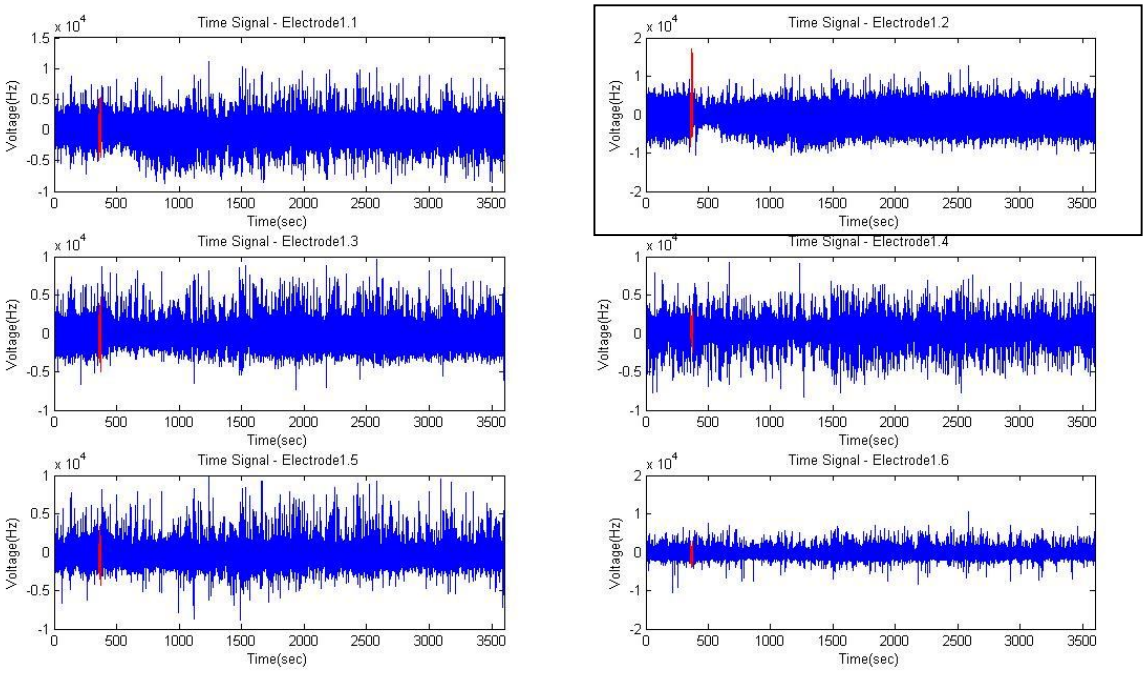
Lisa Jiang
BE301- 12/17/2011
Final Project

Epilepsy is a disorder characterized by the occurrence of epileptic seizures and their neurobiological, cognitive, psychological, and social consequences (Fisher, 2005). It affects nearly 3 million Americans and 50 million people worldwide. While medications reduce the symptoms of epilepsy, more than a million people continue to have seizures that severely limit how they live their lives ("Epilepsy Foundation", 2010). Developing a mechanism for seizure detection would result in reliable seizure tracking and can lead to steps for prevention and counteraction. Due to the prevalence of Epilepsy and the utility of being able to detect seizures, for my final project, I developed an algorithm for seizure detection. Seizure events can be discriminated from normal brain activity by analyzing electroencephalographic (EEG) signals collected from the brain (Oweis, 2011). Patient EEG data was analyzed to determine the best method to extract seizure data. Then, the extracted features were used to develop an algorithm to identify seizures. This algorithm was then applied among multiple subjects to test performance.

I obtained EEG recordings from The Seizure Prediction Project with permission from the Freiburg Center for Data Analysis and Modeling (University of Freiburg, 2011). The database consists of EEG recordings of twenty-one patients suffering from medically intractable focal epilepsy. Each patient has approximately twenty four hours of seizure and non-seizure EEG recordings from six different electrodes. I primarily focused on patients suffering from frontal lobe epilepsy, because analyzing seizures originating from varyingr areas of the brain could result in different seizure signals and thus different detection techniques. I used Patient001's EEG data as the training set to develop an algorithm for seizure detection, and I used Patient003 and Patient005's data as the set to test the algorithm.

