Signal Processing Project

Fast and Fourier

Members: Vedant P, Varun S, Siddarth G

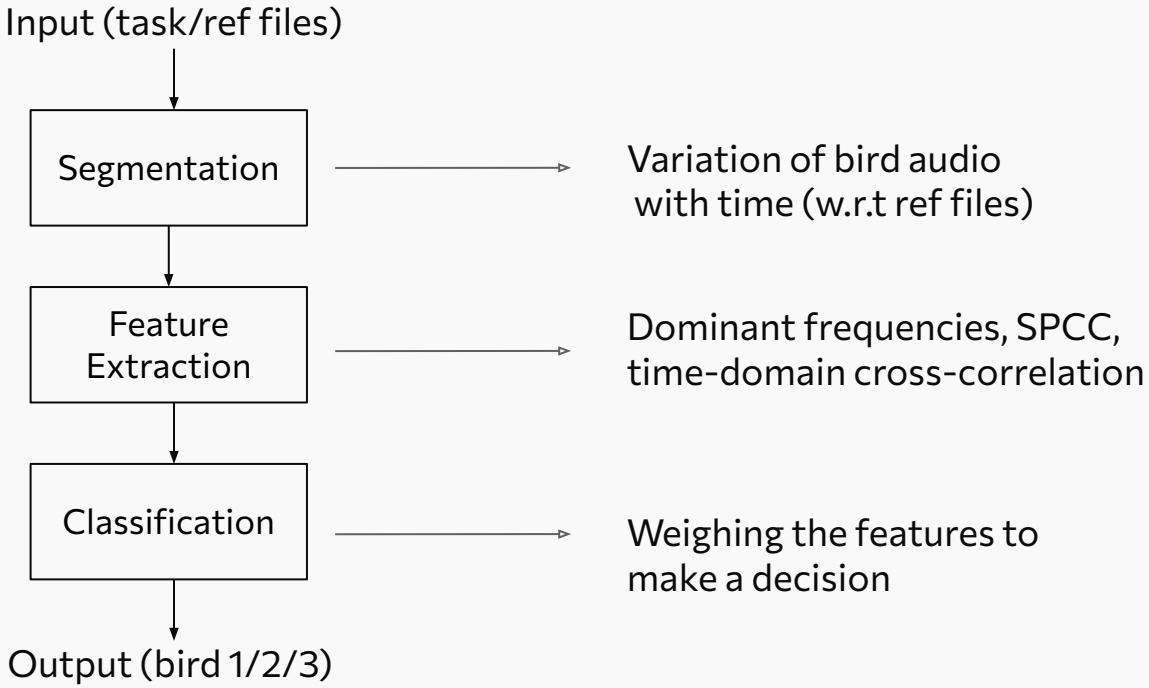
Contents

01

Bird Recognition

Mapping task bird chirps to
reference chirp samples

Workflow


Input (task/ref files)
Segmentation
Variation of bird audio with time (w.r.t ref files)
Feature Extraction
Dominant frequencies, SPCC, time-domain cross-correlation
Classification
Weighing the features to make a decision
Output (bird 1/2/3)

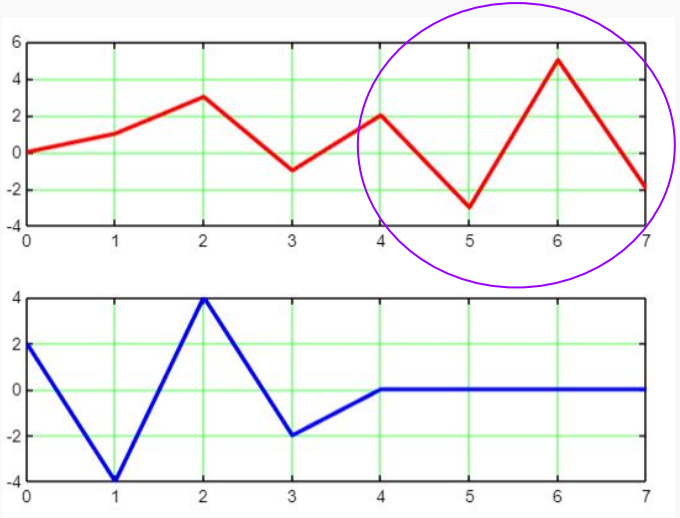
Features

- Initial attempt
  - - Simple Frequency Matching
  - - The problem of non-stationarity!
- Spectrogram and Dominant Frequency Analysis
  - - Merlin Bird ID by Cornell University
- Cross-Correlation
  - - Why Correlation?
  - - Problem with time-domain cross correlation
- Importance of Normalization!
- A natural extension to a weighted average

Cross-Correlation!

