```
% ROLL NO: CB.EN.P2DSC21016
% NAME   : N.SAI DHANUSH
```

1A

```
[y1,Fs]=audioread('arctic_a0016.wav');
sp=y1(:,1);
egg=y1(:,2);
m=resample(sp,8000,Fs);
Fs=8000
```

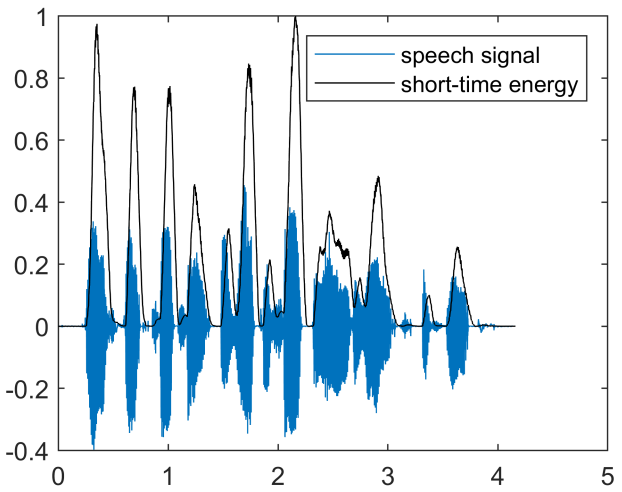
```
Fs = 8000
```

```
t = 0:1/Fs:(length(m)-1)/Fs;
seg = buffer(m,30*10^(-3)*Fs);
seg = buffer(m,481,480);
[l,b] = size(seg);
% hammingw = 0.5 - 0.5*cos(2*pi*[0:l-1]/(l-1));
% hm = repmat(hammingw',1,b);
m_wind = seg;
```

SHORT TIME ENERGY

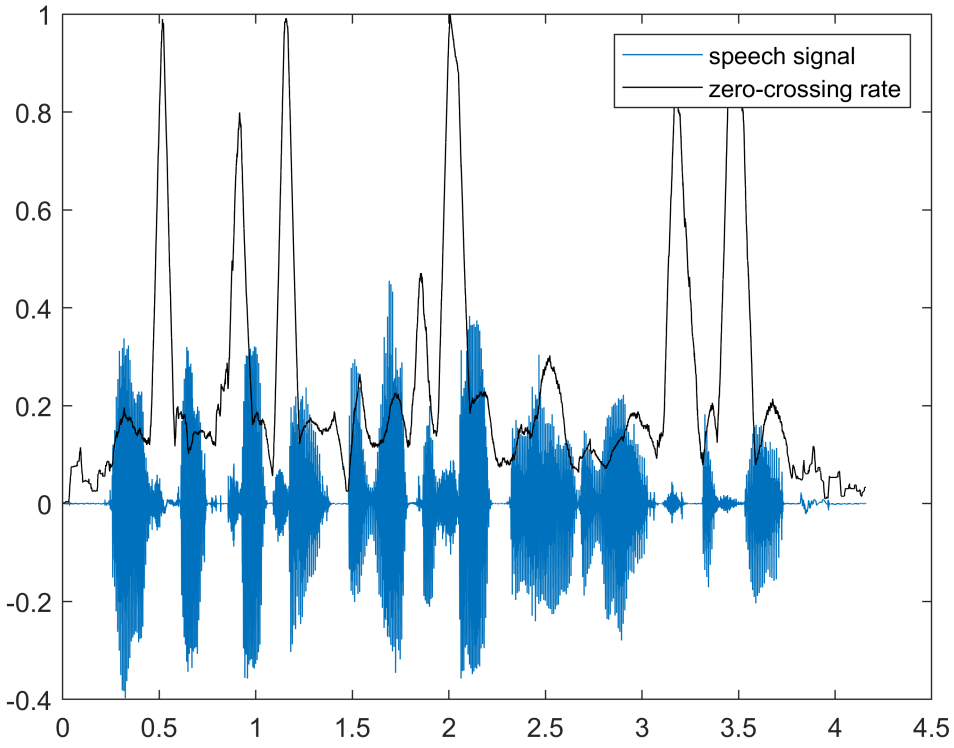
```
% SHORT TIME ENERGY

st_E = sum(m_wind.^2,1);
st_E = st_E / max(st_E);
figure();
plot(t,m);
hold on;
plot(t,st_E(1:length(m)),'k')
legend('speech signal','short-time energy');
```



ZERO CROSSING RATE

```
% ZERO CROSSING RATE
len=l;
st_ZCR = sum(abs(sign(m_wind(1:end-1,:))-sign(m_wind(2:end,:))),1)/2/len;
st_ZCR = st_ZCR/max(st_ZCR);
figure();
plot(t,m);
hold on;
plot(t,st_ZCR(1:length(m)),'k')
legend('speech signal','zero-crossing rate');
```



