OCT Image Extraction Project Report

Jeffrey Wong

March 25, 2025

1 Guiding Computations

1.1 Lateral Resolution

The lateral resolution of the system is primarily limited by the diffraction limit imposed by the detector system. Numerically, the lateral resolution is $\Delta x = 0.37\frac{\lambda_0}{NA} = 0.37(\frac{1.3\mu m}{0.055}) = 8.75\ \mu m$.

1.2 Axial Resolution

The axial resolution is limited by the (Fourier Transform of the) spectrum of the source, which acts as a PSF on the locations of the sources. Based on the parameters given, the axial resolution should be $\Delta z = \frac{2ln2}{\pi}\frac{\lambda_0^2}{\Delta\lambda} = \frac{2ln2}{\pi}\frac{(1.3\mu m)^2}{0.1\mu m} = 7.45\ \mu m$.

1.3 B-Scan Aspect Ratio

We note that the B-scan has 10,175 total samples and 175 background samples, meaning 10,000 of the samples correspond to actual A-scans. The overall B-scan is 1 mm wide, so the lateral spacing (and thus the lateral pixel spacing dx) of the A-scans is $\frac{10^{-3}m}{10^4} = 10^{-7}m = 100\ nm$. Since the image will end up being 10000 x 2048 without accounting for the mirror image effect, we get a pixel aspect ratio of 1 : 36 and an overall size of 1 mm x 7.37 mm or an aspect ratio of 1 : 7.4. Accounting for mirror image artifacts halves the overall axial length giving us an overall size of 1 mm x 3.69 mm or an aspect ratio of 1 : 3.7.

2 A-Scan

2.1 The A-Scan Function and Time

The A-Scan function utilized largely follows the standard OCT pipeline with a few steps shuffled around to avoid redundancy in storing values (notably doing background subtraction before smoothing because we need the raw average background for subtraction). The function accepts a data array or matrix (which is treated as just a series of arrays), a background matrix, and three logical arguments, which determine whether to display the time elapsed on four major steps, whether to perform deconvolution, and whether to perform background subtraction. The arguments block is used to give default values of false for verbosity (we'll only display it once to analyze the performance in this section), and true for enabling deconvolution and background subtraction. The amount of time it takes to run the function is largely dependent on how many slices we pass in to the background and data arguments, which largely affect the background processing and FFT stages, respectively. The windowing and deconvolution also scale with the amount of data, but generally take much less time to perform than the FFT as for N scans with M points, windowing and deconvolution are $O(MN)$ while FFT is $O(Mlog(M)N)$. As an example to illustrate where bottlenecks occur, if we run the A-scan function with a single slice, windowing and background processing take the longest at 3.6 and 7 ms, while the FFT averaging step takes 1.1 ms. In comparison, with the entire M-scan most of the time is spent on the FFT/averaging step, which takes 20 seconds where the other steps take at most 1 second. Windowing likely scales very slowly because it takes some time to just generate the Hamming window, and comparatively less time to apply it to each column, while the FFTs that have to be performed at each column are more expensive due to the MlogM factor.

