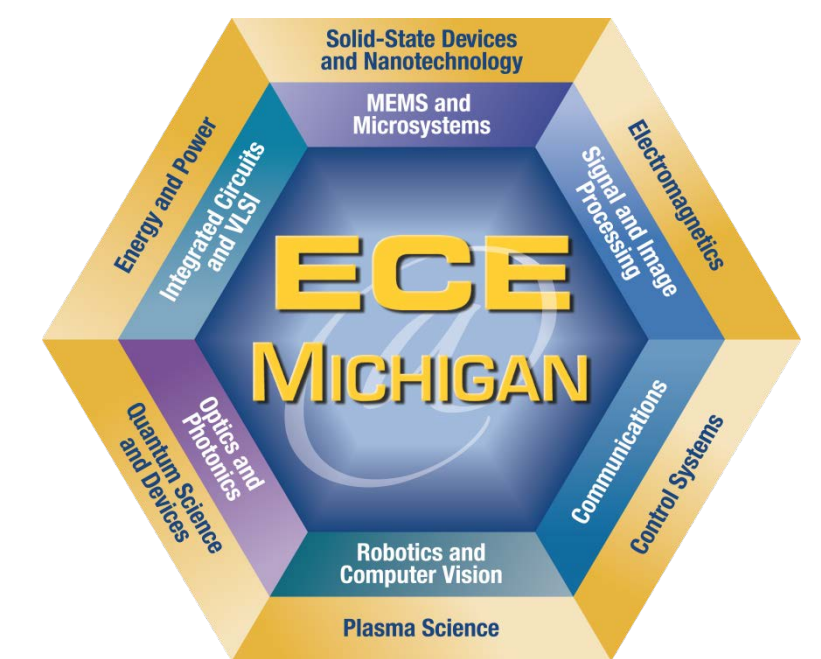# Verax Speech Enhancer

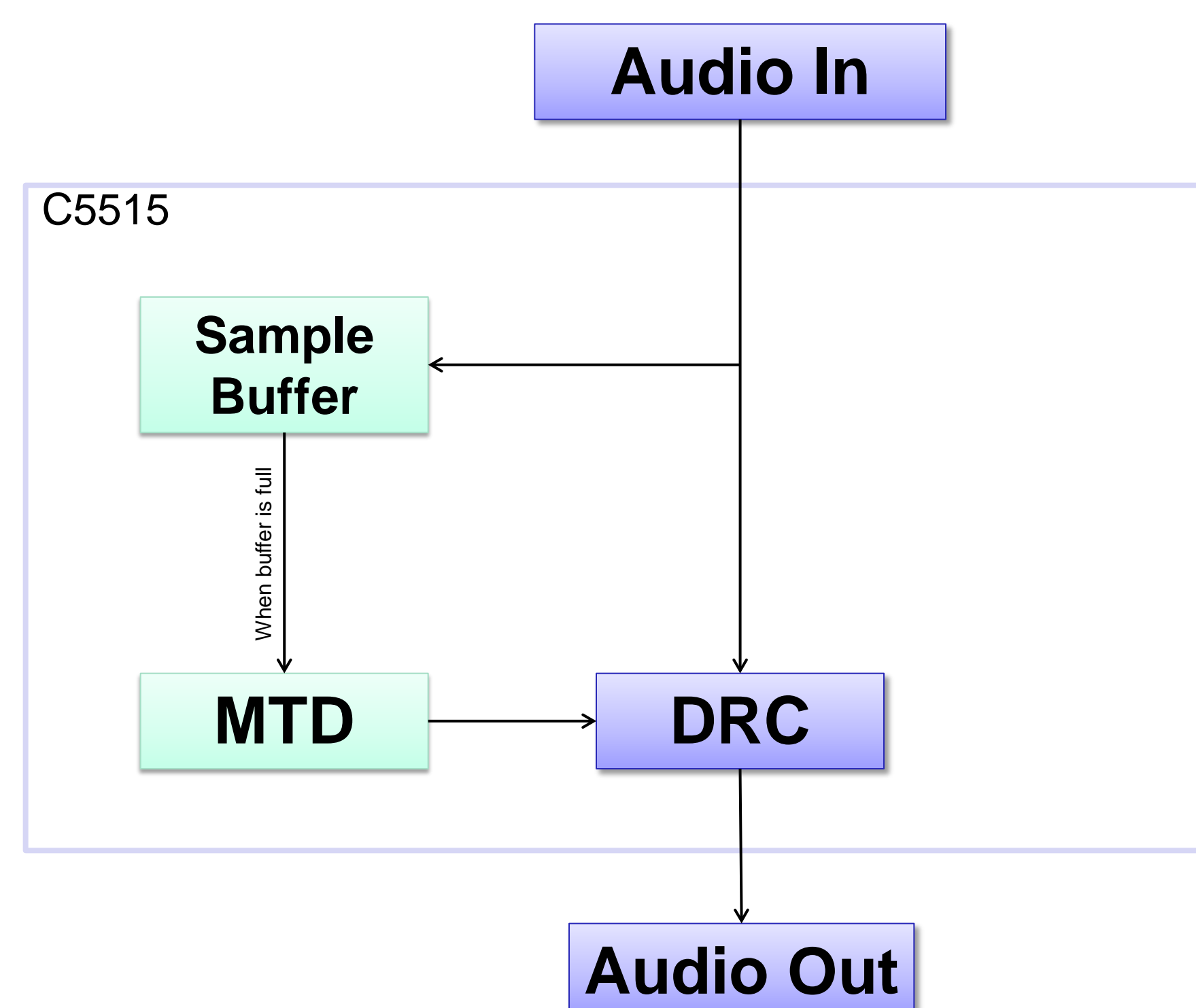## Amisha Agarwal, Vincent Fasburg, Robert Liu, Daniel Miller, Kathleen Walker
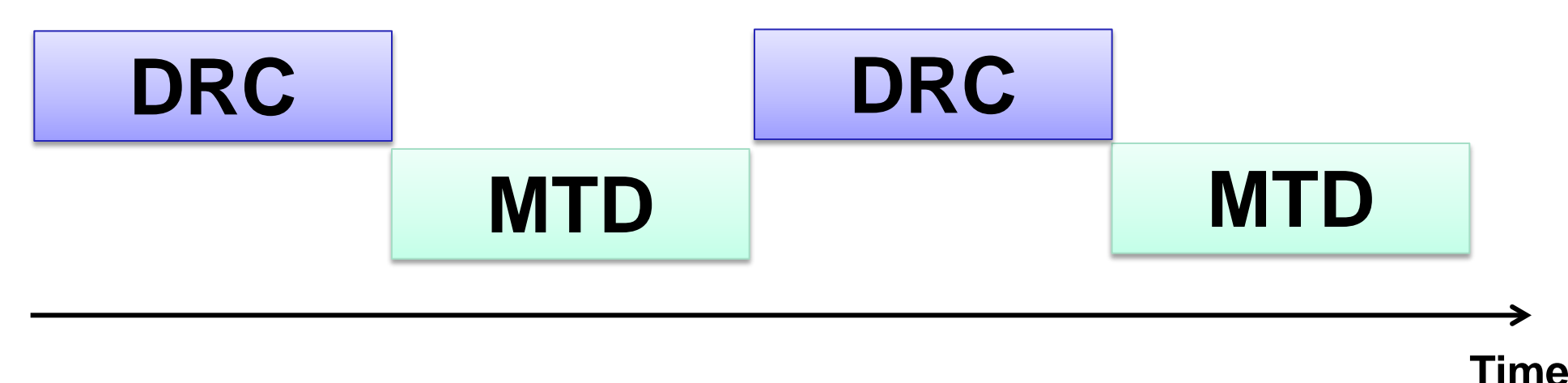
EECS 452 – Winter 2012 Design Expo

Abstract

The Verax Speech Enhancer is an apparatus designed to improve speech intelligibility and balance sound levels in home audio systems. It intercepts the audio signal emitted by a sound source, employs dynamic range compression and media type detection, and outputs the modified audio signal to speakers. This product may be of interest to consumers who wish to avoid disturbing others in the vicinity due to unbalanced sound levels. It may also be of interest to consumers who may have trouble hearing spoken voices in home audio systems due to hearing impairment.

