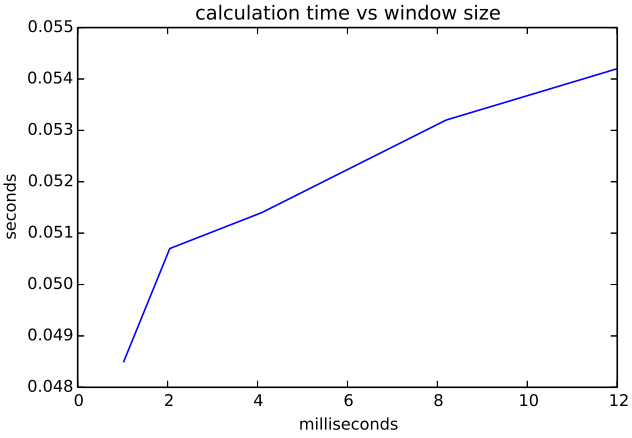


Fig. 3: speed versus window size

$(x = 0, y = -0.3)$. Fig 5 shows a picture of the sestup. A sound source is placed on the edge of rotating disk and the arrays localizes the sound source as it rotates in a circle. In this experiment, we want to test how accuracy changes with:

- different window sizes
- different audio sources
- different movement tracking filters
- different movement speeds

Fig 7 gives an intuitive representation of how accuracy changes with window size. When window size is small(1.02

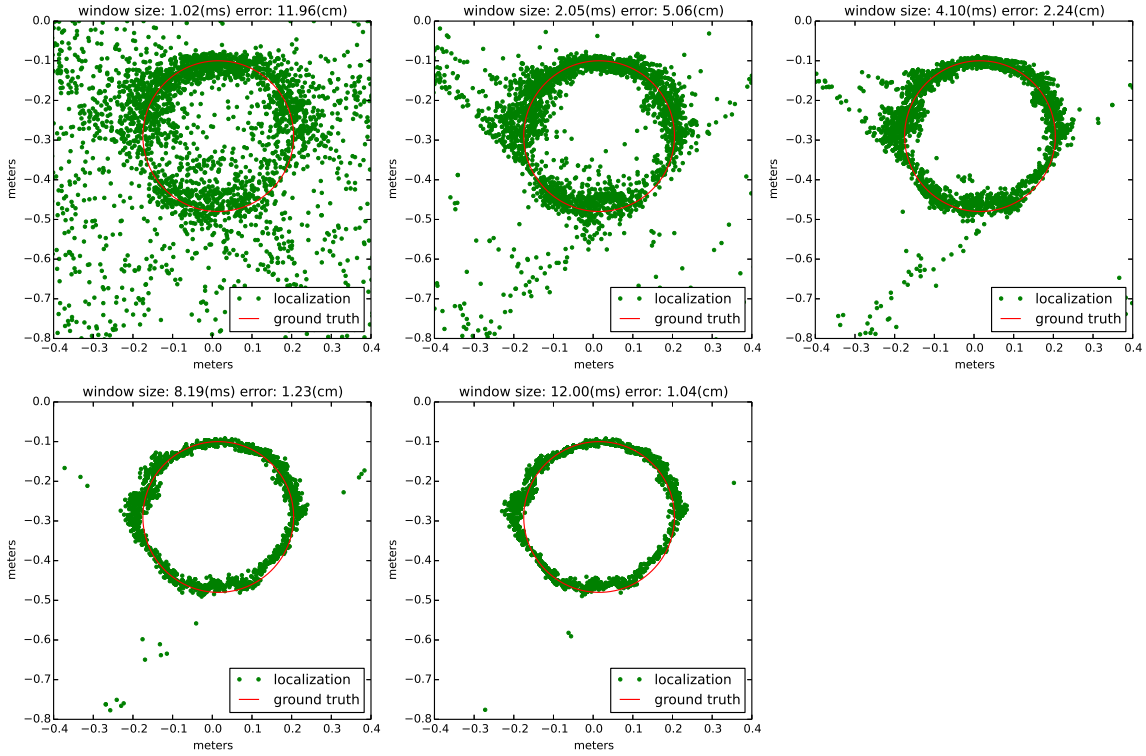


Fig. 6: Localization quality versus window size

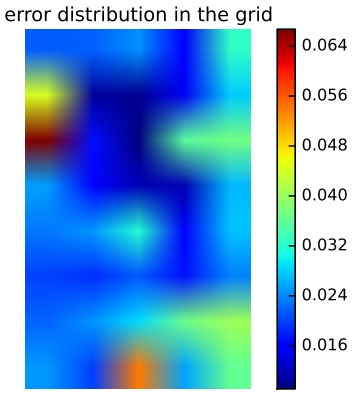


Fig. 4: error distribution in the grid

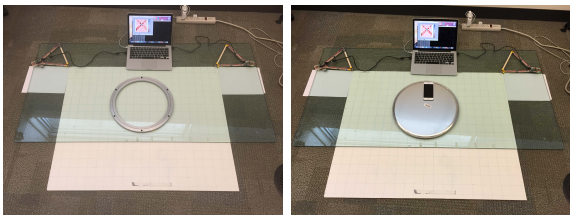


Fig. 5: Setup for circle movement localization

millisecond), the audio does not contain enough information to reliably estimate arrival time differences. The localization is very noisy. As window size increases, the localization converges to the shape of ground truth circle. Fig 7 shows how the error changes with window size. The general trend is similar to that in point localization case. The error decreases as window size increases and plateaus after window size exceeds around 10 milliseconds.