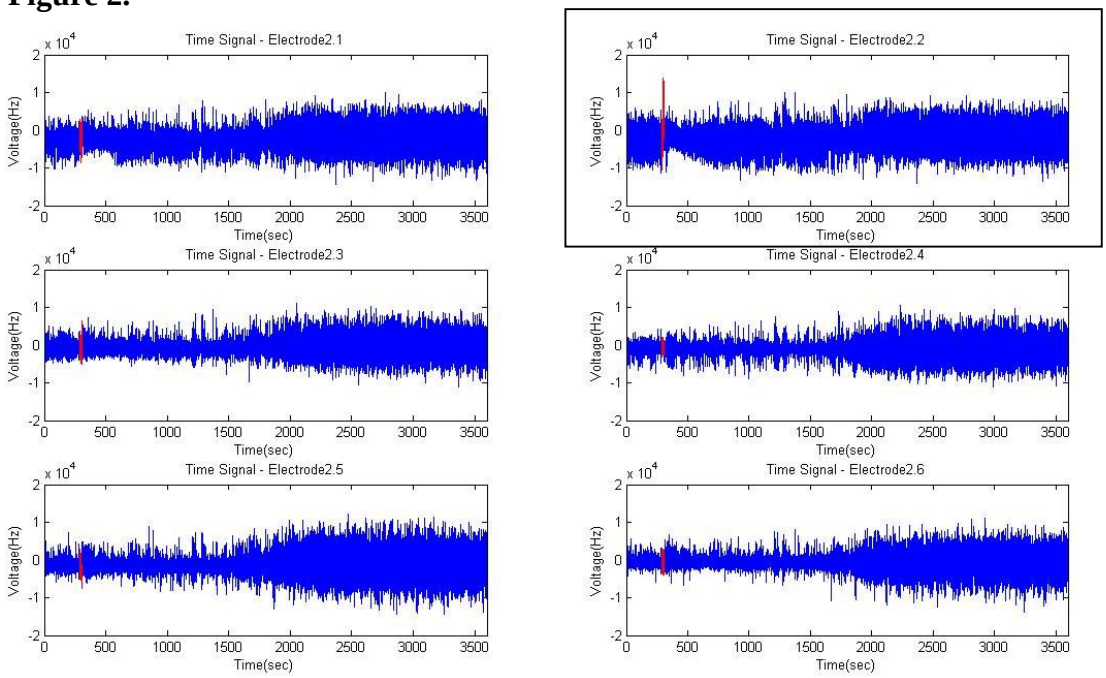
Figure 1 shows the MATLAB plot of the first time segment of Patient001's EEG signal. There are six plots, representing the six different electrode channels recorded. The blue signal represents the entire signal, and the red represents the highlighted seizure events.

Figure 1.



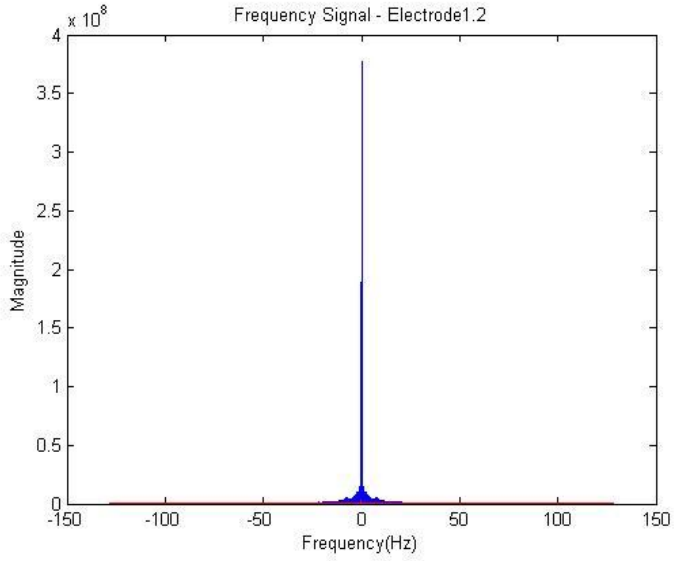
Based on the time-domain plots, Electrode2 is the best channel to use for seizure detection, because it has the most distinct seizure signal. A plot of the second time segment (Figure 2) for Patient001 for electrodes 1-6 also confirms this. This data makes sense, because the electrodes were placed based on three focal and three extra-focal points. Focal points represent sites initially involved in ictal (seizure) activity, and extra-focal points represent sites not involved in the seizure or involved late in the spread of ictal activity. For this reason, some of the electrodes may be recording signal irrelevant to seizure detection. All following analysis will be performed based on the signals from Electrode2, since it represents the most relevant EEG seizure data.

