Niharika V
2018122008

Speech Signalling Processing
Assignment 5
Report

Q3 .Consider a voiced segment and plot its GVV waveform in MATLAB.

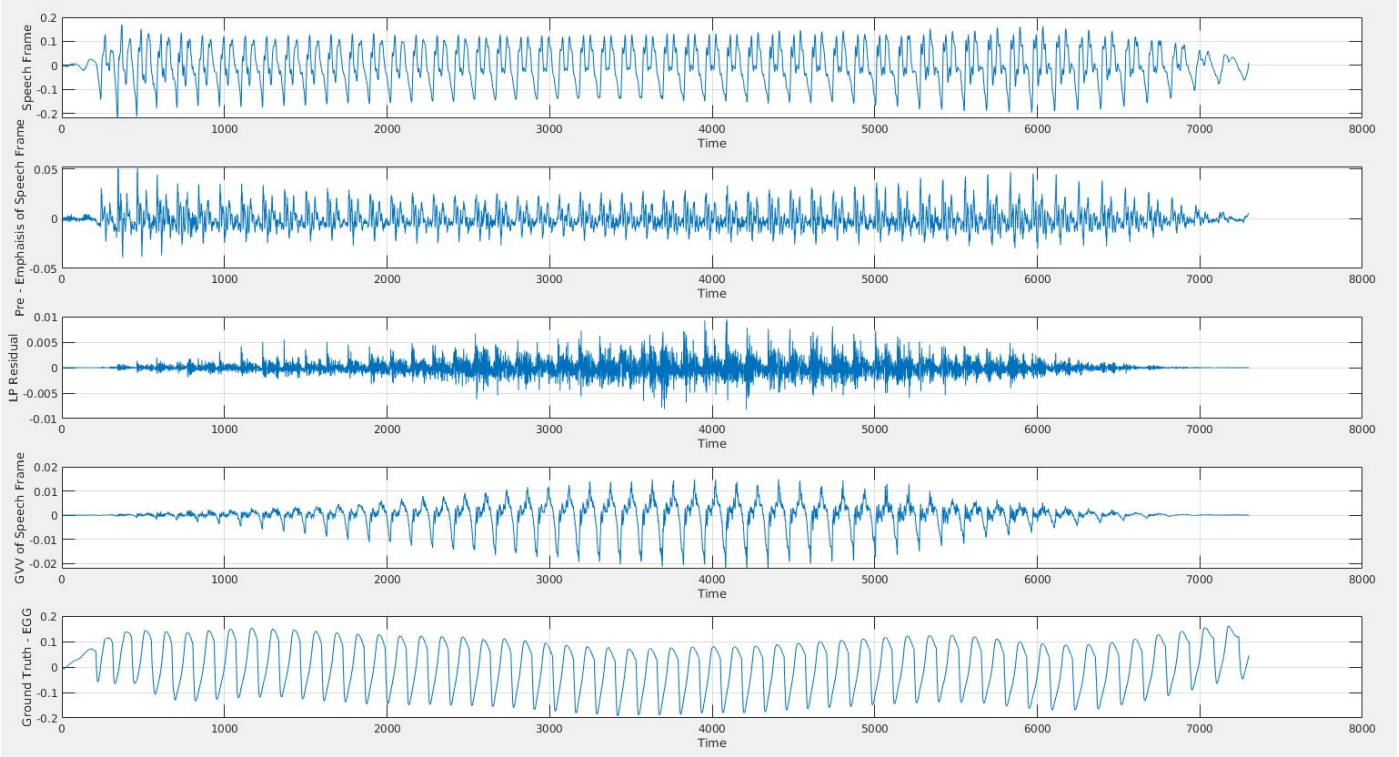
A : Glottal Volume Velocity helps us in understanding the behavior of vocal folds during speech production.