System Diagram

Audio In

C5515

Sample Buffer

When buffer is full

MTD → DRC

Audio Out

Timing Diagram

DRC   DRC

MTD   MTD

Time

Dynamic Range Compression (DRC)

DRC is a method by which the volume of loud portions of an audio signal is reduced without affecting quiet portions of the signal. This is accomplished by taking the average amplitude over a very short portion of the signal and comparing it to a predetermined threshold value. If the signal level is above this threshold, the amplitude of the signal is reduced according to a preset ratio.

In this device, the DRC has been implemented separately on three frequency bands. The signal is separated into low, mid-range, and high frequencies, and DRC is performed on each band, using parameters which are most appropriate for those frequencies. This allows us to independently manipulate bass, vocal, and high pitch sounds to maximize intelligibility of speech and minimize startling, abrupt changes in volume. The parameters used with each frequency band are dependent on the detected media type.
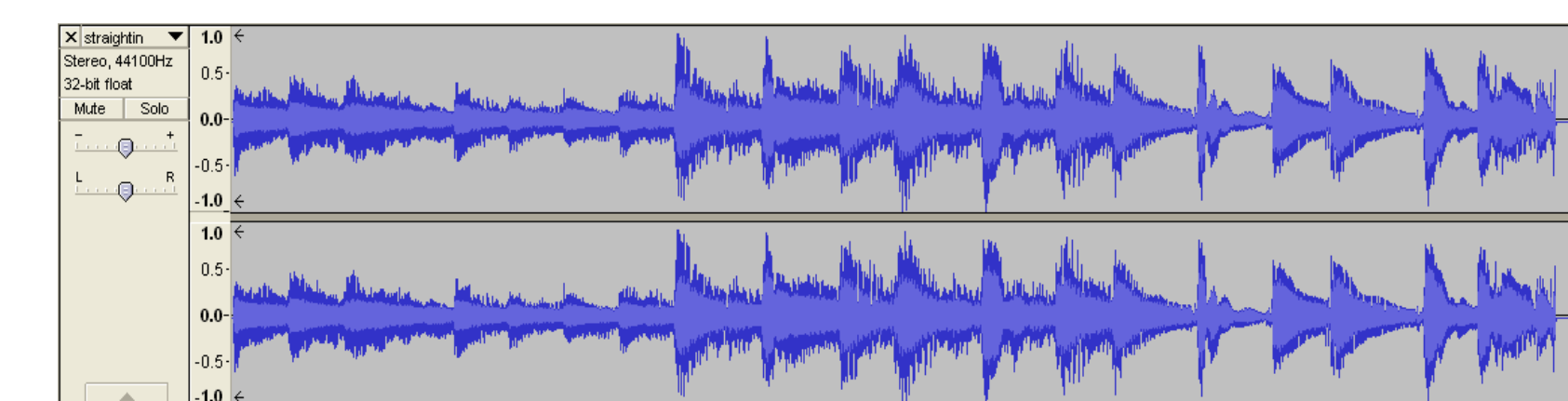
Original Audio Signal
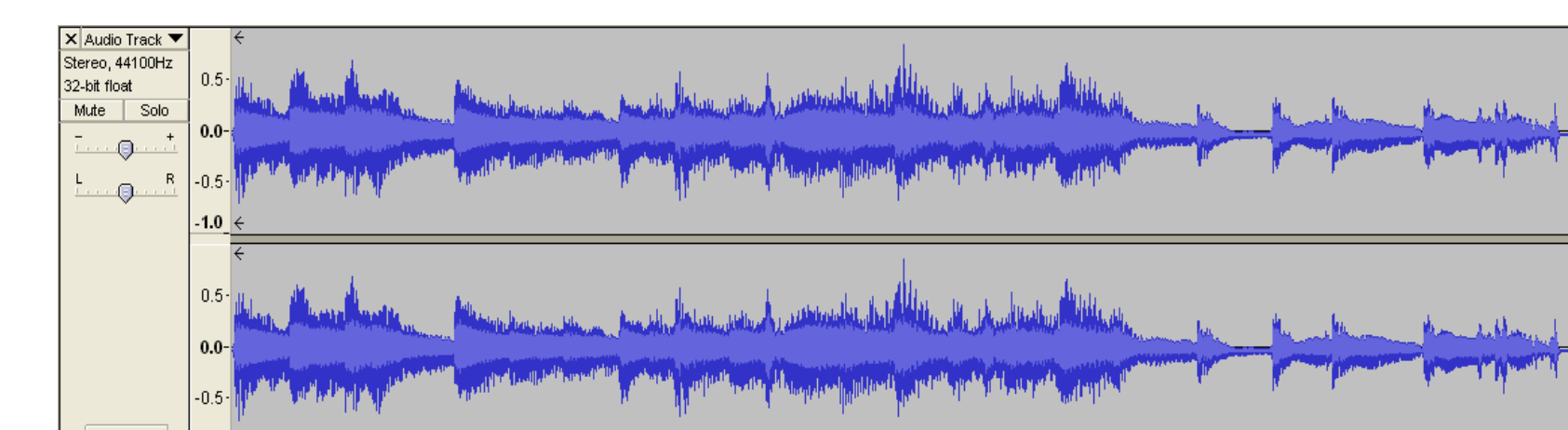
Compressed Audio Signal
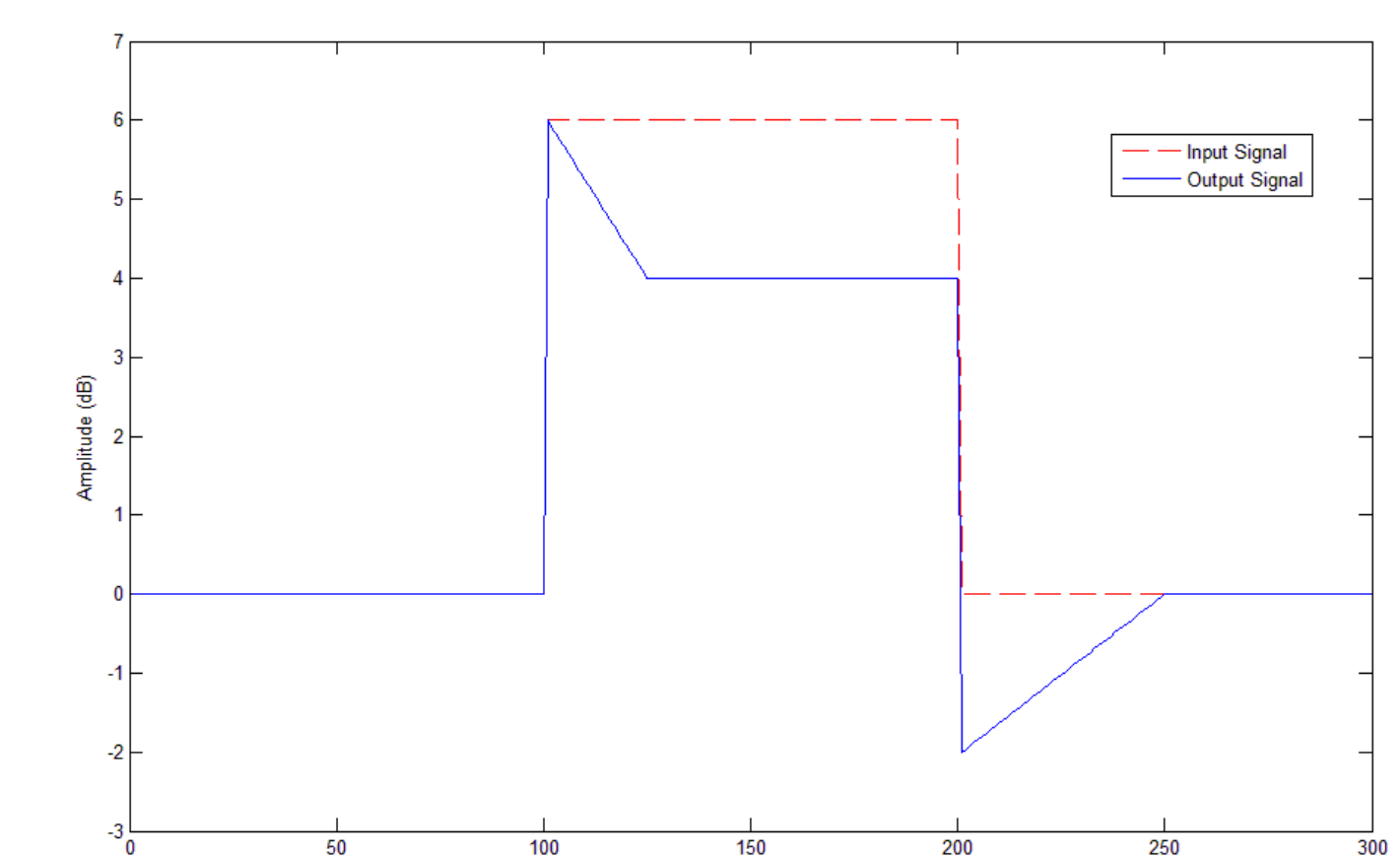
Dynamic Range Compression Illustration

