1 Signal-to-Noise Ratio Before and After MCRA

As a quick update from the last report, the SNR for each type of call [stemming from STE-063] before and after the MCRA (one iteration) is as follows:

Table 1: SNR Improvement via MCRA (1 Iteration)

Call	**Original SNR (dB)**	**SNR with MCRA (1 Iteration) (dB)**	**SNR Improvement (dB)**
Combo	-24.645578	-24.590096	0.055482
No Voice	-28.977140	-26.805875	2.171265
Peep	-24.340737	-24.272304	0.068433
Phee	-20.642118	-20.583845	0.058273
Trill	-22.023286	-22.102499	0.092939
Tsik	-32.165536	-32.159946	0.005590
Twitter	-28.685297	-28.652038	0.033259

Table 2: SNR Improvement via MCRA (2 Iterations)

Call	**Original SNR (dB)**	**SNR with MCRA (2 Iterations) (dB)**	**SNR Improvement (dB)**
Combo	-24.645578	-24.584944	0.060634
No Voice	-28.977140	-26.795295	2.181845
Peep	-24.340737	-24.271530	0.069207
Phee	-20.642118	-20.567935	0.074183
Trill	-22.023286	-22.116225	0.092939
Tsik	-32.165536	-32.156996	0.008540
Twitter	-28.685297	-28.646333	0.038964

Corresponding to the above tables, Figures 1-3 show the SNR via the spectrograms of phee, tsik, and twitter calls.