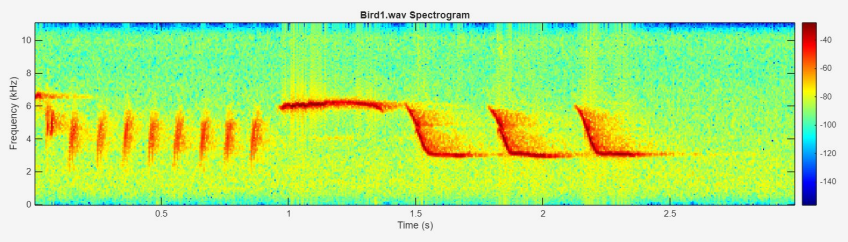
$$R_{xy}[n] = \Sigma_m x[m].y[m+n]$$

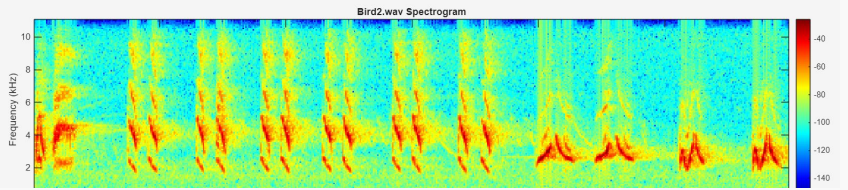
- Correlation with shifts in time!
- Measures the similarity between two signals as a function of time lag
- Helps identify overlapping patterns!

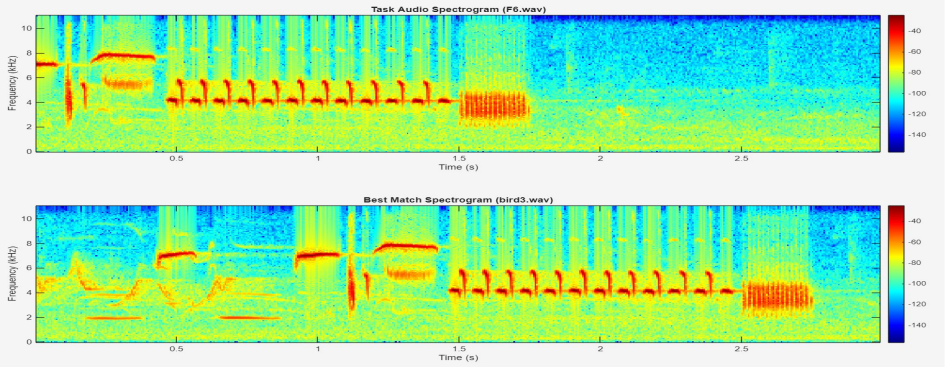


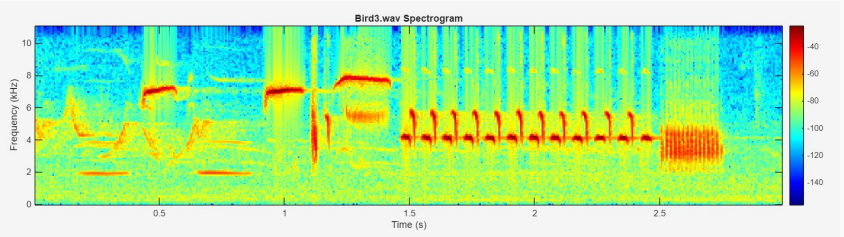
Spectrograms and Enhancements

- Spectrogram plots to the right! (Eg. F6.wav)
- The anomalies we find in F7.wav









Results and References

- 1) “Bird Species Identification Using Signal Processing” Chalmers University of Technology
  2) “Bird Chirps Annotation Using Time-Frequency Domain Analysis” - Suveen Kumar Vundavalli Sri Rama Srinivasa Varma Danthuluri

- Possible improvements going ahead
  - Neural Networks? Deep Learning?
  - Shazam but for birds?
    - DTW Algorithms

