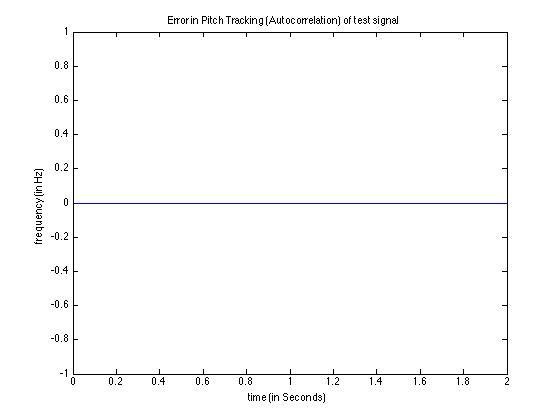
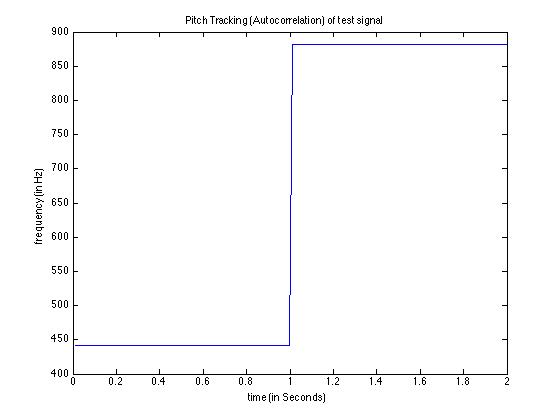
Chris Latina + Liang Tang

MUSI-6201: Assignment 1

September 14, 2015

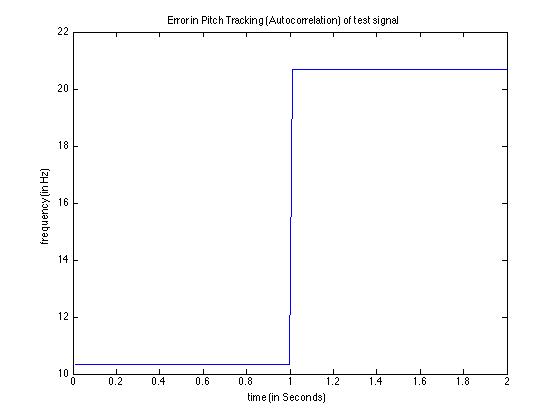
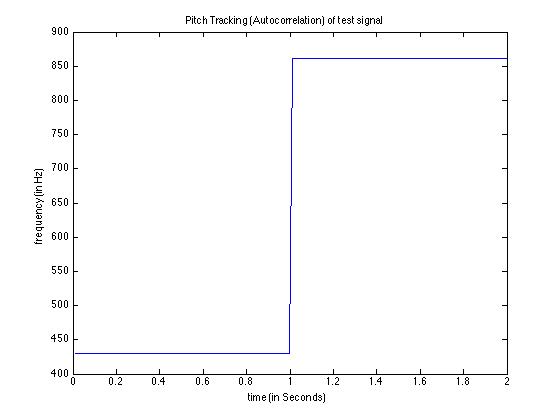
1. **Plot the f0 curve. Plot the absolute error and discuss the possible causes for the deviation.**

Part a, For the test signal which is the sine wave of 441 Hz during 0 to 1s and 882 Hz during 1 to 2s, this ACF pitch tracker works very well. (Parameters: window size: 1024, hop size: 512, sample frequency: 44100)

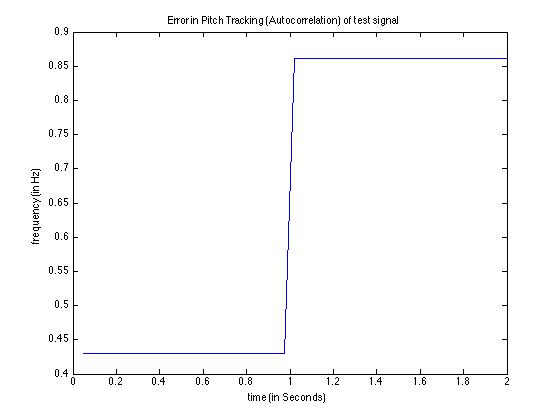


1. **Plot the f0 curve. Plot the absolute error and discuss the possible causes for the deviation.**

Part a, For the test signal, this max spectrum pitch tracker still works, and the error is below 5%, however, it is not as well as ACF one. It runs much faster. The following is with the following parameters: window size: 1024, hop size: 512, sample frequency: 44100.



Raising the window size and hop size to 4096 increases the resolution in the frequency domain and this reduces the error significantly.