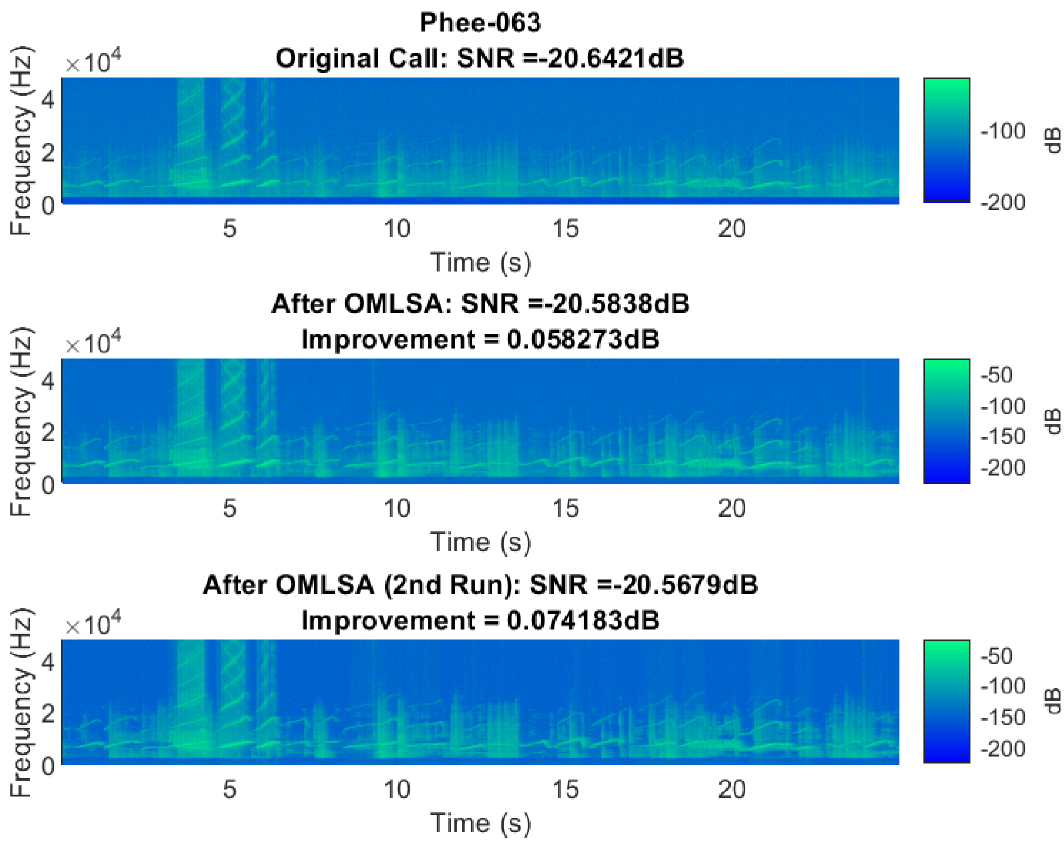


Figure 1: Phee call SNR Improvement via MCRA

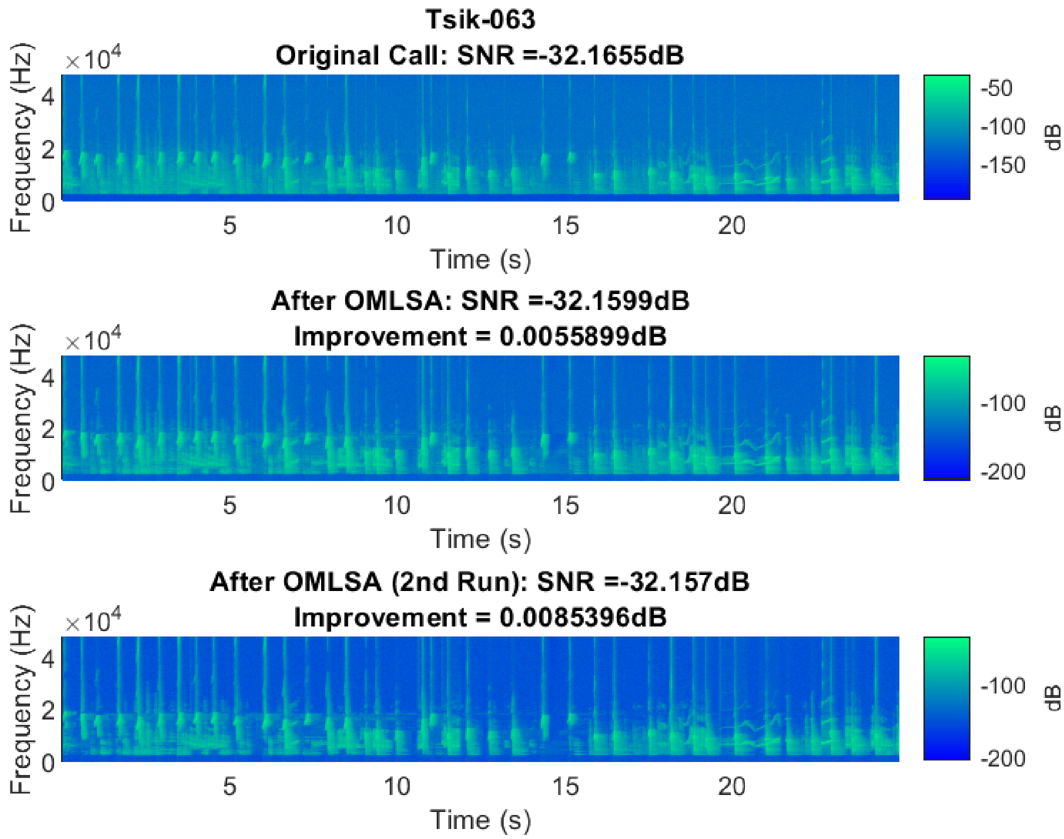


Figure 2: Tsik call SNR Improvement via MCRA

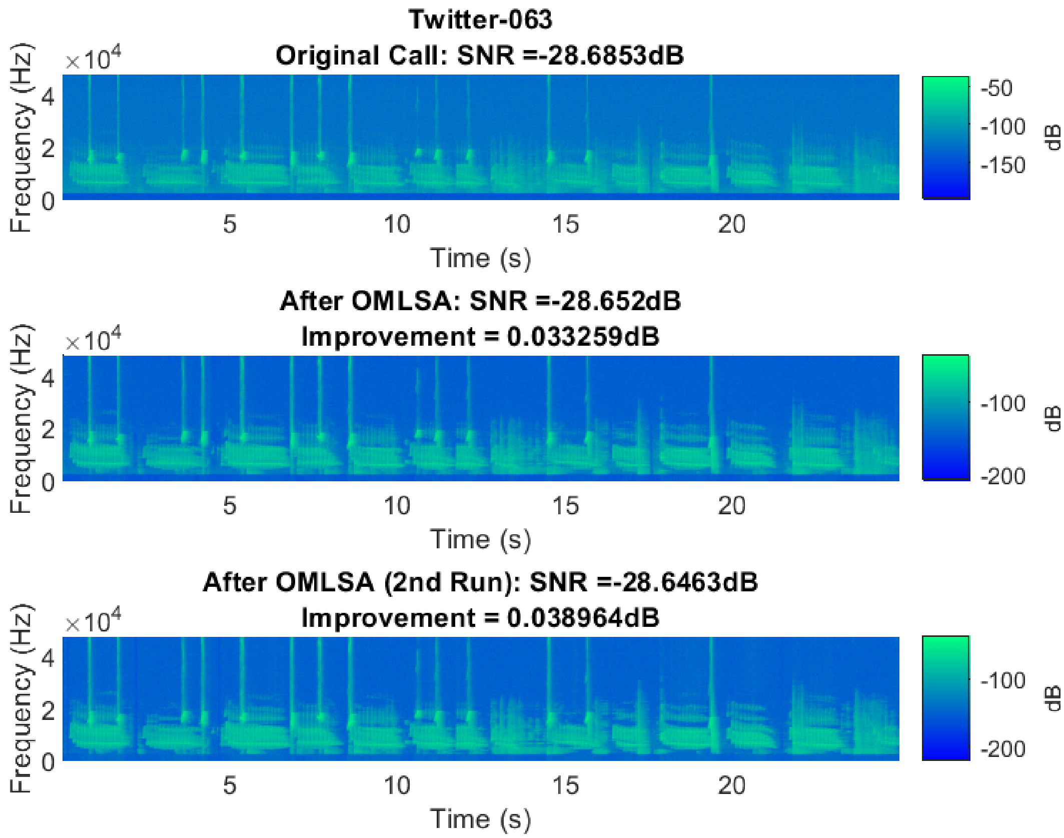
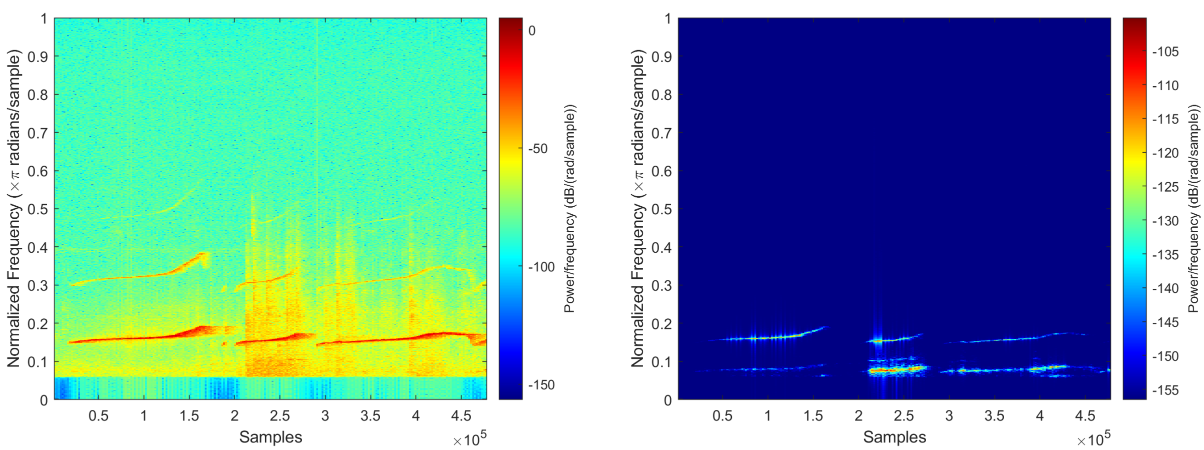