Here, I have taken a voice file from CMU-Arctic Database , and extracted a voiced region -- 'arctic_voiced.wav'. [PFA gvv.m file]

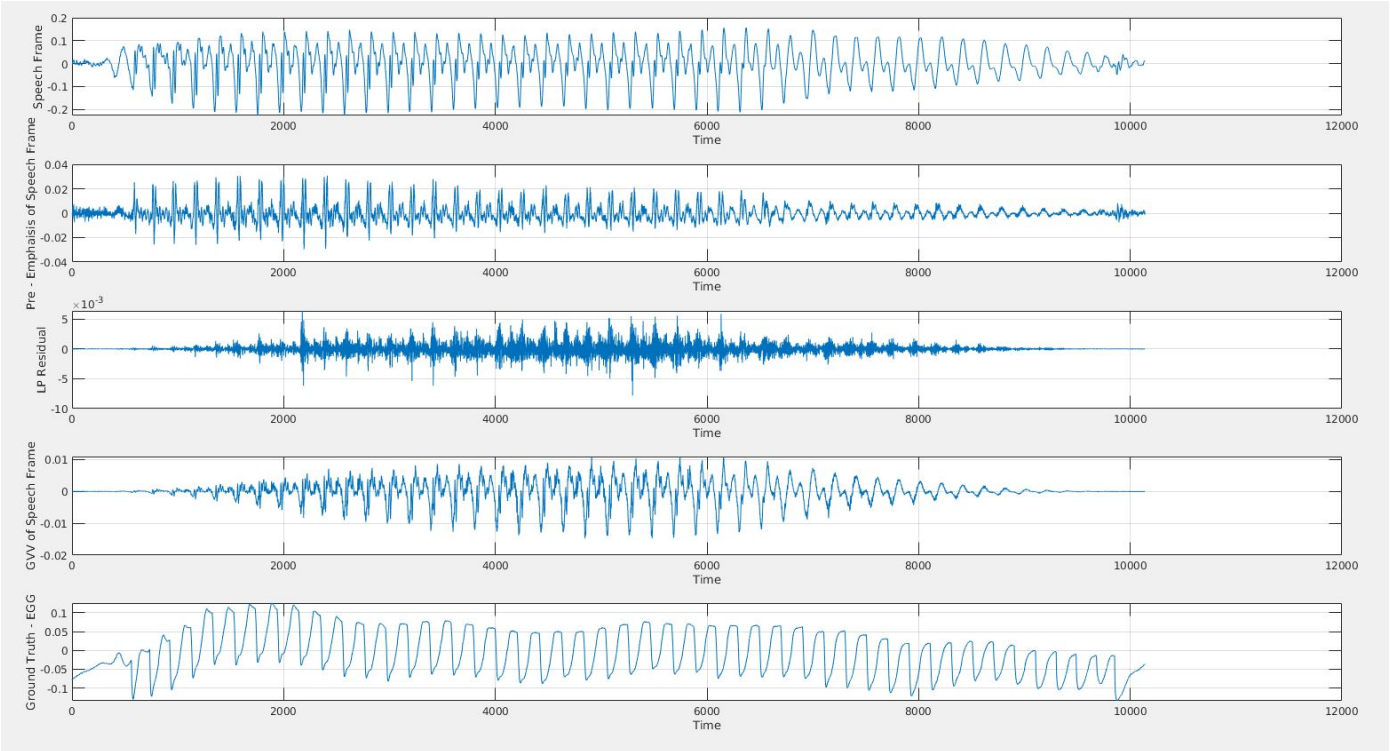
MATLAB Code :

```
function gvv(file)

    close all ;

    flag=1;

    [data, Fs] = audioread(file);

    sig=data(:,1);
    egg=data(:,2);

    % Pre-emphasis of signal
    a=1;
    pre_sig=filter([1, -a], 1, sig);

    Lsig = length(sig);
    OrderLPC =24;   %order of LPC
    out = zeros(size(sig)); % initialization

    Win = hanning(Lsig);  % analysis windo
    sigLPC = Win.*pre_sig;
    en = sum(sigLPC.^2); % get the short - term energy of the input

    % LPC analysis
    r =  xcorr(sigLPC); % correlation
    r = r./max(abs(r));
    [a,g]=lpc(sigLPC,OrderLPC);% LPC coef.
    G = 1 ;% gain
    ex = filter([0 -1*a(2:end)],1,sigLPC); % inverse filter
    % LP Residual Part
    lp_res=sigLPC-ex;
```