2.2 Plots

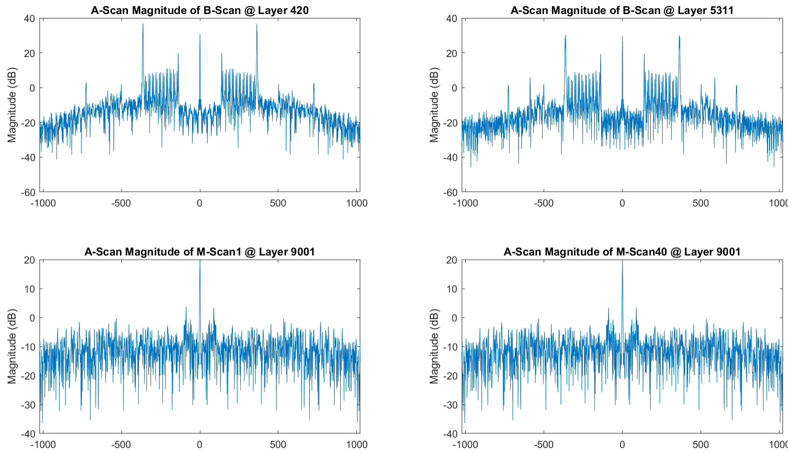

Figure 1: Magnitude Plots of selected A-Scans

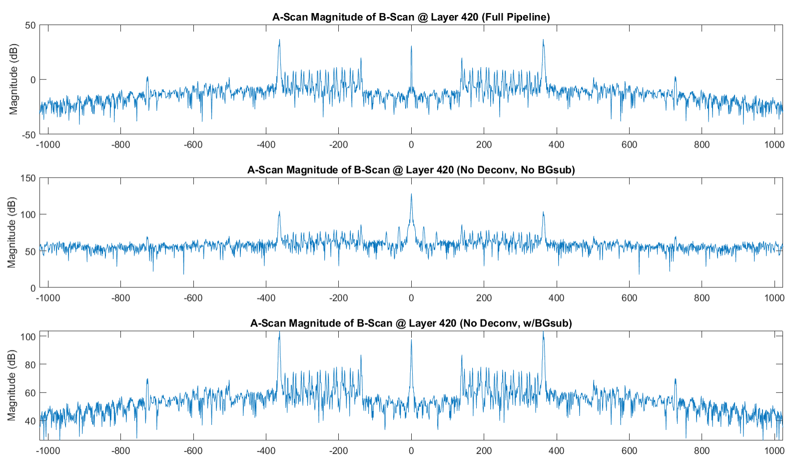

Figure 2: Magnitude Plots of A-Scans with Incomplete Pipelines

2.3 Images with Incomplete Pipelines

For comparison's sake, plots of the A-scan obtained from the B-scan were produced, one without deconvolution or background subtraction, and one without deconvolution but with background subtraction. Without deconvolution, but with background subtraction, it appears to have mostly just shifted the overall amplitude up by 70 dB, which could be explained by the "PSF" associated with the spectrum being particularly narrow and tall in the k-domain, which leads to sharpening in the ζ-domain. By removing the background subtraction, we mostly just observe a larger DC peak amplitude and nothing outside of a narrow band at the center, which makes sense because a lot of the DC component comes from the reflected intensity associated with E_R from the galvo-mirror. Overall, the "texture" of the scans is largely unaffected, only the amplitudes.

3 B-Scan

3.1 Scans

Figure 3: Raw B-scan Image

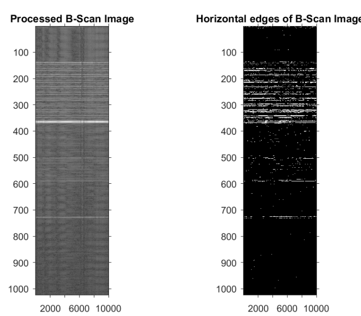

Figure 4: Processed B-Scan Image with horizontal edges highlighted

3.2 Image Processing Pipeline

Note: A discussion of pixel sizes can be found in section 1.3 of the report.

First of all, only a few layers of tape show up in the raw image, so the image had to be gamma corrected with $\gamma = 0.15$ in order to show the rest of the layers. Since the mirror image effect essentially produces a reflected copy of the image, we can take just half of the image as representative of the actual sample, meaning we can crop out the top half of the image. Because the axial and lateral resolutions are much larger than the pixel sizes, the image is already "smoothed out" so applying a smoothing filter such as a mean or Gaussian filter makes no sense. In order to clarify what parts actually represent tape and what is noise, I additionally performed horizontal edge detection on the image using the Sobel method, which highlighted only the upper third of the edges as actually corresponding to layers of tape. Based on close analysis of the image, there appears to be 10 (or 11, if you count the bright white band) layers of tape, with each layer of tape appearing to be $7\frac{1}{2}$ pixels or 27 microns thick, and the layers appearing to be separated by 14 pixels or 50 microns.