Figure 3: Twitter call SNR Improvement via MCRA

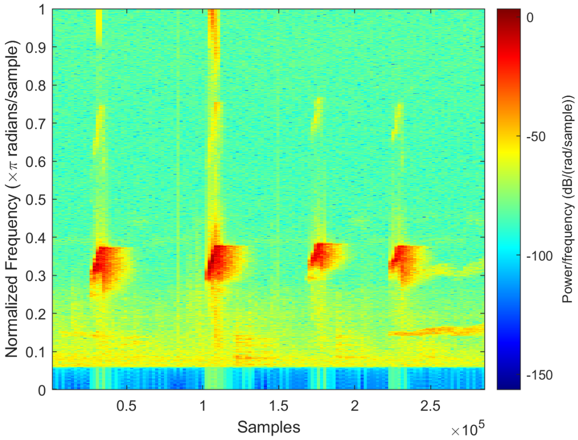
2 Harmonic Product Spectrum

A large focus in the past couple of weeks has been on working on extracting the fundamental frequencies of the calls using the harmonic product spectrum (HPS). This pursuit has seen mixed results. Theoretically, the Harmonic Product Spectrum is ideal for signals with high harmonicity such as the ones the team is dealing with, but there is no uniform number of harmonics that works for all the calls. This corresponds to the number of decimations or products of the spectrum and while for one call, 3 harmonics may produce good results, another call with weak 3rd harmonics may experience a decrease in SNR overall. One approach that is currently being investigated is to merely preserve the estimated fundamental frequency of each frame, while discarding the remaining data. One unintended artifact observed while using this algorithm is a smearing effect in the spectrogram (see Figures 4-6 below). While this may reduce distinguishability for call vs. call classification, it may very well improve the marmoset call vs. no call classification.