To test how different sound sources impact localization quality, three experiments on the same movement track were conducted with three different sound sources:

White Noise A recording of white noise.
Music A A music that has normal audio amplitude throughout experiment period was chosen. ”Honest Eyes” by Black Tide was chosen
Music B A music with intermittent low amplitude sections was chosen. ”Canon” was used.

To test how audio movement source affects localization quality, each of three experiments were conducted at two different speeds:

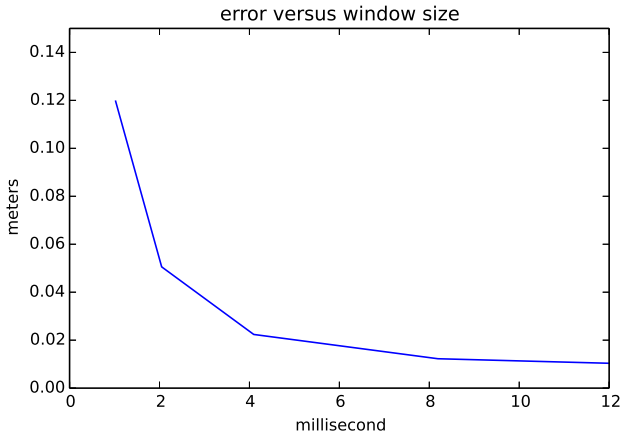


Fig. 7: Localization error versus window size

Normal An angular speed of 0.5 rad/s was maintained, which translates to linear speed of 10 cm/s.

Fast An angular speed of 1.0 rad/s was maintained, which translates to linear speed of 20 cm/s.

For each experiment conducted, two different movement filters are tests:

Averaging filter Location estimation for past 0.5 seconds are averaged and outputed as current estimate.

Kalman filter A 2nd order Kalman filtering is used.

Fig 8 shows results for experiments with three different audio sources at normal speed.

with three different audio source were used for the same circular movement. In the first trial, white noise is used to set a baseline for audio sources, since our understanding is that white noise performs the best under GCC_PHAT. In the second trial, a randomly picked music is used. In the third trial, a music that contains low amplitude segments is picked to reflect real life situations. Fig **??** shows the result for all three trials. From the result, white noise does perform the best and the localized shape matches well with the shape of the rotating disk. There exists a bit of jiggling, but can be improved with filtering. Fig **??** also shows the result with averaging filtering and Kalman filtering.

For the music A, the array still localizaes correctly and the output represetns a circular shape, but the localizaed points are more spread out with more jiggling. This is because music does not have components in all frequency bands and GCC_PHAT estimate is less accuracy with such audio sources. Similar to the white noise case, we can improve the raw tracking estimates from the array with filtering.

For music B, the music has a low amplitude section and it is also shown in the traced result. The ”blank” region in the output corresponds to the low amplitude segment in the audio. The general tracking quality of this music is not as good as that with Music A due to these low amplitude audio segments.

III. Conclusion

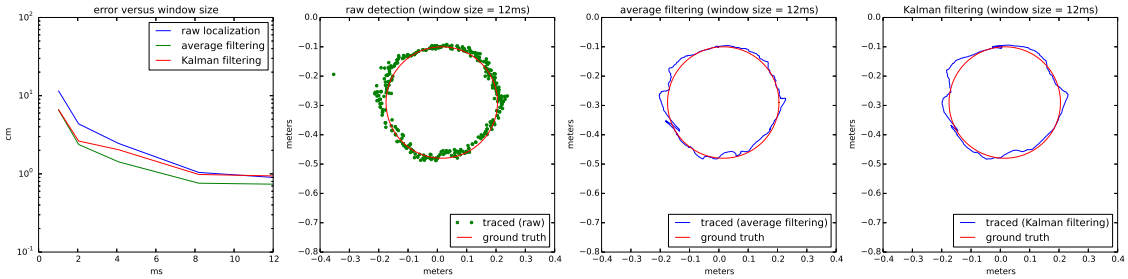
The conclusion goes here.