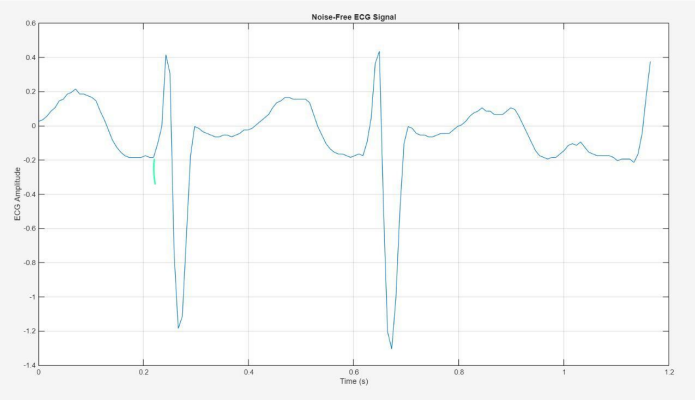
Results:

F1.wav → Bird 3
F2.wav → Bird 1
F3.wav → Bird 2
F4.wav → Bird 3
F5.wav → Bird 1
F6.wav → Bird 3
F7.wav → Bird 2
F8.wav → Bird 2

02

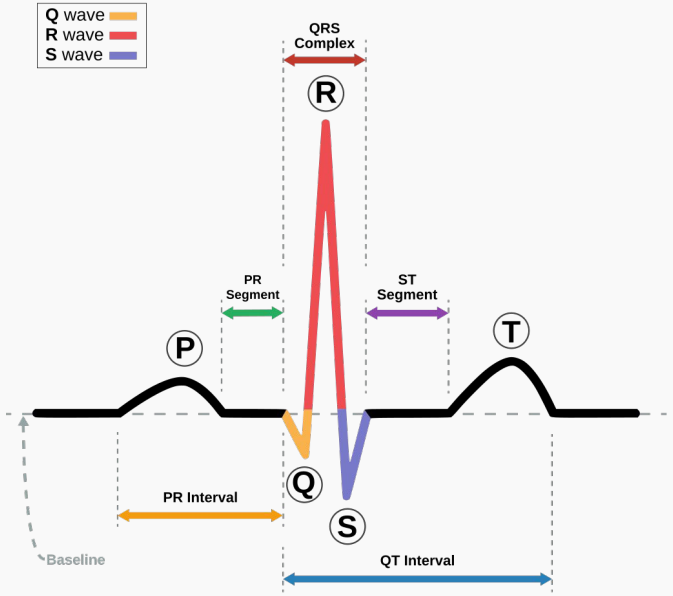
Heart Rate Estimation

Estimating HR from noiseless and noisy ECG data



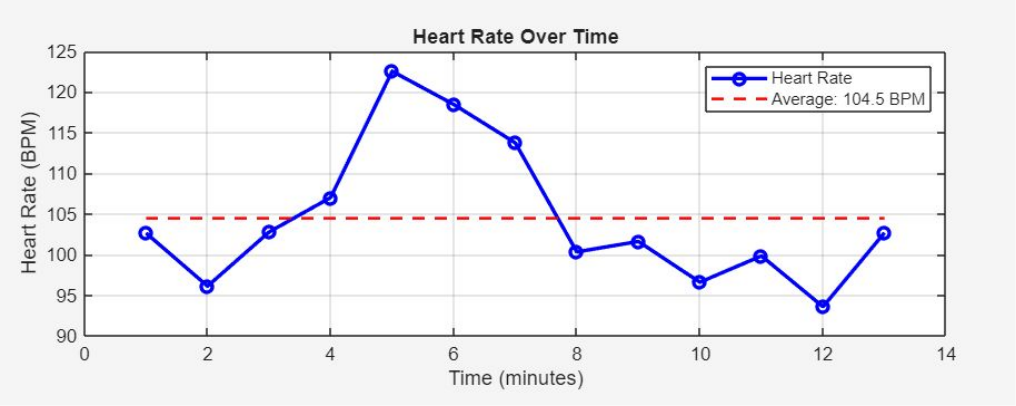
QRS Complex and R-R peak detection

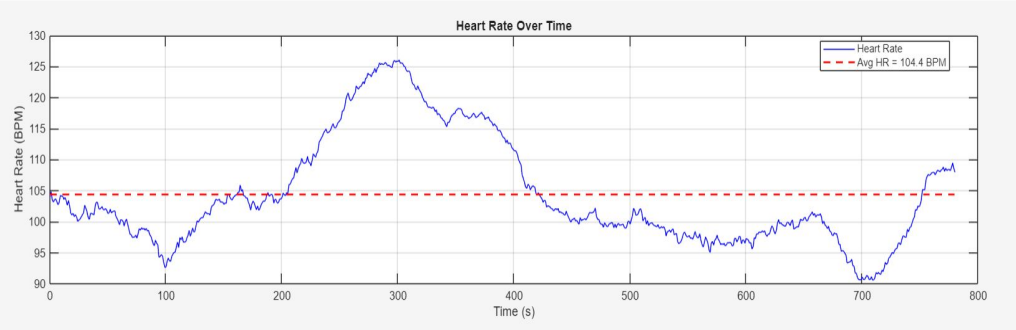
- Combination of three graphical deflections seen in a typical ECG
- R wave: Peak of ventricular depolarization (most significant electrical activity) as the ventricles contract.