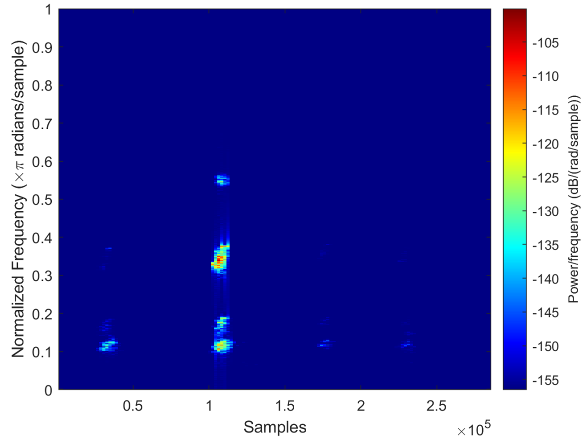


(a) Base Spectrum (b) Harmonic Product Spectrum

Figure 4: Phee call with no MCRA clean up

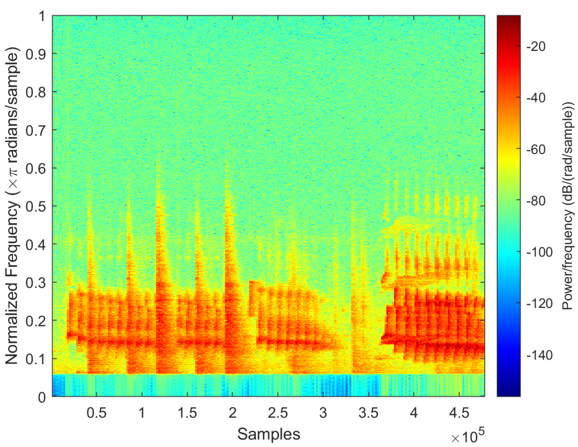


(a) Base Spectrum

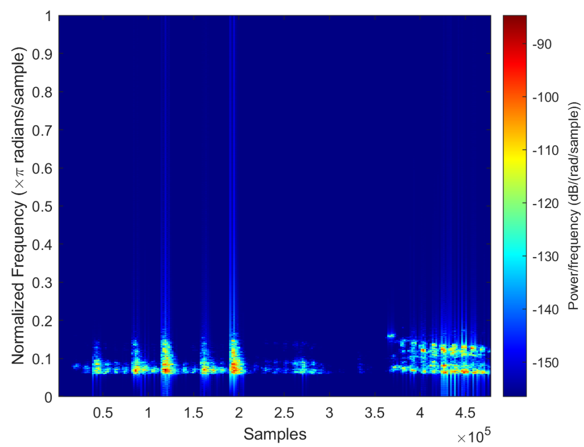


(b) Harmonic Product Spectrum

Figure 5: Tsik call with no MCRA clean up



(a) Base Spectrum



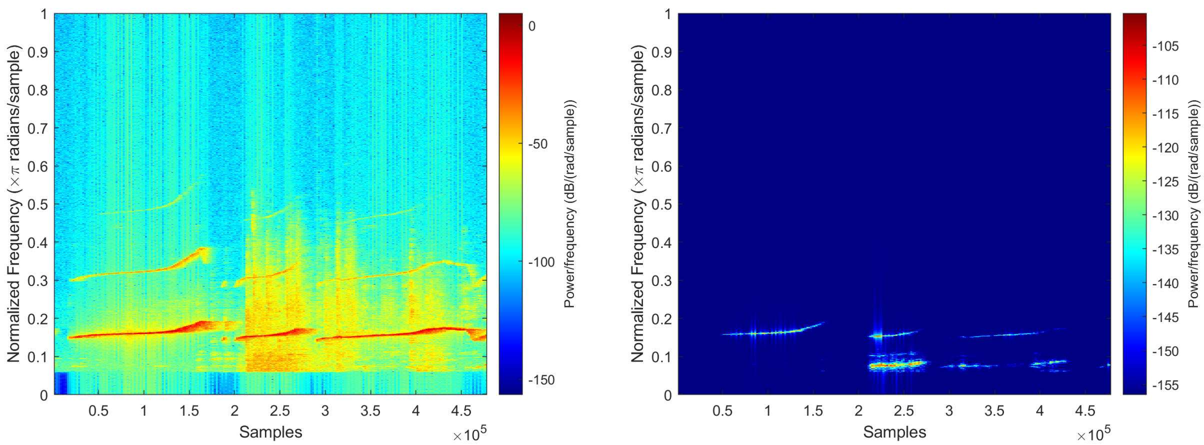
(b) Harmonic Product Spectrum

Figure 6: Twitter call with no MCRA clean up

3 Harmonic Product Spectrum with MCRA

In an effort to clean up the HPS results, as well as simply clean up the spectrograms in general, the MCRA algorithm from Dr. Israel Cohen [1] (as noted in a previous email) was used for both one and two iterations. Since this will serve as a minor step in the team's overall goal, it has been decided that with credits, the team will use this open-source code [for only noise reduction]. While the team is confident that a lightweight version of this algorithm could be written, it was decided that the time commitment for a much weaker version of this robust algorithm is not seen as a priority for this semester. In Figures 7-9, one iteration of the MCRA algorithm can be seen for the same calls shown before, and in Figures 10-12, a second iteration.

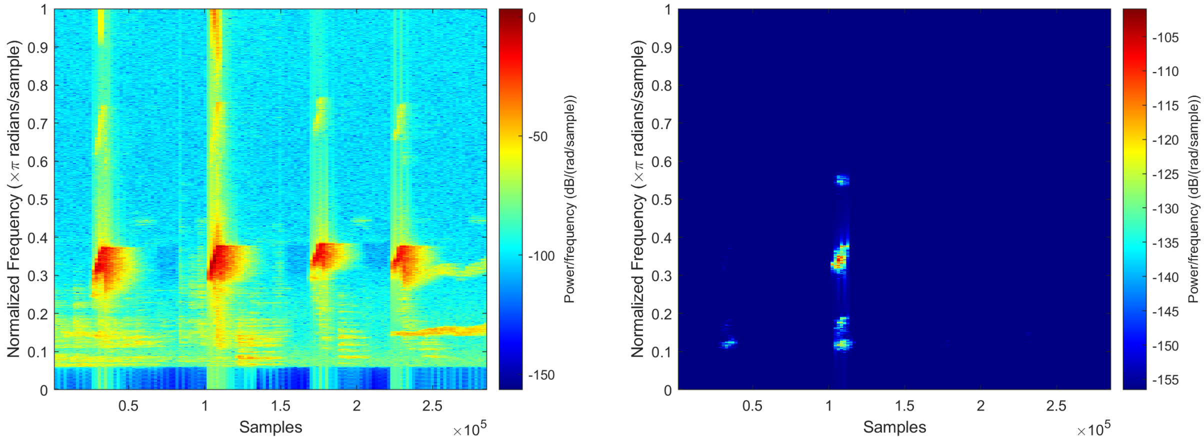
Note: It is clear that one iteration makes for a large improvement in terms of noise reduction and two iterations has benefit, but the reduction is marginal in comparison.