Acknowledgment

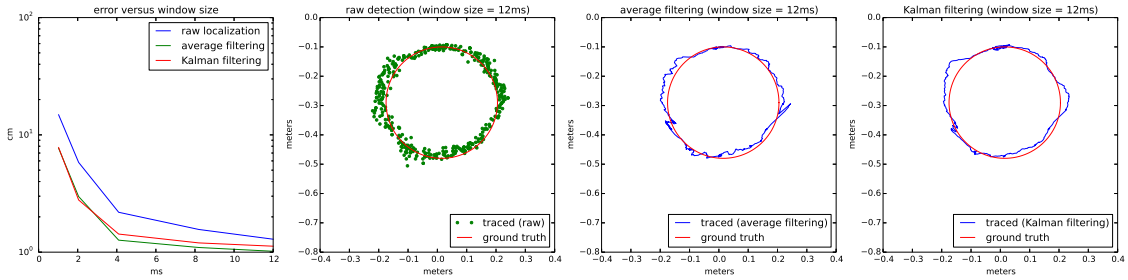
The authors would like to thank...

References

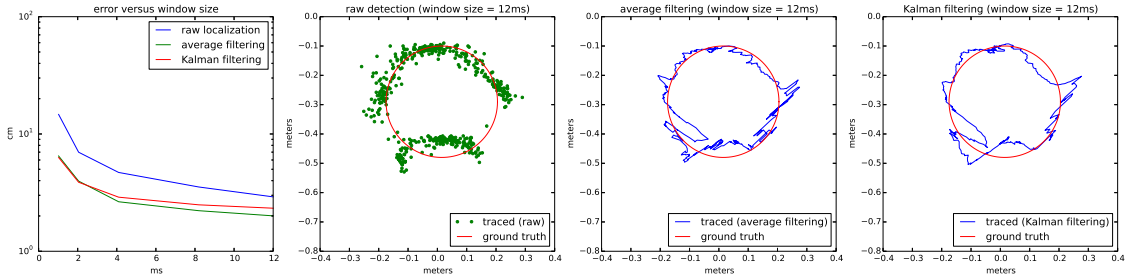
[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.



(a) white noise

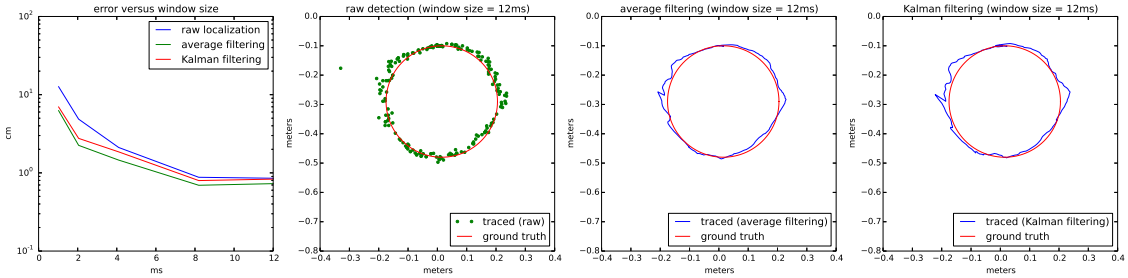


(b) music A

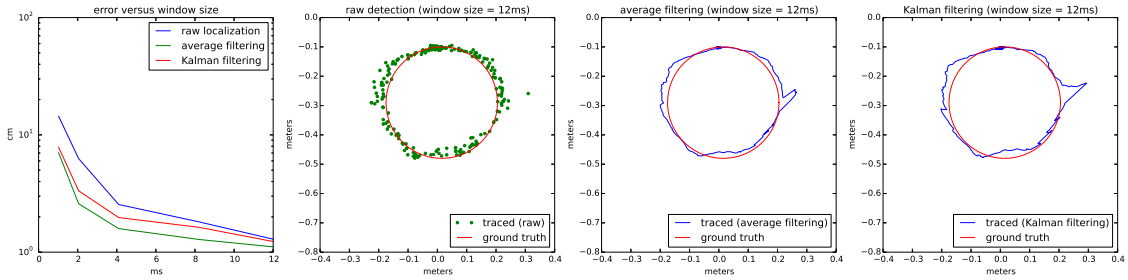


(c) music B

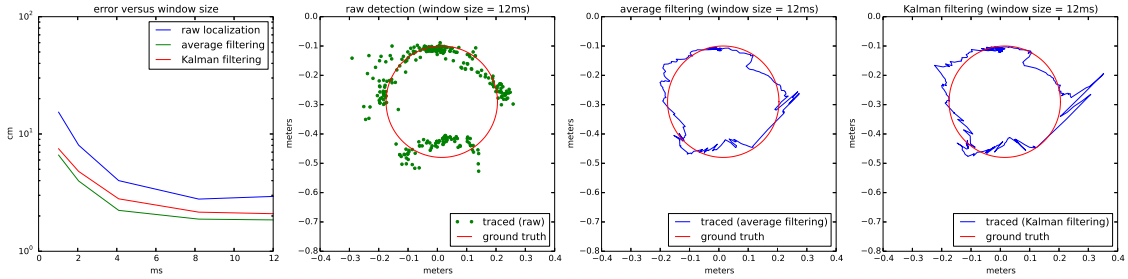
Fig. 8: Localization of circle movement with different sound sources. Sound source is moving at 10 cm per second



(a) white noise



(b) music A



(c) music B

Fig. 9: Localization of circle movement with different sound sources. Sound source is moving at 20 cm per second