- Most prominent and easily detectable, hence the R-R peak detection.
- Number of R-peaks in 1 minute, or equivalently, $\frac{60}{\text{R-R interval (sec)}}$



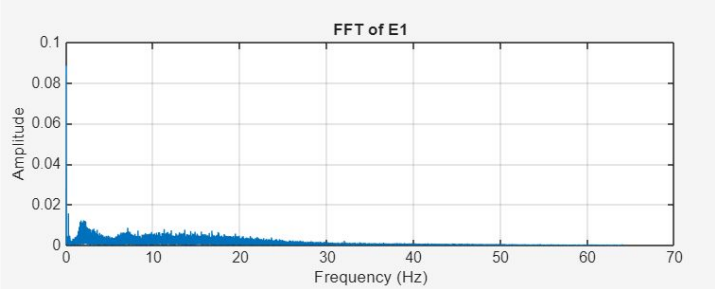
Noiseless ECG

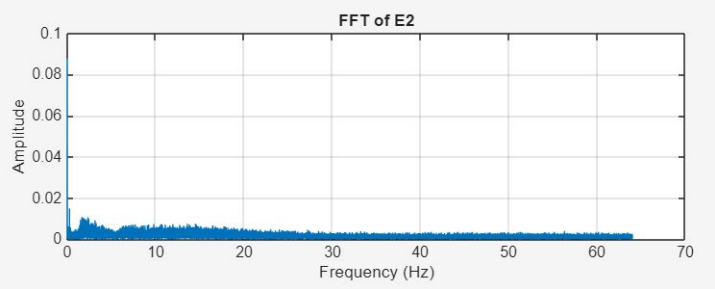
- Assumption of E1 as a reference.
- E1 - noiseless. Straight-forward implementation.
- Sliding window for estimation
- Estimated avg. HR over 13 minutes: 104.5 BPM
- Upon looking at QRS complex, it's a normal rhythm but higher than typical 72-80 BPM.
- Is such an HR possible?