Figure 2.



To view the signals in the frequency domain, the fft function in MATLAB was used. The EEG data was sampled at 256Hz, and it was nice to be able to clearly see the bandwidth of 128 when the frequency was plotted. Similar to the time domain data, I tried to highlight the seizure in the frequency domain as well (Figure 3).

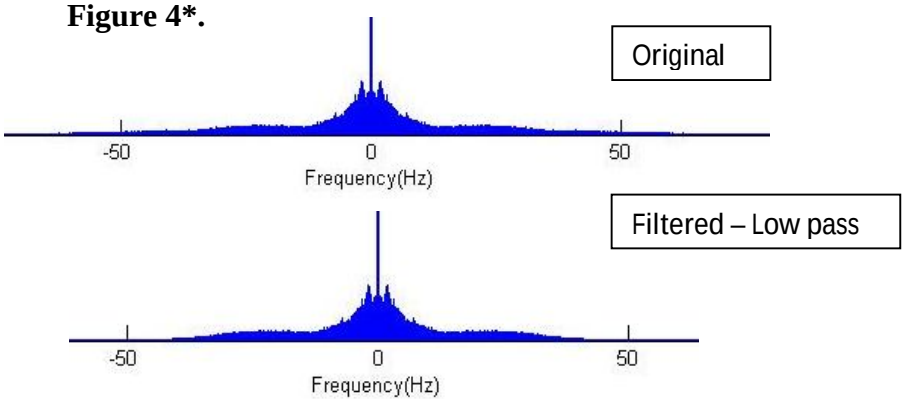
Figure 3.



The blue curve represents the entire signal, and the red curve on the bottom represents the "highlighted" seizure. I understood mathematically how to convert back and forth from the frequency and the time domain, and I understood how the values of the frequency changed based on the time domain. However, I realized that I didn't understand physically what the frequency domain actually meant. I had tried to translate time values directly into frequency values by taking the fast Fourier transform of the seizure data separately and plotting it against a frequency vector that I had converted from a time vector. When I zoomed into the signal in MATLAB, I saw that the red and the blue did not overlap. Instead, the red curve was actually below all values of the blue curve. This was an important learning step, because it really helped me understand that frequency domain was based on separating the signal into a spectrum of frequencies. The magnitude at each frequency shows how much of the total signal in the time domain actually lies within each frequency. The fact the red curve was below the blue curve outlined that the summation of truncated signal would give a lower magnitude at each frequency than a summation of the whole signal. This was further enforced by the fact that the range of frequencies was actually the same for both signals, from -128 to 128.

According to the Freiburg Center, each of the signals included noise at around 50Hz. So, I tried applying different filters that would eliminate signals at the 50Hz frequency. The two filters I tried were the low pass filter and the band stop filter. For each, I used the Butterworth filter, because it has an ideal, flat top. For the low pass filter, I filtered out all frequencies above 40Hz. The frequency domain plot accurately represented the filter (Figure 4). Comparing the original and the new frequency domain plots, the range clearly decreased from about 60Hz to about 40Hz, but the magnitude remained the same. In addition, it showed that real world filters cannot be ideal, because, although insignificant, there was still signal left above 40Hz.

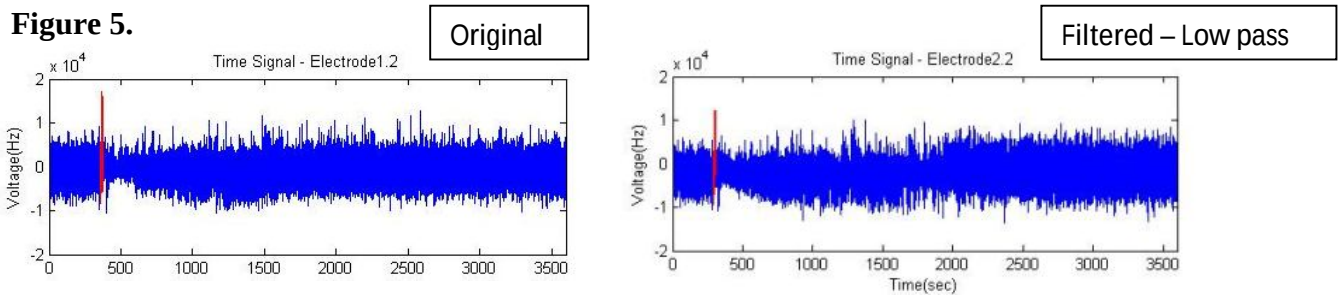
Figure 4*.