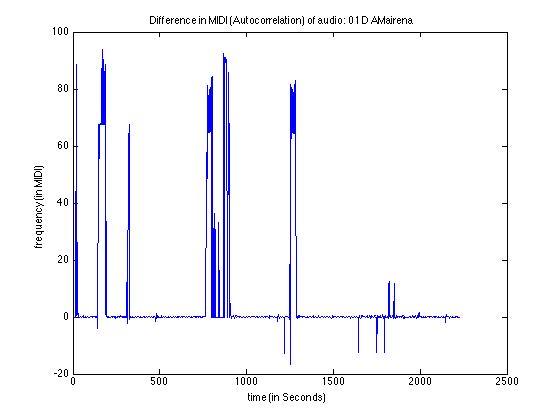
1. **Please report the overall errCent\_rms of the training set for each method. What are the differences between the two methods? What are the potential solutions to improve their performances?**

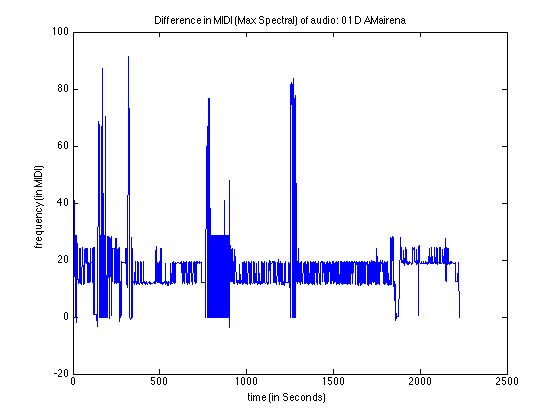
File: 01-D\_AMairena.wav

*rmsACF*: 19.114

*rmsMaxSpec*: 20.189

The error shown below is primarily during silence. Otherwise, the ACF approach is very accurate.