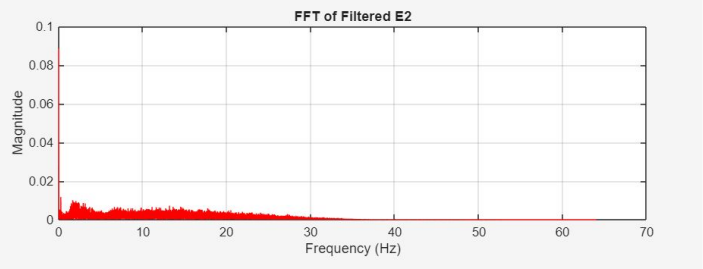


Noisy ECG signal: E2

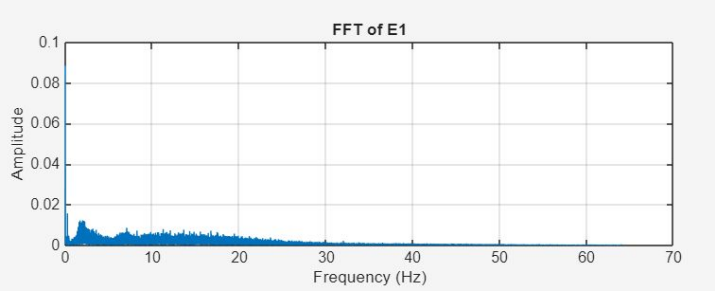


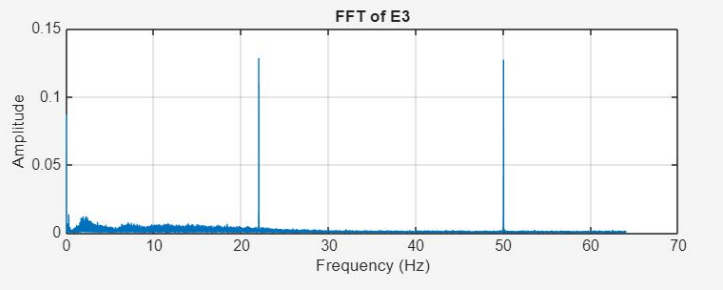


- Low-pass butterworth filter
  Cutoff ~30Hz
- HR drops from ~116 BPM to ~91 BPM

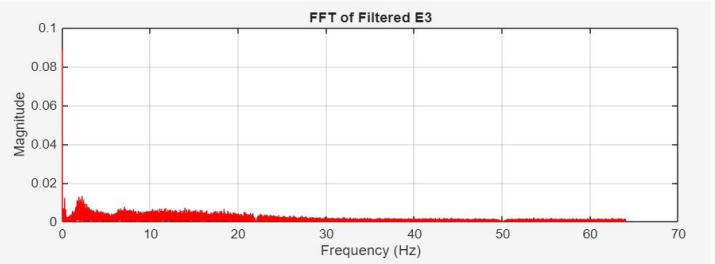


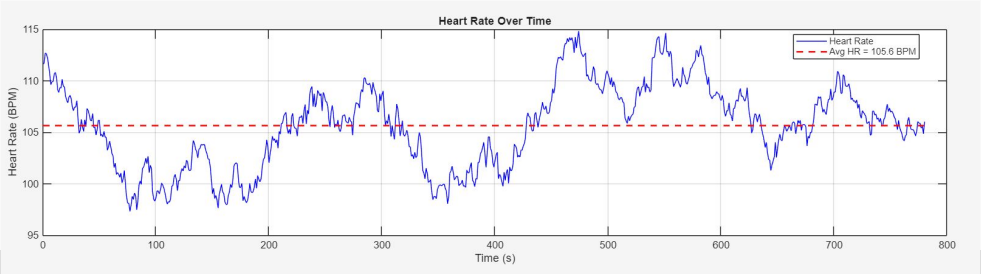
Noisy ECG signal: E3



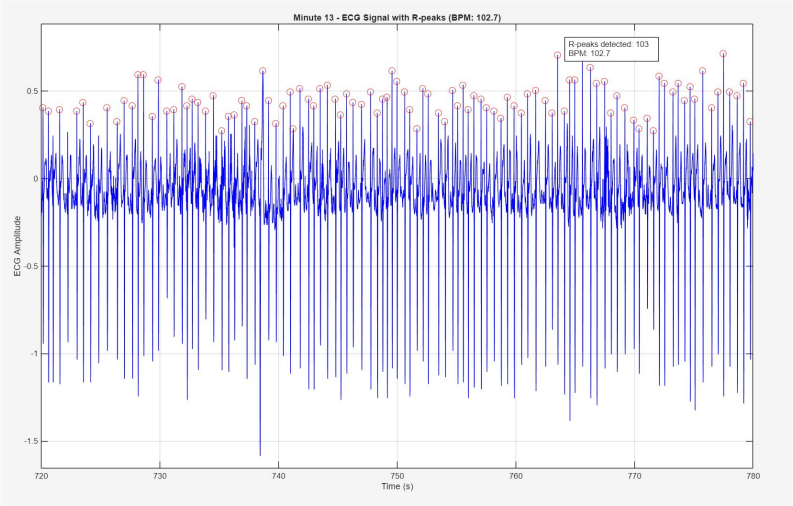


- IIR Notch filter with notches at 22 Hz and 50 Hz, with weak bandwidth.
- HR drops from ~119.4 BPM to ~105.6 BPM





Tabulation and Enhancements



- Current filter implementations work successfully.
- Pan Tompkins Algorithm for QRS detection
  - Peak Enhancements

File Name	Minimum HR		Average HR		Maximum HR	
	Before	After	Before	After	Before	After
E2.mat	104.4	72.3	115.9	91	124.1	112.7
E3.mat	109.6	99	119.4	105.6	129	109.7

03

Loudness Segmentation

Detection of loud words in human voice samples

Identification and Segmentation

Parameters analysed:

- Peak amplitude
- Energy
- Normalised Band Energy

Why normalised and band energy?