*plots were zoomed in to show difference in frequency range

Though there was a clear difference in the frequency domain representations, the time domain representation of the filtered signal looked relatively similar to the original (Figure 5). However, the signal is actually decreased. This shows that significant parts of the signal actually exist above the 50Hz frequency, and a low pass filter is not a good option for seizure extraction.

Figure 5.



The band stop filter was a much better option. To implement it in MATLAB, I had to eliminate signals from 49Hz to 51Hz (Figure 6). Even though the Butterworth filter could not perform a sharp band stop for just the 50Hz signals, it cut out a lot less vital signal than the low pass filter. This resulted in a seizure signal of the same size in the time domain (Figure 7).

Figure 6.

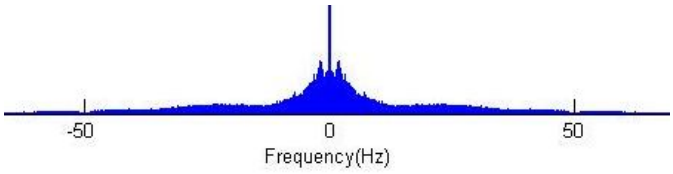
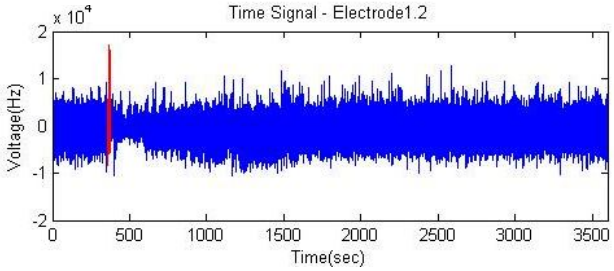


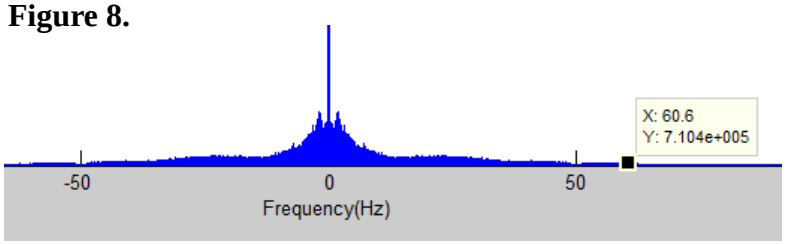
Figure 7.



Since there was so much data for each patient, the signals were each separated into 25200second recordings each. In order for MATLAB to compile all the data into one signal without running out of memory, all the signals needed to be down sampled. The bandwidth of the signal was 128Hz, which gives a Nyquist Rate of 256Hz. By down sampling this data, the sampling rate would drop below 256Hz. This really enforced the concept of aliasing and the need for low pass filters. Without filters, the newly sampled data would not be an accurate representation of the original signal, because the distribution of signal among each frequency in overlapping frequencies would be greater compared to the original.

To determine the sampling factor, M, for down sampling I looked at the frequency signal. Most of the signal was in the range of -60Hz to 60Hz of the signal (Figure 8). There seemed to be very little signal from 60Hz to 128Hz. Therefore, I determined that the signal could run through a low pass, antialiasing filter of 60Hz. This would result in a sampling rate of 120. Therefore, M could be 2 in this case. Each recording was sampled at 120Hz and combined.

Figure 8.



Once the combined signal was filtered, I examined a variety of signal manipulations to measure signal size and extract important features (equations or a set of equations). I used several of the most common features in signal analysis. I plotted the energy, power, area, entropy, line-length, and time aligned line-length features. On the same plot, I also plotted the actual seizure locations to determine if seizure locations became more distinct based on the extracted features. The two features which highlighted seizure locations best were the energy (Figure 9) and the time aligned line-length feature (Figure 10). The energy can be described by the equation $E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt$, and the line length feature computes the total distance between successive points in a series. The time aligned line-length feature also normalizes the data set so that a set of values centered along y = 10 can be treated the same as a set of values centered along y = 0. By comparing the blue signals in the plots to the black lines that indicate actual

seizure location (Figure 9,10), it was determined that line length was a better indicator of seizure location.

Figure 9.