(a) Base Spectrum (b) Harmonic Product Spectrum

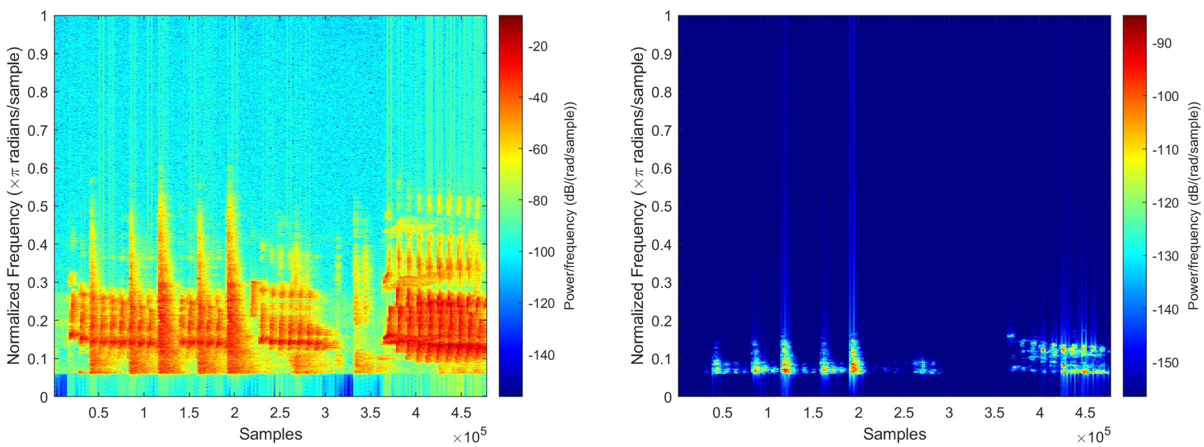
Figure 7: Phee call with one iteration of MCRA

[1]The open-source OMSLA code for MCRA can be found at https://israelcohen.com/software/



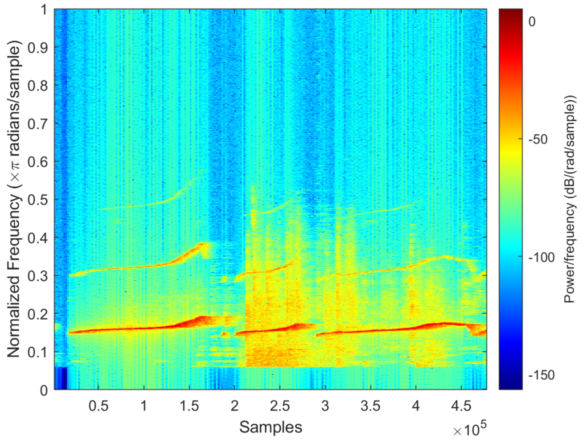
(a) Base Spectrum (b) Harmonic Product Spectrum