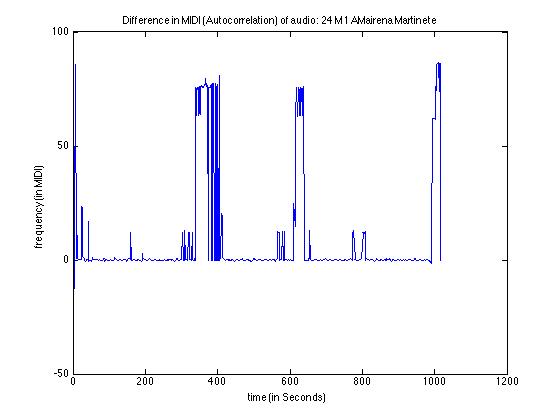


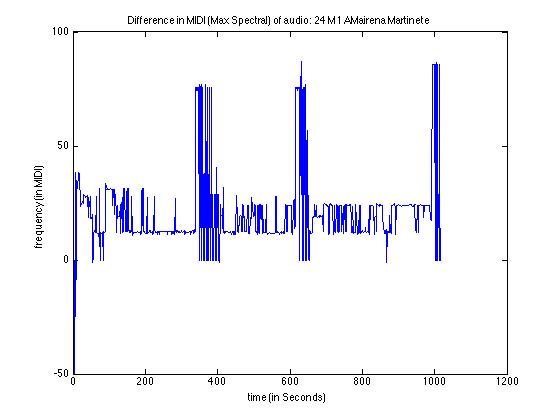


File: 24-M1\_AMairena-Martinete.wav

*rmsACF*: 23.402

*rmsMaxSpec*: 26.073

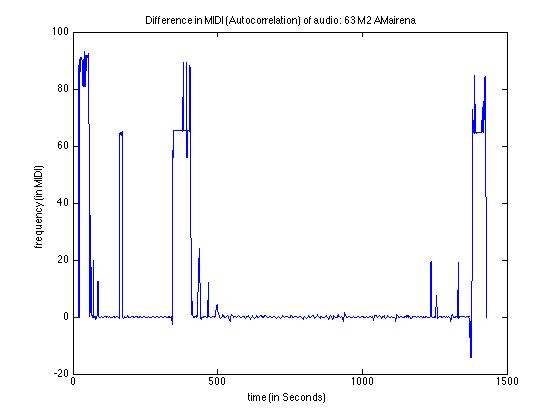


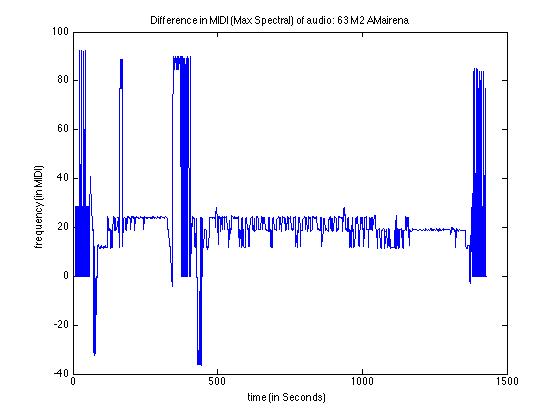


File: 63-M2\_AMairena.wav

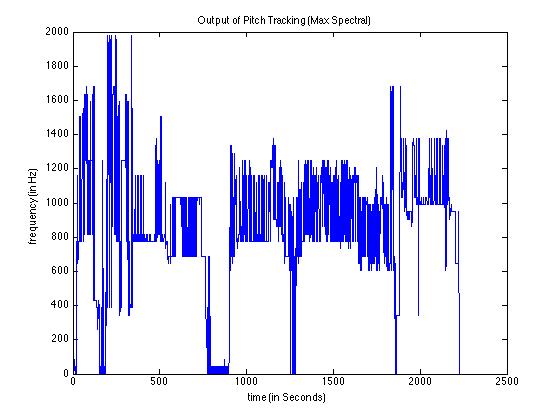
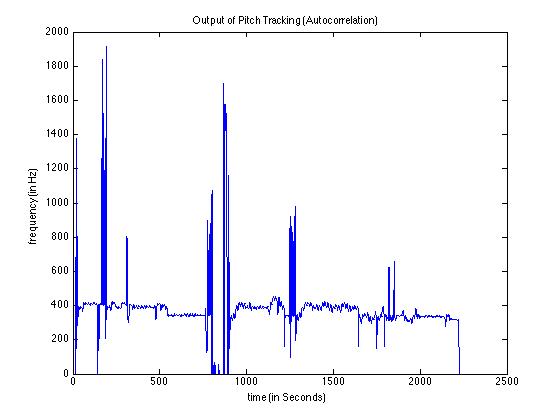
*rmsACF*: 23.894

*rmsMaxSpec*: 28.127





*Below are the plots for the raw f0 data for the first audio file:* **01-D\_AMairena.wav***.* You can see that the silence causes the significant error in the ACF approach. The Max Spectral approach is finding the 2nd and third harmonics, vacillating between around 800 and 1200 Hz.



To improve performance, preprocessing would significantly help. By filtering the spectral envelope in the frequency domain, one could achieve more accurate results using MaxSpec. MaxSpec, however, does require a larger window size for accuracy. On the other hand, it is much faster than the ACF.

Another improvement could be taking the autocorrelation in the frequency domain to speed up the ACF. Comparing the ACF results to the MaxSpec results could also lead to accurate readings.   
  
Lastly, a cepstral approach to pitch shifting could be another approach.