- The problem of perception
- Part-2 as a superset of Part-1
- Energy Over Time and Threshold conditions
    - Current issues and Improvements

Problem of Perception

```
Audio - 6
i               0.457101        0.721523        0
told            0.721523        1.004392        0
you             1.004392        1.184773        0
this            1.184773        1.455344        0
would           1.539385        1.781259        1
happen          1.781259        2.357247        0


Mean Values:

Mean Peak Amplitude: 0.41384
Mean Energy: 146.20462
Mean Normalized Band Energy: 109087.26801

Word Analysis:

Word: i         Peak Amplitude: 0.4502          Energy: 112.2807        Normalised Band Energy: 82647.2500      Is Loud: 0
Word: told      Peak Amplitude: 0.5827          Energy: 347.2068        Normalised Band Energy: 245017.9549     Is Loud: 1
Word: you       Peak Amplitude: 0.2595          Energy: 69.7855         Normalised Band Energy: 74735.7111      Is Loud: 0
Word: this      Peak Amplitude: 0.2210          Energy: 28.9711         Normalised Band Energy: 18903.4507      Is Loud: 0
Word: would     Peak Amplitude: 0.5696          Energy: 244.7116        Normalised Band Energy: 207731.4977     Is Loud: 0
Word: happen    Peak Amplitude: 0.4001          Energy: 74.2721         Normalised Band Energy: 25487.7437      Is Loud: 0
```

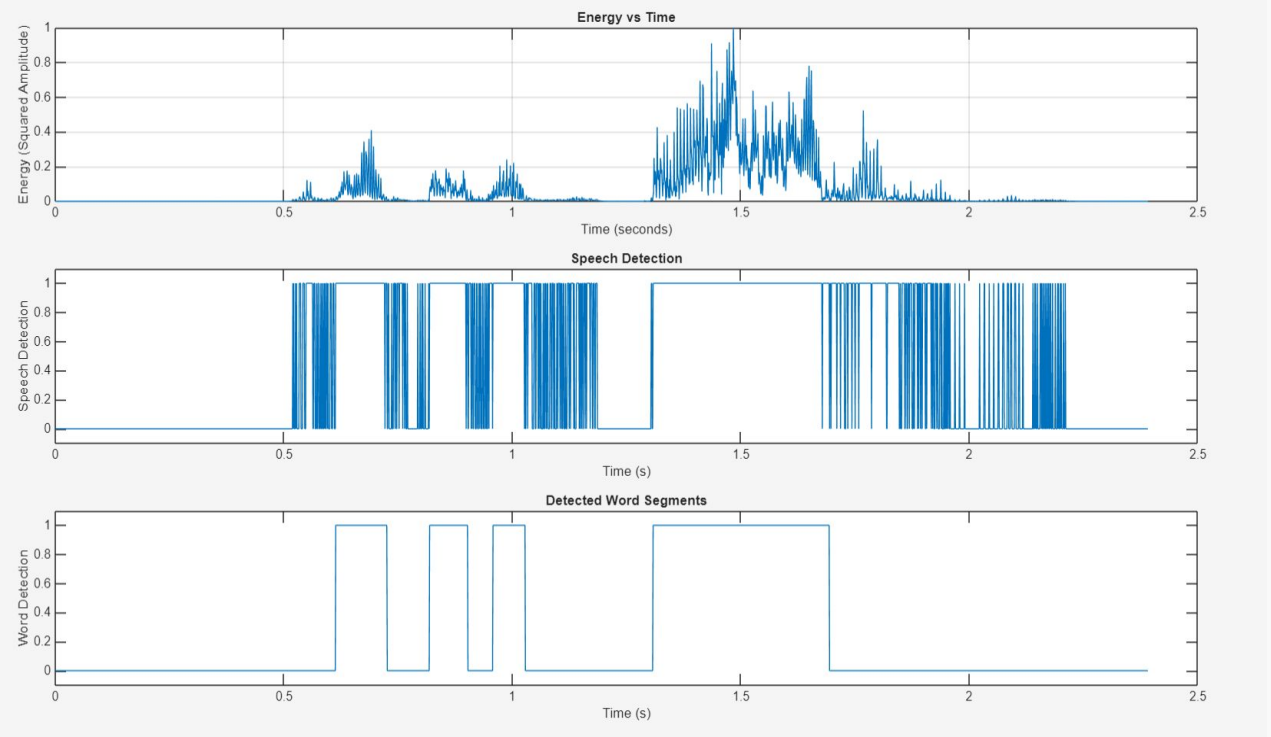
Problem of Perception

```
Audio - 7
i               0.449975        0.556333        0
didnt           0.556333        0.744504        0
say             0.744504        1.026761        0
he              1.186297        1.362197        1
stole           1.362197        1.703768        0
the             1.703768        1.818307        0
money           1.818307        2.125108        0


Mean Values:

Mean Peak Amplitude: 0.25806
Mean Energy: 66.92602
Mean Normalized Band Energy: 52929.30905

Word Analysis:

Word: i         Peak Amplitude: 0.1055          Energy: 1.7946          Normalised Band Energy: 2692.4786       Is Loud: 0
Word: didnt     Peak Amplitude: 0.3721          Energy: 124.1207        Normalised Band Energy: 118663.8192     Is Loud: 1
Word: say       Peak Amplitude: 0.3773          Energy: 87.5915         Normalised Band Energy: 54545.1057      Is Loud: 0
Word: he        Peak Amplitude: 0.3280          Energy: 114.4198        Normalised Band Energy: 111355.3829     Is Loud: 1
Word: stole     Peak Amplitude: 0.3292          Energy: 120.0210        Normalised Band Energy: 65655.5063      Is Loud: 0
Word: the       Peak Amplitude: 0.1531          Energy: 6.5767          Normalised Band Energy: 9822.0274       Is Loud: 0
Word: money     Peak Amplitude: 0.1412          Energy: 13.9579         Normalised Band Energy: 7770.8431       Is Loud: 0
```