Figure 8: Tsik call with one iteration of MCRA

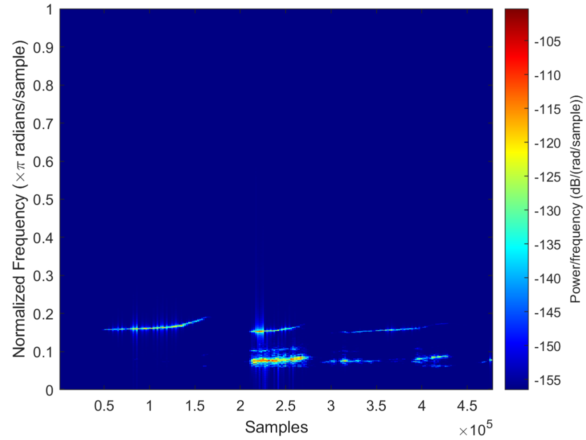


(a) Base Spectrum (b) Harmonic Product Spectrum

Figure 9: Twitter call with one iteration of MCRA

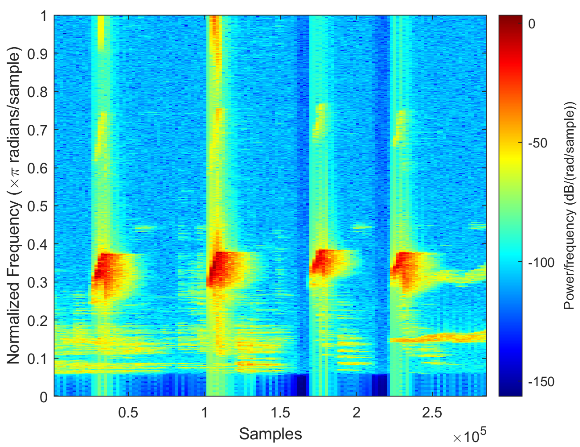


(a) Base Spectrum

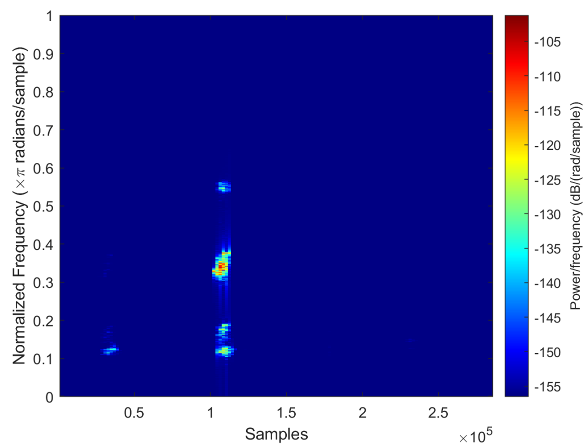


(b) Harmonic Product Spectrum

Figure 10: Phee call with two iterations of MCRA

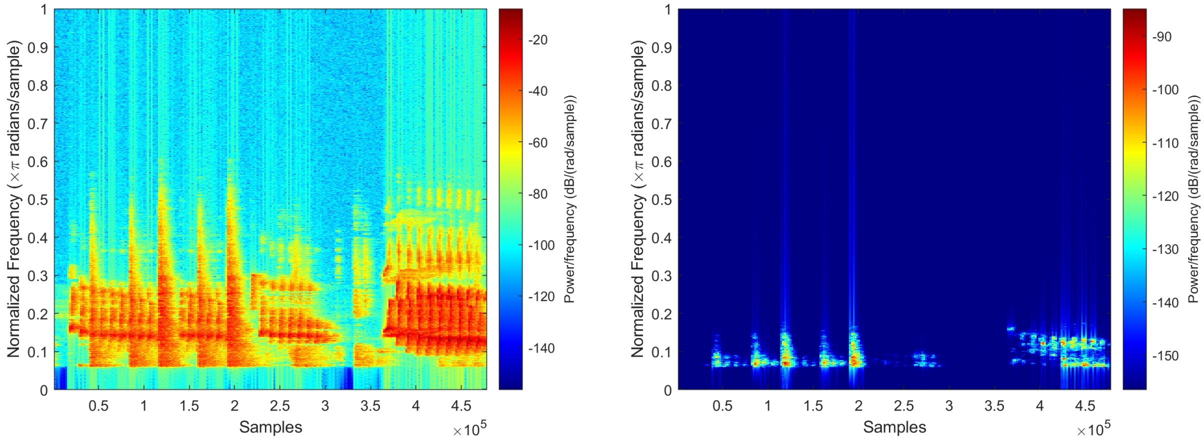


(a) Base Spectrum



(b) Harmonic Product Spectrum

Figure 11: Tsik call with two iterations of MCRA



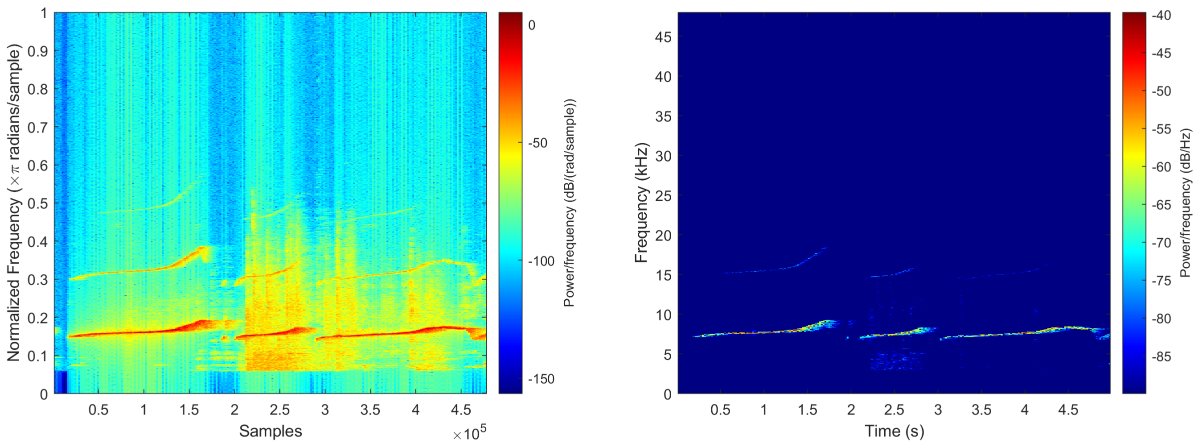
(a) Base Spectrum

(b) Harmonic Product Spectrum

Figure 12: Twitter call with two iterations of MCRA

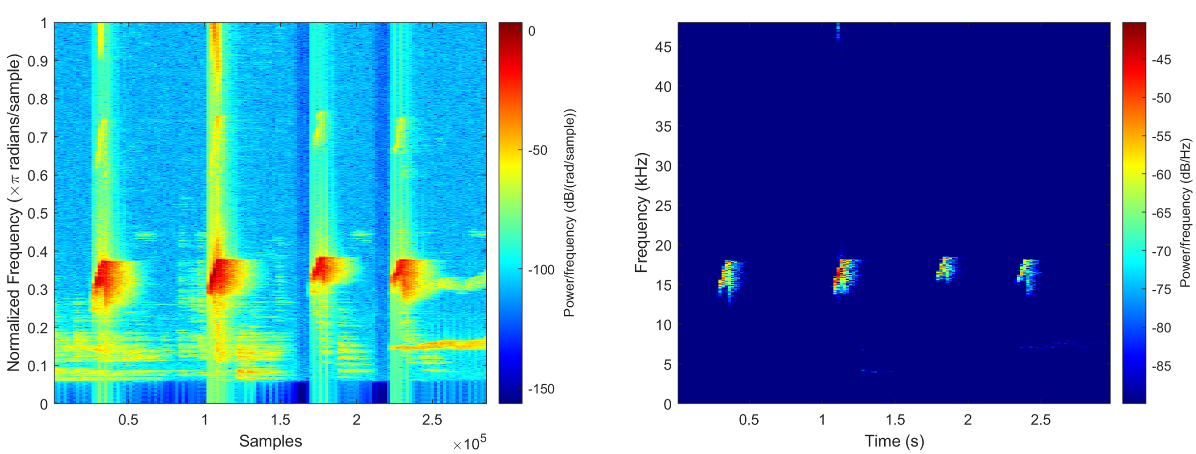
4 Thresholded Spectrum

What the team looked at here was simply setting an energy threshold on the spectrum, below which all data was discarded. While thresholding is far from a good solution for a problem