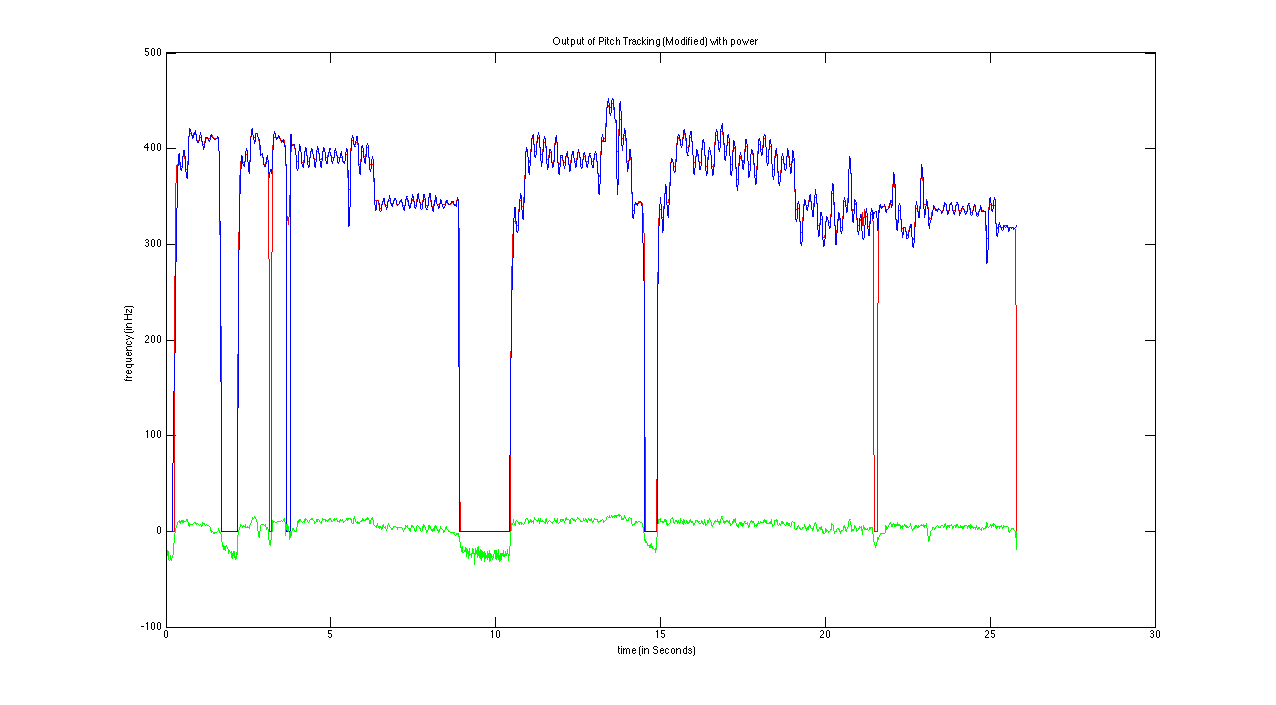
1. **Modified version**

After attempted to filter in the magnitude spectrum, we decided to instead use the ACF as our basis because it was more accurate. Since our ACF was closest, we added additional checks to remove the noise elements in the output (the silence). After calculating the power of the magnitude spectrum, we conditionally set the frequency to 0 if it was below -8 dB after converting with the pow2db function. This threshold worked well.

We also applied a 10th order medium filter to remove any spikes.

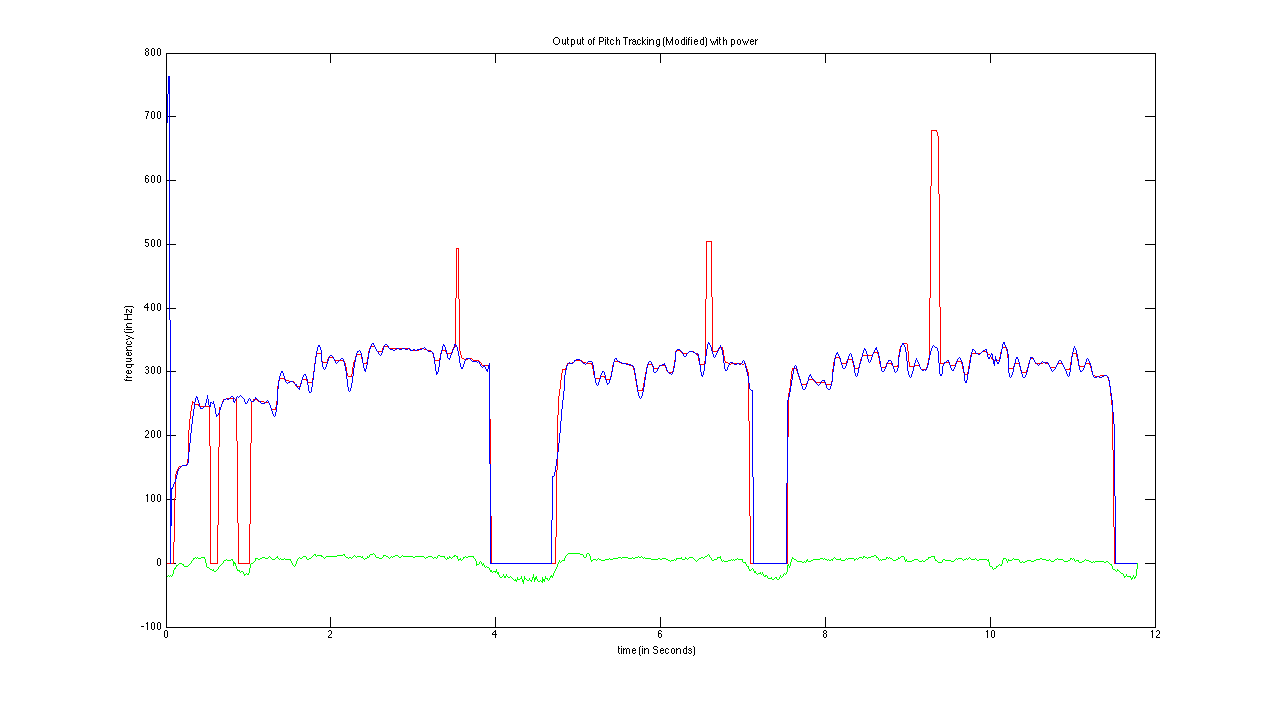
*Filename:* 01-D\_AMairena

*rmsMod*: 5.328



*Filename:* 24-M1\_AMairena-Martinete

*rmsMod:* 2.373



*Filename: 63-M2\_AMairena*

*rmsMod: 7.062*