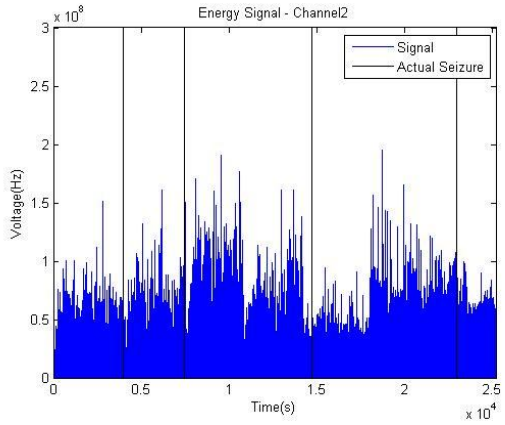
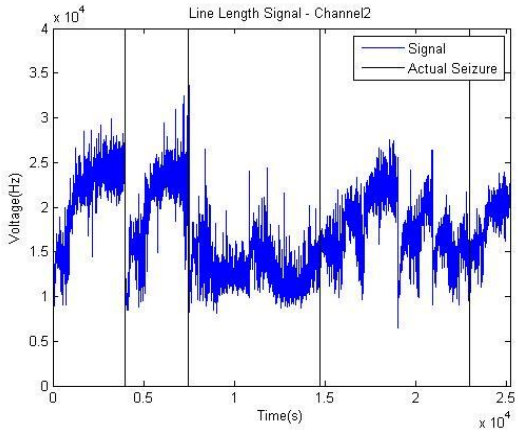


Figure 10.



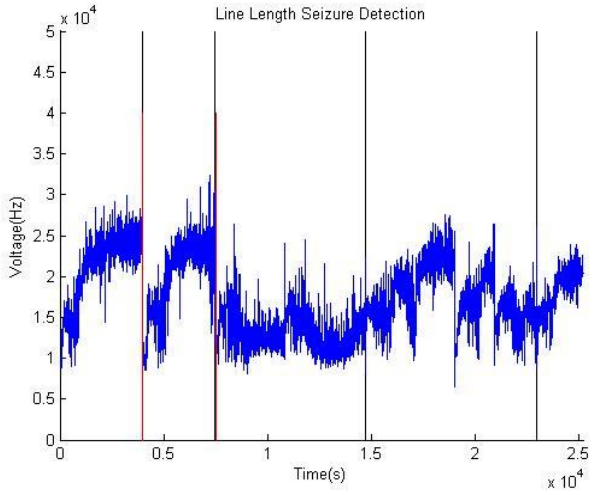
The time aligned line length feature was calculated and scaled with the help of the zohinterp and rfeature functions:

```
%window and displacement
L = 10; d = 1;
%line length feature: rf = running feature, ni = number ignored from left
[rf,ni]
=rfeature(data,inline('sum(abs(diff(data)))'),L*frequency,d*frequency);
%scales up the graph/aligns
llFeat = [NaN(1, ni+L*frequency-1) zohinterp(rf(1:end-1), d*frequency)
rf(end)];
%normalize
plot(time, llFeat*2*max(data)/max(llFeat), 'b')
```

For seizure detection, a threshold for rf was set at 9.0E5Hz, and a seizure was considered detected every time the signal crossed over the threshold. The next point was not counted as a seizure until after the signal crossed back under the threshold. If the threshold level was set too low, there would be too many seizures detected, and if the threshold level was set too high then there wouldn't be enough. The threshold of 9.0E5Hz served as a good balance threshold for Patient001. In Patient001, where the threshold was set, two of the four seizures were detected (Figure 11). When the same algorithm and threshold were used on Patient003 and Patient005, four of four and zero of four seizures were detected respectively. However, when the threshold was lowered to 1.5E5Hz for Patient005, four of four seizures were detected with two false positive (Figure 12). This shows that the algorithm could remain the same for seizure detection, but the threshold may need to be set individually based on the patient.

Understanding signal manipulation, the frequency domain, aliasing, bandwidth, the Nyquist rate, and filters were all vital in the ability to understand and develop an algorithm for signal detection. Without the prior knowledge and skills in those areas, it would have been hard to compile and obtain a directly usable seizure signal. Having a background in signals and systems gave me the foundation to eliminate the noise and compile the most useful, accurate portion of the signal to analyze. It was extremely rewarding to see that, with just the tools from this semester, it was actually possible to solve a problem with an immediate, real world application

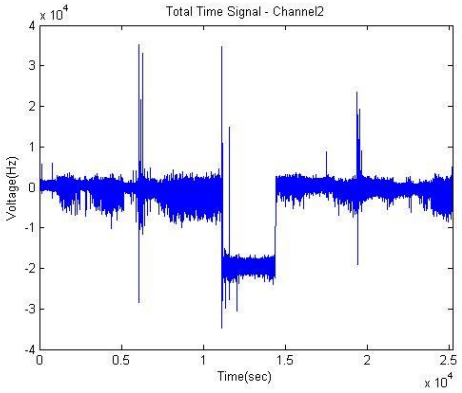
Figure 11*.