like this, we think that eliminating the very low energies can act as a distinguishing feature that can be fed into the classifiers. The spectrograms resulting from this approach show great promise, and analysis will be done as detailed below to measure the effectiveness of the thresholded spectrum as a feature. Figures 13-15 show this reassignment and dB thresholding with two MCRA cycles.



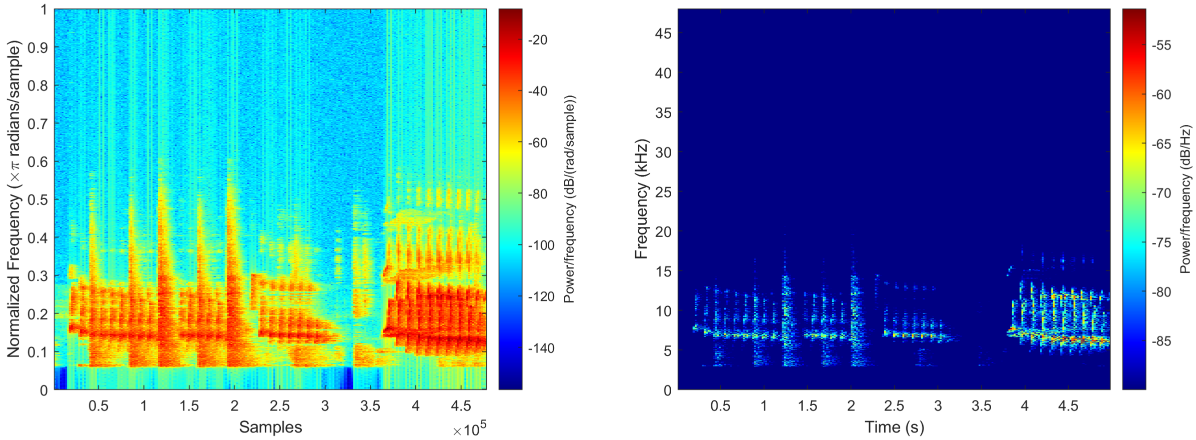
(a) Base Spectrum (b) Reassignment and dB Thresholding

Figure 13: Phee call with two iterations of MCRA - Reassignment and dB Thresholding



(a) Base Spectrum (b) Reassignment and dB Thresholding

Figure 14: Tsik call with two iterations of MCRA - Reassignment and dB Thresholding



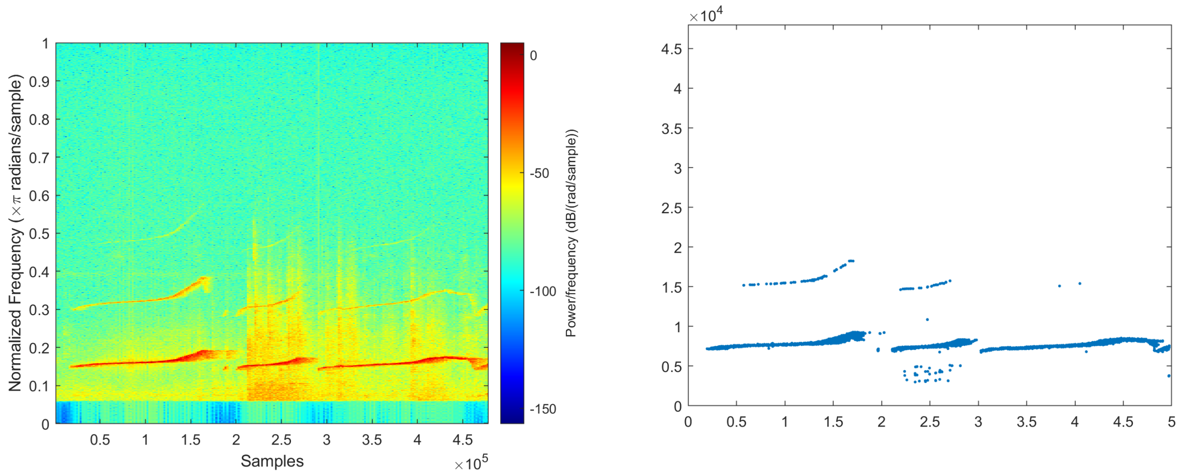
(a) Base Spectrum

(b) Reassignment and dB Thresholding

Figure 15: Twitter call with two iterations of MCRA - Reassignment and dB Thresholding