3.3 Scans without Deconvolution

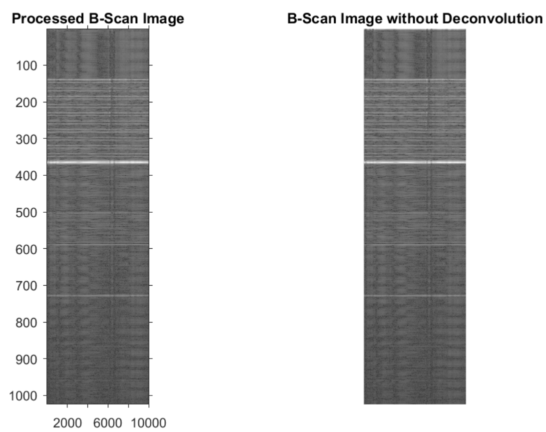

Figure 5: Processed B-Scan image with and without deconvolution

Skipping the deconvolution step in the pipeline seems to have no meaningful effect on the final processed image, other than making the noisy regions slightly brighter. This is disappointing, but not necessarily surprising, considering that removing deconvolution in the individual A-scans only shifted the overall amplitude up, and since all of the images are being normalized by the highest amplitude over all the scans, it makes sense that the relative magnitudes are the same and thus that the images appear to be the same with and without deconvolution.

4 M-Scan

4.1 1-Tone M-Scan

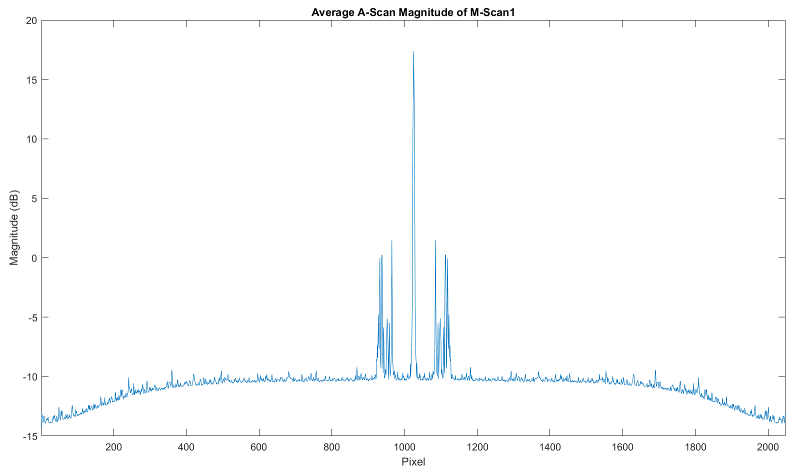

Figure 6: Magnitude Plots of average of A-Scan from MScan1

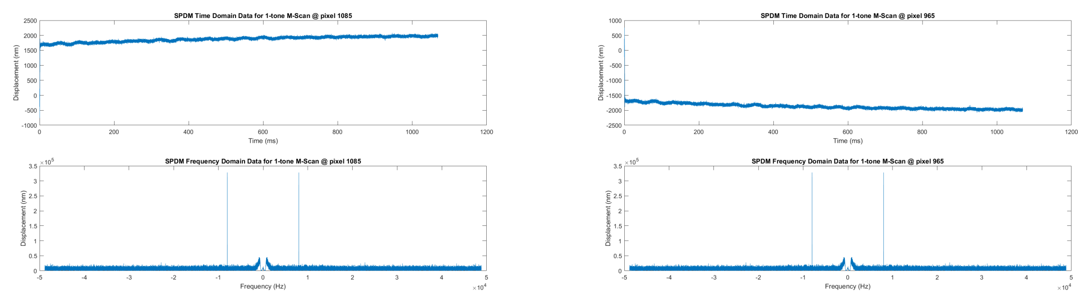

Figure 7: SPDM Time and Frequency Domain Data for MScan1

The A-scan magnitude highlighted two pixels that were equidistant to the center. It thus makes sense that performing SPDM at those two pixels produces the same frequency-domain data and time-domain data that differs only by their sign, as those areas should move in opposite directions during the normal use of the speaker. Based on the frequency-domain data, it looks like an 8 kHz tone was played. Note that the time domain data presented is the raw (unwrapped) angle, but the frequency-domain data for both the 1-tone and 40-tone scans is filtered using a 5th order Chebyshev II filter with 30 dB stopband attenuation and a 500 Hz cutoff frequency to eliminate spurious near-DC components to produce the frequency domain data.