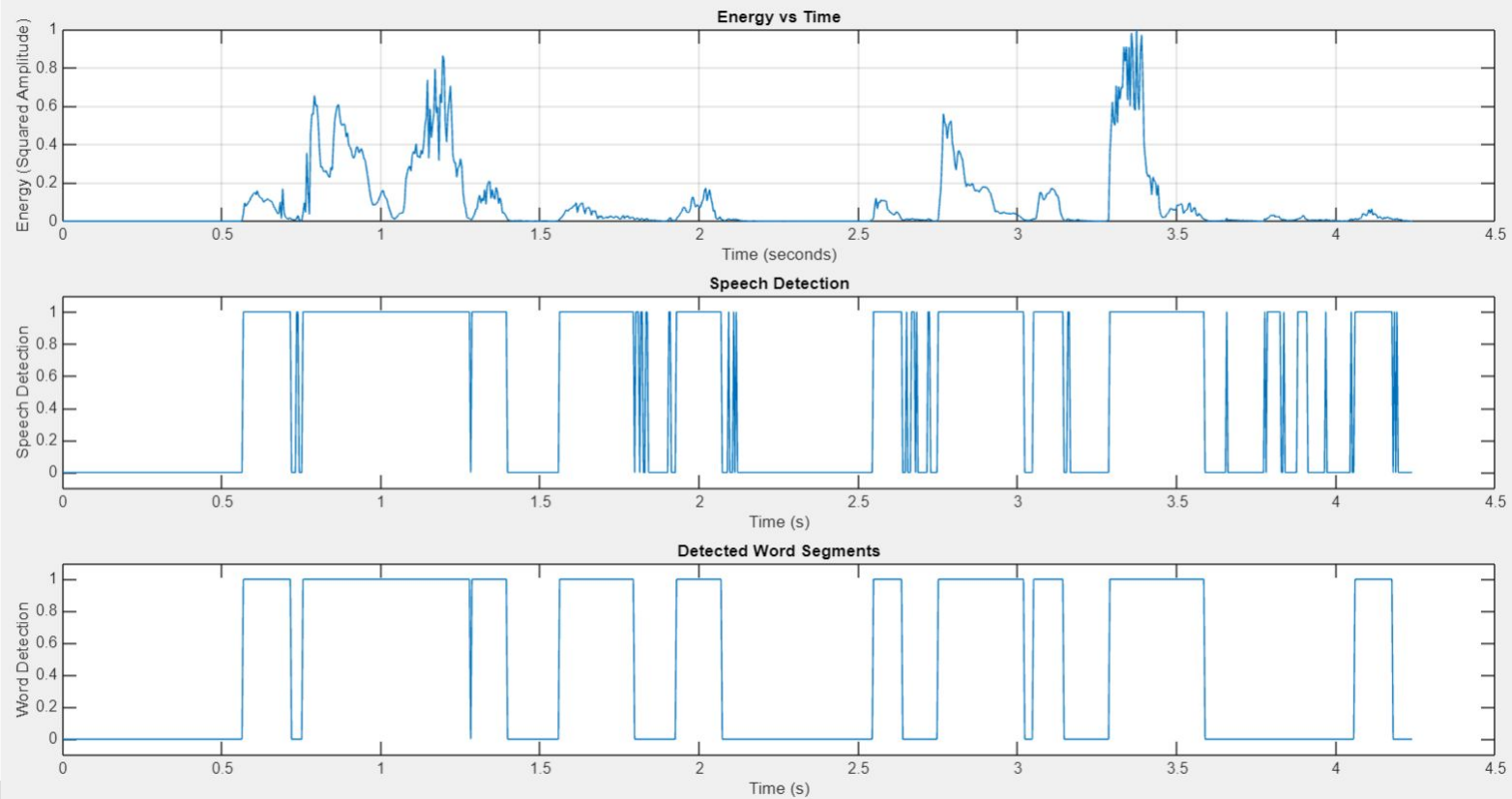
Energy vs Time Plot

Importance of the Window Size:

- Small window size causes lot of fluctuations
- Large window size doesn't give accurate energies and doesn't handle the small fluctuations



Energy Plots and Segmentation Results



Improvements

- Differentiating between syllables and a complete Word
- No silence zone between two words

```
Audio - 2
he              0.502840        0.726325        0
always          0.726325        1.471274        1
manages         1.555081        2.163455        0
to              2.532826        2.638360        0
find            2.638360        3.032562        0
the             3.032562        3.209488        0
best            3.209488        3.721640        0
deals           3.721640        4.227584        0


Mean Values:

Mean Peak Amplitude: 0.35493
Mean Energy: 123.85982
Mean normalizedEnergy: 440.43621
Mean normalizedBand_Energy: 78533.98150

Word Analysis:

From t=0.57 to 0.72     Peak Amplitude: 0.2148          Energy: 44.3940         Normalised Band Energy: 56302.8905      Is Loud: 0