5 Extraction of Thresholded Spectrum

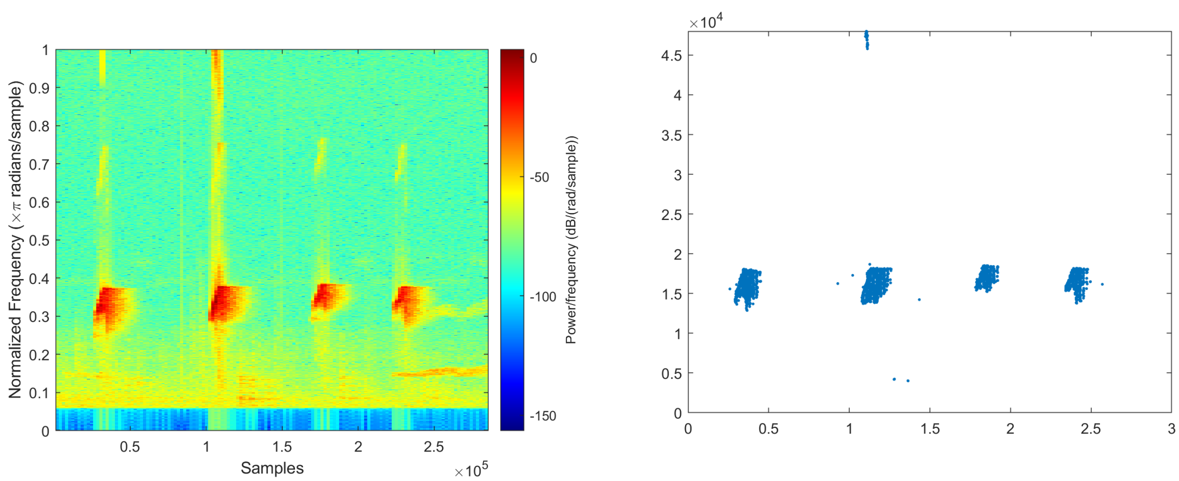
After thresholding the spectrum, another visual representation and promising feature for classification is the extracted data from the thresholded spectrum. This is basically a representation of all non-zero spectrum values left after thresholding, with equal magnitude representation. This provides a visually distinct representation of the data, and opens the possibility for the use of CNNs. Results are quite promising and can be seen in Figures 16-18 below (each call has been run through the MCRA algorithm twice here).



(a) Base Spectrum

(b) Reassignment and dB Thresholding Extraction Points

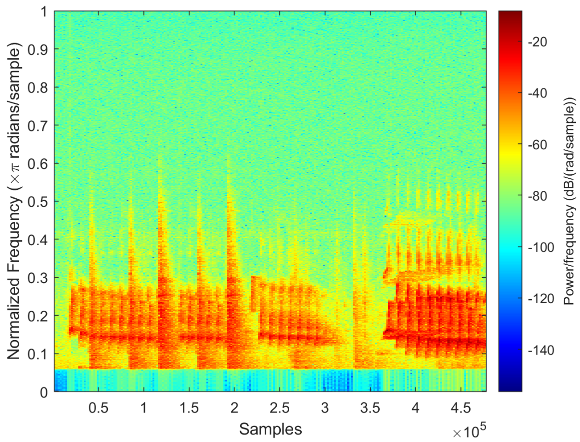
Figure 16: Phee call with two iterations of MCRA - Extraction



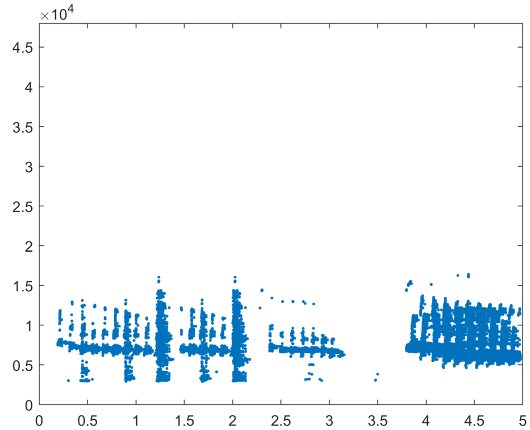
(a) Base Spectrum

(b) Reassignment and dB Thresholding Extraction Points

Figure 17: Tsik call with two iterations of MCRA - Extraction



(a) Base Spectrum



(b) Reassignment and dB Thresholding Extraction Points

Figure 18: Twitter call with two iterations of MCRA - Extraction

6 What's Next

Next steps include extracting the above mentioned features, along with select features that were explored last semester. Specifically the MFCCs, with a revised pre-emphasis. After the features for all the calls have been extracted and normalized, we will perform some statistical analysis and explore the feature correlation and separability. Once this is complete we will train the classifiers using the data and check against the cross-validation sets to evaluate performance on a feature by feature basis.