4.2 40-Tone M-Scan

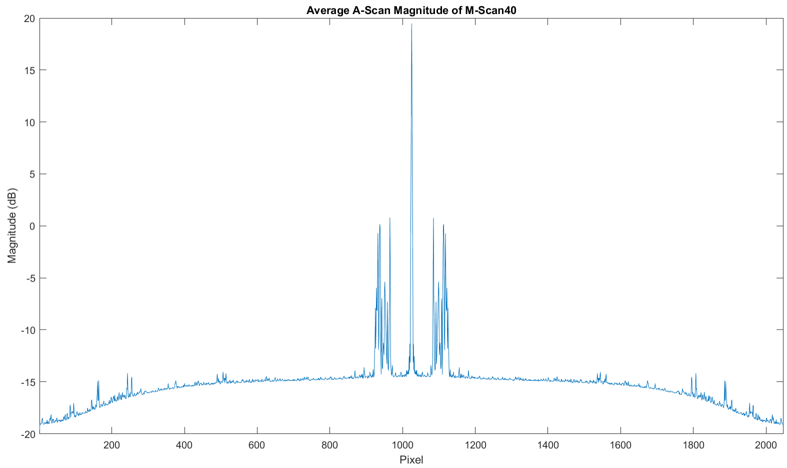

Figure 8: Magnitude Plots of average of A-Scan from MScan40

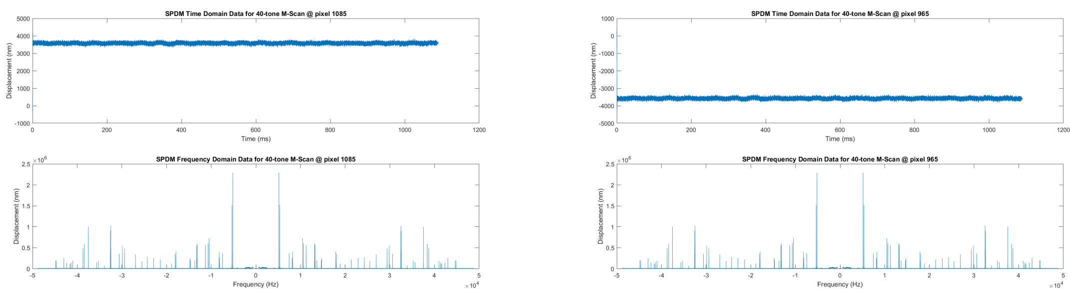

Figure 9: SPDM Time and Frequency Domain Data for MScan40

The `extract_tones` function is used to obtain (positive) frequencies, in kHz, with a frequency response above a certain threshold, which would correspond to frequencies where a tone was played. Of course, it may show a few consecutive nearby frequencies, which could be treated as a single tone rather than multiple. For example, based on the output of `extract_tones`, we can observe tones at 5.16 kHz, 5.35 kHz, 10.45 kHz, and 32.47 kHz, among many others, with 5.16 kHz being the lowest observed tone and 44.8 kHz being the highest observed tone. We note that since the frequency response is not identical for all tones even though all tones were inputted at the same amplitude, it is clearly not all-pass. However, using the `extract_tones`, and noting that the lowest frequency response of a tone is about 78000 (arbitrary units), with a threshold of 70000 we find exactly 40 tones, which aligns with superposition of responses and thus would suggest linear behavior. However, to confirm linearity we would need to observe responses at other SPLs to determine if the system response observes homogeneity.