From t=0.76 to 1.28     Peak Amplitude: 0.5864          Energy: 563.8950        Normalised Band Energy: 217835.5171     Is Loud: 1
From t=1.29 to 1.40     Peak Amplitude: 0.3985          Energy: 38.9488         Normalised Band Energy: 7851.6117       Is Loud: 0
From t=1.57 to 1.80     Peak Amplitude: 0.2639          Energy: 30.8534         Normalised Band Energy: 25667.7260      Is Loud: 0
From t=1.94 to 2.08     Peak Amplitude: 0.2515          Energy: 34.9200         Normalised Band Energy: 50069.9552      Is Loud: 0
From t=2.56 to 2.64     Peak Amplitude: 0.1868          Energy: 18.2439         Normalised Band Energy: 38544.3422      Is Loud: 0
From t=2.76 to 3.03     Peak Amplitude: 0.5680          Energy: 162.0258        Normalised Band Energy: 120122.4712     Is Loud: 0
From t=3.06 to 3.15     Peak Amplitude: 0.2804          Energy: 33.9308         Normalised Band Energy: 73044.1911      Is Loud: 0
From t=3.30 to 3.60     Peak Amplitude: 0.5891          Energy: 301.3667        Normalised Band Energy: 193809.6489     Is Loud: 1
From t=4.07 to 4.19     Peak Amplitude: 0.2101          Energy: 10.0199         Normalised Band Energy: 2091.4612       Is Loud: 0
```

THANK YOU