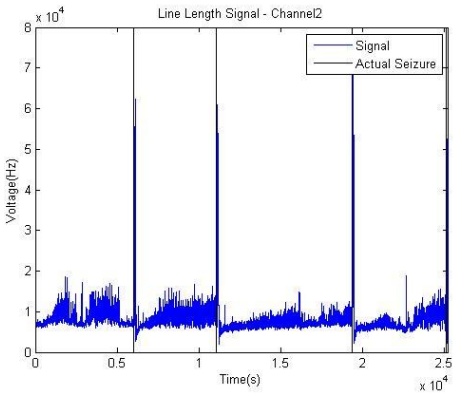
*In the figures, blue represents the signal. Black represents the actual seizure locations, and red represents the predicted seizure locations.

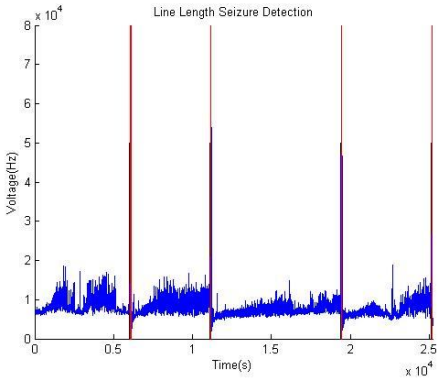
Figure 12*.

Patient003 – Time Signal

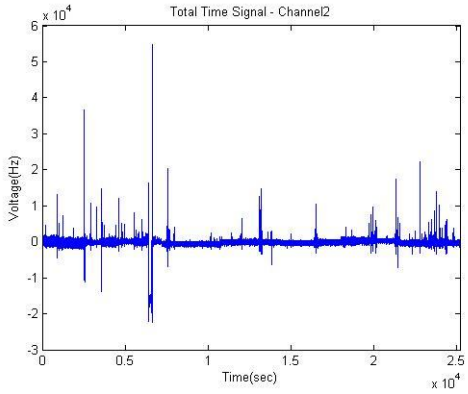


Patient003 – Line Length Signal

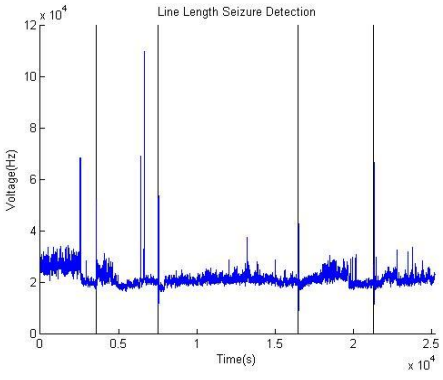


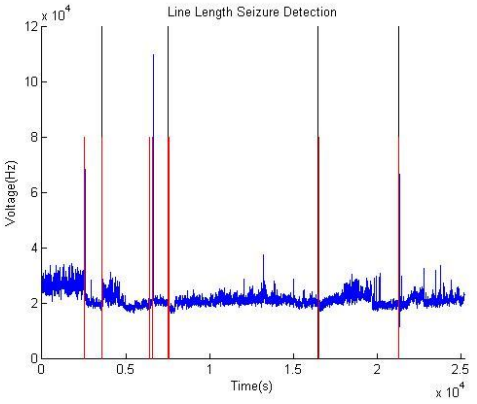


Patient005 – Time Signal



Patient005 – Line Length Signal





Works Cited

Epilepsy foundation. (2010). Retrieved from http://www.epilepsyfoundation.org/aboutepilepsy/.

Fisher RS, Van Emde Boas W, Blume W, Elger C, Genton P, Lee P, Engel J. (2005) Epileptic seizures and epilepsy: definition proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE). Epilepsia 46:470–472.

Oweis RJ, Abdulhay EW. (2011). Seizure classification in EEG signals utilizing Hilbert-Huang transform. Biomed Eng Online.

University of Freiburg. (2011). EEG Database. Freiburg Seizure Prediction Project. Retrieved from http://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database.