Work-in-progress:
Vocal Intonation Regeneration through Heterodyne Mixing of Overtone Series

Eric Freudenthal, Raul Avarez-Lopez, Vladik Krenovich, Bryan Usevitch,
University of Texas at El Paso
David Roundy, *Oregon State University*

Intonation and why it matters
Human vocalizations are largely overtone series resulting from vocal cord oscillations. Vocaliation pitch (f0) is commonly referred to as intonation. Groups of harmonics which resonate in the vocal tract are frequently far louder than intonation frequency.

Intonation is a key aspect of prosody primarily used to express emotion in many European languages. Thus, people who lack the ability to reliably identify and control it in their own speech are at risk of mis-conveying emotion in their spoken communication.

This deficit, colloquially referred to as “geek syndrome,” can impede effective communication in social and professional contexts, and can even elicit social stigma. Augmentation of intonation pitch in human speech has recently been identified as a potential strategy for attenuating communication deficits related to intonation perception and production and to facilitate on-pitch singing.

Why intonation regeneration is needed
Implementations of Vocal Intonation Boosting (VIB) require the real-time accurate synthesis or selective amplification of intonation within vocalizations.

The low relative amplitude of f0 reduces signal-to-noise ratio and therefore increases the complexity and number of samples (and therefore time) required to accurately determine its frequency. Our preliminary study indicates that it is practical to exploit heterodyne mixing of vocal overtones to boost the relative amplitude of f0 in a manner that facilitates both its rapid isolation and frequency measurement.

Description of Algorithm
Vocalizations are captured using a microphone sampled at 48kHz. As is common, in heterodyne mixers, a non-linear transform is applied to the input signal. Our implementation squares each input sample. The amplitude of the regenerated f0 is of similar amplitude as the overtones.

These higher frequency harmonics are isolated by a second order Butterworth filter whose cutoff frequency is set to 70Hz, which is below f0 in human voice. This filter’s 12 dB per octave roll off ensures that f0 dominates the output signal.

The filtered signal appears to be suitable for the intended purpose of intonation boosting. Furthermore, it appears that f0’s frequency can be quickly and accurately determined to the necessary accuracy from the period of one full f0 cycle detected from the timing of zero-crossings.

Data and relevant citations are available at *http://www.freudensong.com/research/heterodyne1*