Graphs : Case : arctic_voiced.wav



(Case : artic2_voiced.wav)



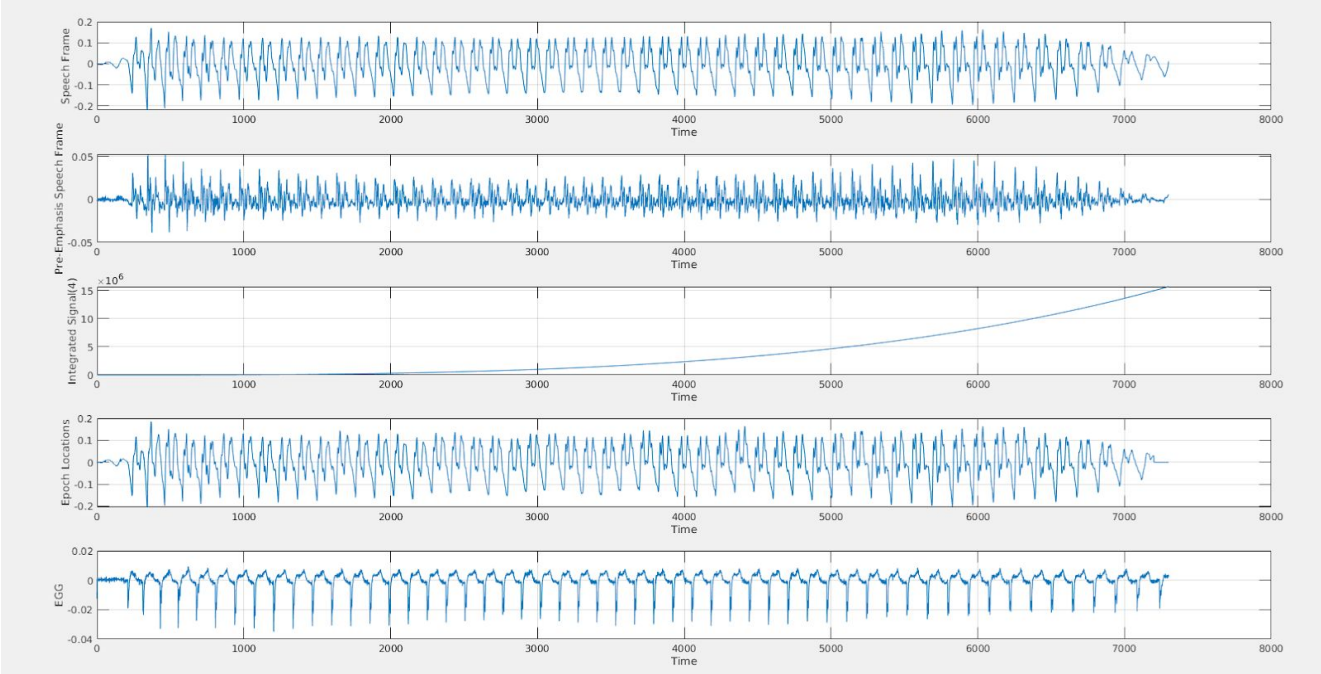
Q2 : Implement the ZFF algorithm in MATLAB and plot the output for one voiced segment .

A: ZFF algorithm is used for epoch extraction , which provides information regarding the excitation source . Here, I've pre-emphasised the voiced signal , and integrated 4 times (2 Cascaded Resonators) , and subtracted the mean-trend of every 10 ms frame of the signal .