Media Type Detection (MTD)

The incoming audio signal is analyzed and placed into one of three categories: music, vocal, and default. The type is determined by computing the average energy of the entire spectrum and three different frequency ranges (low, mid-range, and high frequency), then comparing these averages. These values are run through a series of comparison tests twenty times; majority logic determines the final media type decision for subsequent DRC as described above.

DRC is implemented in an interrupt routine that runs between rounds of MTD. Every third sample is collected from the incoming signal until a buffer is full, then the frequency spectrum of the sample buffer is obtained using a 2048 point fast Fourier transform and analyzed.
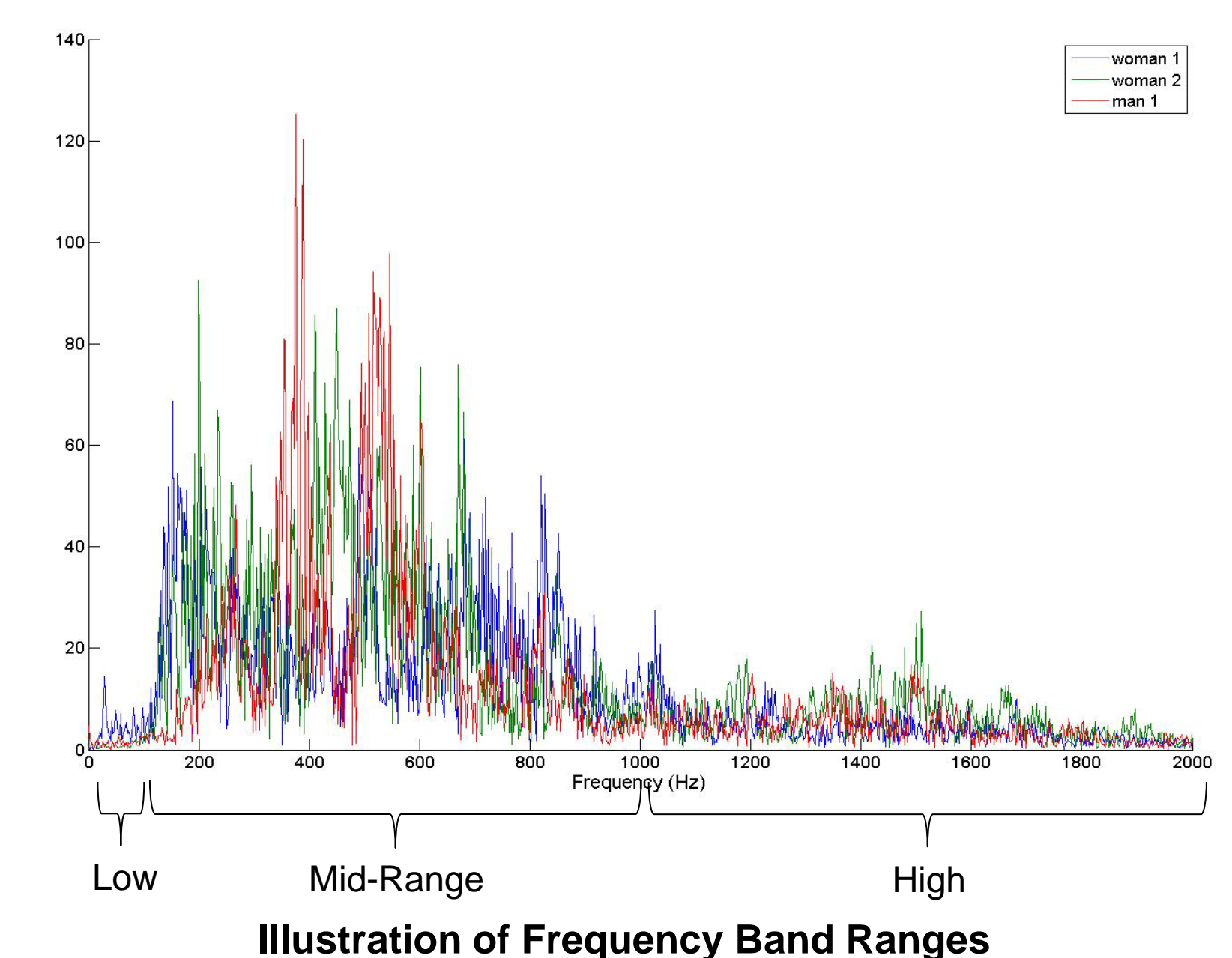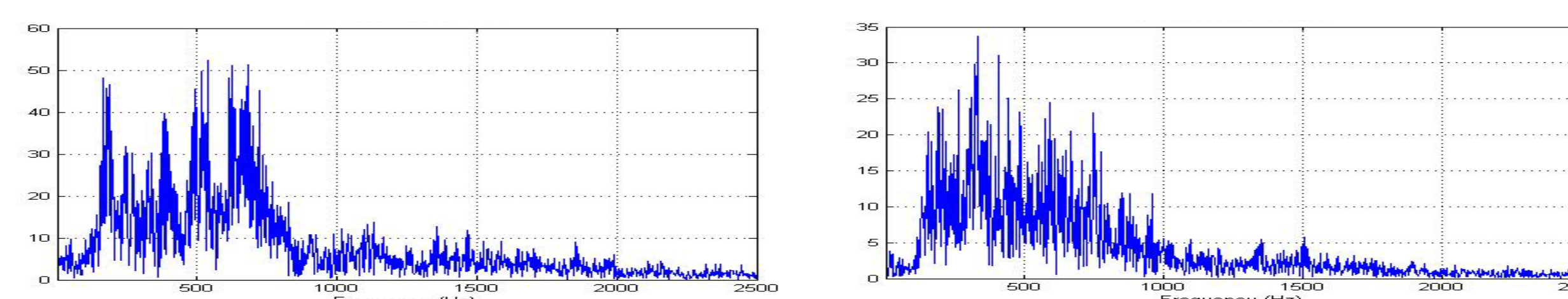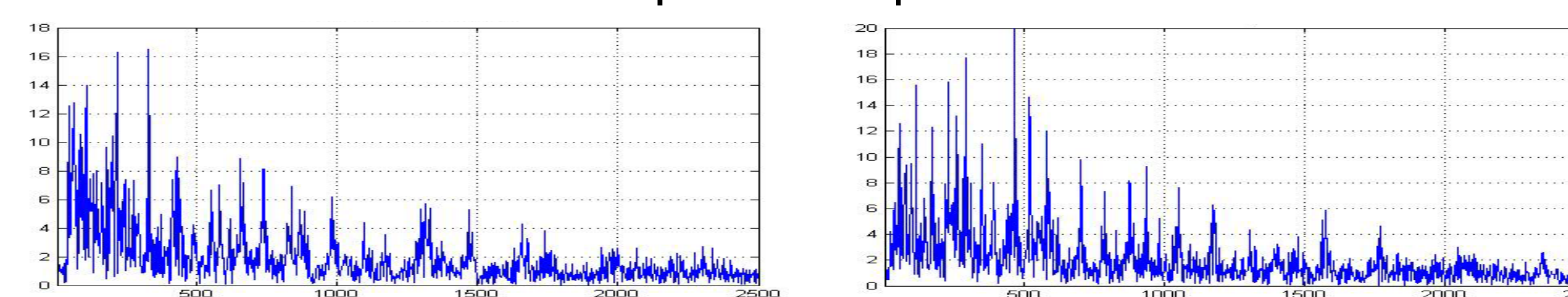
Low   Mid-Range   High

Illustration of Frequency Band Ranges

Examples of Vocal Spectrums

Examples of Music Spectrums

Contact Information

Feel free to contact us at veraxenhancer@umich.edu with questions or requests for additional information.

Acknowledgments