Pitch Correction Tool Using YIN and PSOLA

Ben Shafran, Josh Miller, and Brent Ikei

Abstract

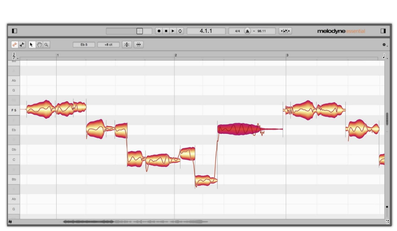
For our final project, we were able to create a functional pitch correction or “auto-tune” software programmed in MATLAB. The three main components of this project are pitch detection (determining the pitch which is being played or sung), the pitch shifting algorithm used to change the pitch of audio without changing the speed at which it is played, and finally pitch correction: the “tuning” aspect of the auto-tune.

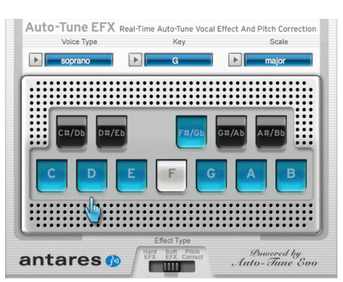
Objectives

Our goal for this project was to make a pitch correction software using YIN and Pitch Synchronous Overlap-Add (PSOLA) algorithms. The YIN algorithm was used to detect pitches. From then, the detected pitches were sent to the PSOLA function to do pitch correction. We set out to have minimal distortion in our output signal and an effect that could be easily applied to any signal. Additionally, we wanted to have flexible methods for choosing how pitches are corrected: either by a defined major or minor scale, or by a drawn in pitch curve.

Implementation

Currently our program takes in a 10-second audio input from a microphone. Using YIN to detect the frequencies of the audio given, the user was able to either draw pitches that they would want to shift to on the graph of the frequencies given, or pick a predefined scale, e.g. G minor. After applying the PSOLA algorithm, the main function would output the pitch corrected signal for the user to hear.





YIN (Pitch Detection)

The YIN algorithm is essentially a modified version of the autocorrelation function used in the time domain. Using the difference function on frames of the input signal, YIN is able to accurately use the frames calculated to determine outlier dips which correspond to periods of the signal. After using the cumulative mean normalized difference function to improve accuracy, parabolic interpolation is used to find the exact location of these dips corresponding to periods [2] and output a plot with a clean array of frequencies per given frame as shown in Figure 1.

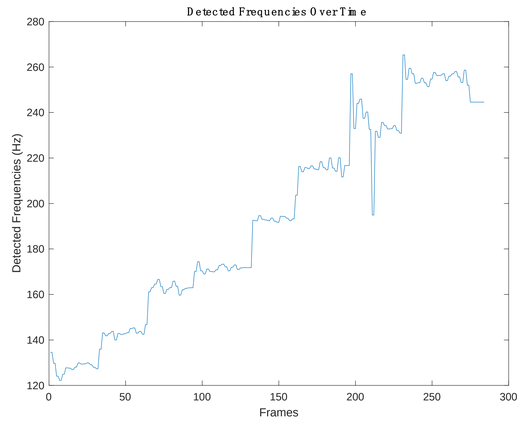


Figure 1: Detected Frequencies for a C major Scale Using YIN

PSOLA (Pitch Shifting)

Pitch Synchronous Overlap Add (PSOLA) is the algorithm we utilized for pitch shifting and correction. Optimized for the human voice, this time domain method preserves the length of the original signal while shifting pitch on a frame by frame basis.

PSOLA achieves pitch shifting by first marking and windowing pitch periods within a frame and then remapping them to contain either more or less periods within the same time range, making the pitch higher or lower [3]. This process can be seen in Figure 2. For this remapping we utilized MATLAB’s k-nearest neighbor function. This algorithm provides very high quality pitch shifting within an octave range.

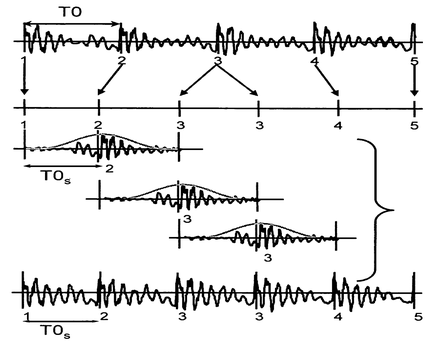


Figure 2: Pitch Periods and their remapping in PSOLA Algorithm

Conclusions and Future Work

Our program is able to successfully pitch correct a given input signal. With user inputs, we were able to record a person’s voice and shift pitches to a value chosen, by drawing a pitch curve or by choosing a scale.

For future work, we would want to improve upon both algorithms used. For YIN, this would involve more effective thresholding for silence detection. For PSOLA, this would involve improved epoch marking. Both algorithms could benefit from further experimentation with window size and filtering techniques. Lastly, we would like to implement some parameters of the pitch correction that the user could modify, e.g. attack time.

References

[1] Cnx.org. (2012). *Auto-Tune*. [Online] Available at: https://cnx.org/exports/22567958-1f9b-4426-8abe-b9e0736df034@1.1.pdf/auto-tune-1.1.pdf [Accessed 29 Apr. 2018].

[2] Alan de Cheveigné, “YIN, a fundamental frequency estimator for speech and music” *Acoustical Society of America*, vol. 111, no. 4, April, 2002. [Online serial]. Available: http://audition.ens.fr/adc/pdf/2002_JASA_YIN.pdf [Accessed 30 Apr. 2018].

[3] Ricardo Gutierrez-Osuna, “L19: Prosodic modification of speech” *research.cs.tamu.edu*, [Online]. Available: http://research.cs.tamu.edu/prism/lectures/sp/l19.pdf [Accessed 30 Apr. 2018].