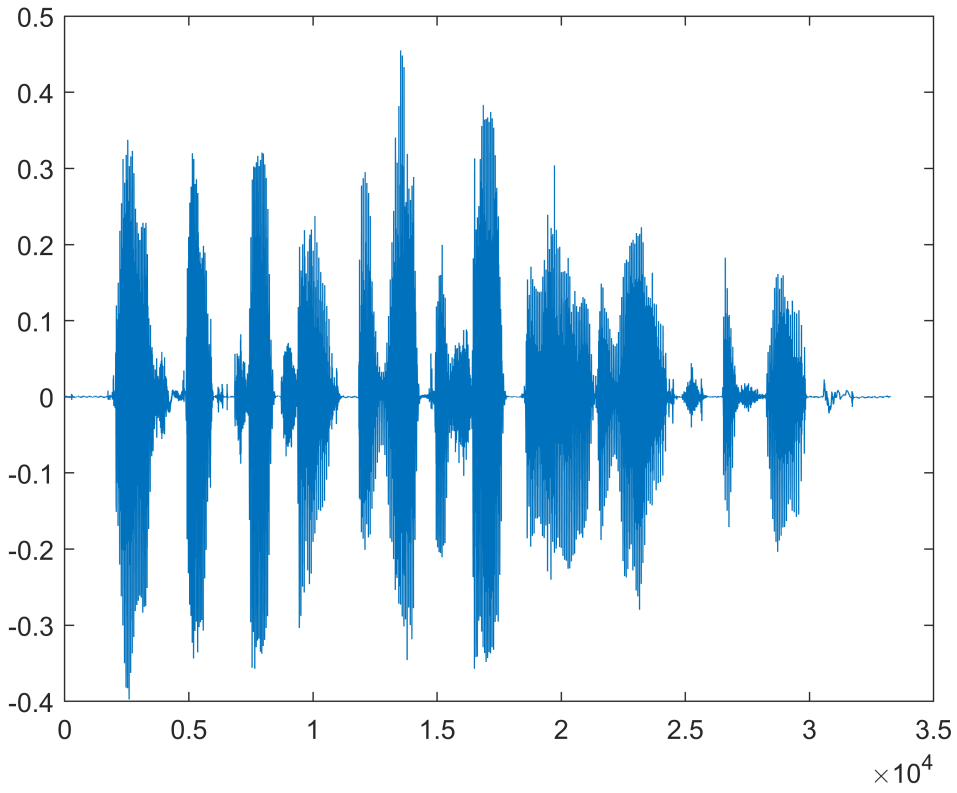
INFERENCE:

By using short time energy analysis method, we can classify voiced and unvoiced regions, where the energy will be high for voiced regions and energy will be less for unvoiced regions, which is clear from the figure.

While using short time zero crossing rate, we can classify voiced and unvoiced regiions clearly, but in this case unvoiced regions have high magnitude, whereas voiced regions have lesser magnitude, which we can clearly infer from the figure obtained, this is because, unvoiced regions are having very less frequency and they oscillate around the x-axis, so thats why we can see laar

1B.

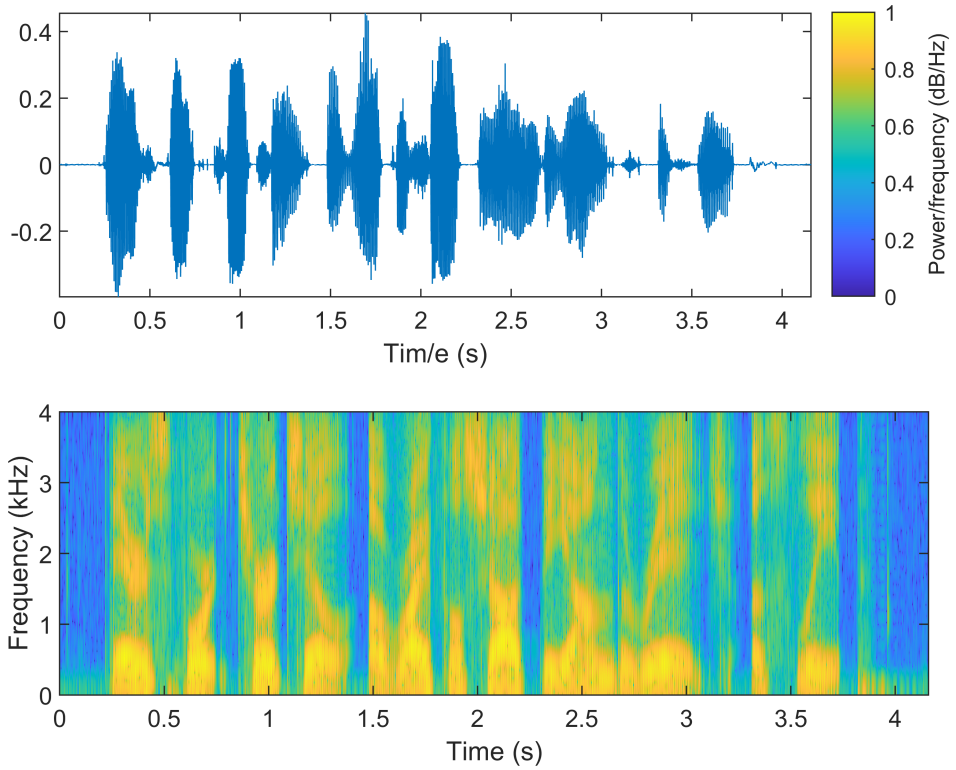
```
% load the audio dataset
[y1,Fs]=audioread('arctic_a0016.wav');
sp=y1(:,1);
s=resample(sp,8000,Fs);
plot(s);
```




```
Fs=8000;
t = 0:1/Fs:length(s)/Fs - 1/Fs;
```

plotting spectogram (WIDEBAND SPECTOGRAM)

```
% window length of 5 milli seconds and step size of 1 milli second
figure; a1=subplot(211);
plot(t,s); xlabel('Tim/e (s)'); axis tight;
a2=subplot(212);
spectrogram(s, 5*10^(-3)*Fs, 1*10^(-3)*Fs, 1024, Fs, 'yaxis')
axis tight;
linkaxes([a1 a2],'x')
```



INFERENCE:

from the figure, it is clear that in voiced regions energy is more,which means we can see vertical stritations which indirectly denoteformant(resonance) frequency, since it is a wideband spectrum it gives good resolution about time, from wideband spectogram we can infer pitch and its formant frequency.

In order to change the charecterstics of spectogram, we can use narrow band spectogram, which will have horizontal stritations, which gives a clearpicture regarding frequency, hence narrowband spectogram gives pitch and its harmonics.