MATLAB Code

```
function zff(file)
close all;
Horizon=10;
flag=1;
[data, Fs] = audioread(file);

% Split the spech file --into signal and EGG.
sig=data(:,1);
egg=data(:,2);
% egg=-1*filter([1 -1],1,egg);
Lsig=length(sig);

% Mean-Frame length
Horizon = Horizon*Fs/1000;

% Pre-Emphasis  of the signal
a=1;
pre_emp_sig= filter([1, -a], 1, sig);

% Passing through cascaded Resonator.
x=cumtrapz(pre_emp_sig);
x=cumtrapz(x);
x=cumtrapz(x);
x=cumtrapz(x);

% Mean Trend Removal

slice = 1:Horizon;
Nfr = floor(Lsig/Horizon);
out = zeros(size(sig));
for l=1:Nfr
    frame=sig(slice);
    out(slice)=frame-mean(frame);
    slice = slice+Horizon;
```

Graphs



Q1 . Consider a vowel region in a speech waveform and extract the prosody features.

A : The characteristics that make us perceive lexical and non lexical information such as lexical tone, prominence, accent and emotion are collectively referred to as prosody. Prosodic features include stress, rhythm and intonation. Each cue is a complex perceptual entity, expressed primarily using three acoustic parameters: pitch, energy and duration.

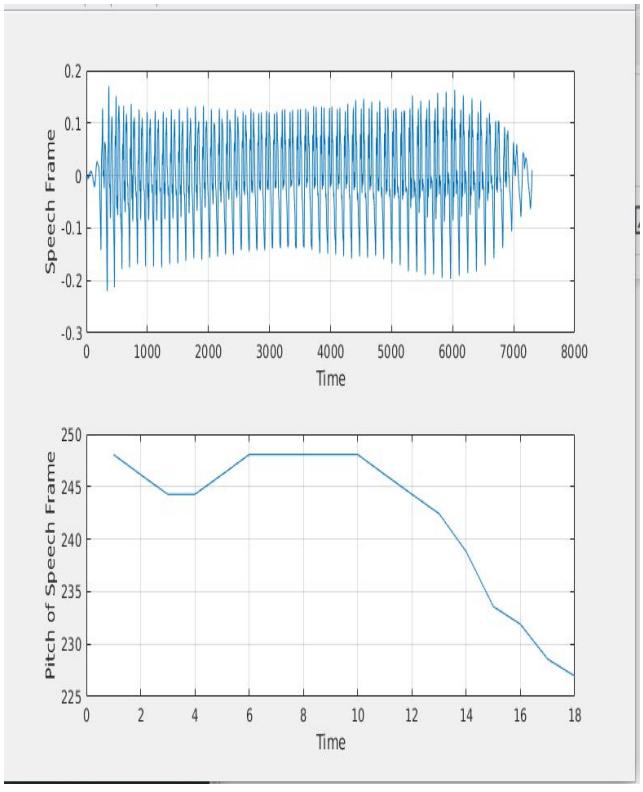
Note : The amplitude tilt,change in F0,etc params can be negative , to signify the slope of the contour is decreasing.

Data File : arctic_voiced.wav [Small Vowel Region]

Code Fie : prosody.m

So, all relevant information is printed out on the command box , also this code is designed for a single vowel detection analysis .

Output : Plot of the speech frame & its contour , Details




```
>> prosody('arctic_voiced.wav')

Intonation Features :
Amplitude Tilt (A): 228.755899
Duration Tilt (D) : 10.500000
Distance of F0 wrt VOP: 2.000000
Change in F0 : 19.490587

Rhytm Features :
Duration of Region : 0.228156
Change in F0 : 19.490587

Stress Features :
Change in Log-Energy : 327.664062
Duration of Region : 0.228156
>>
```