Southern University of Science and Technology

Speech Signal Processing

Lab 3 Report

11510478 郭锦岳

Question 1

Setting L = 21, we can plot the windows, as well as their dB scales as below:

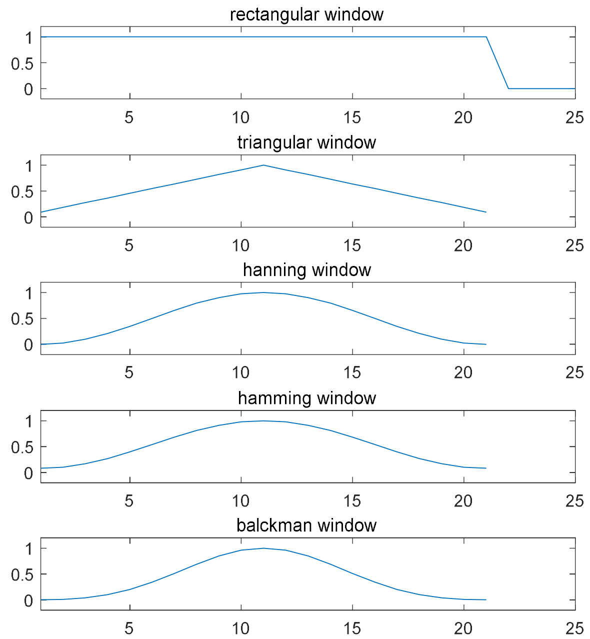


Figure 1

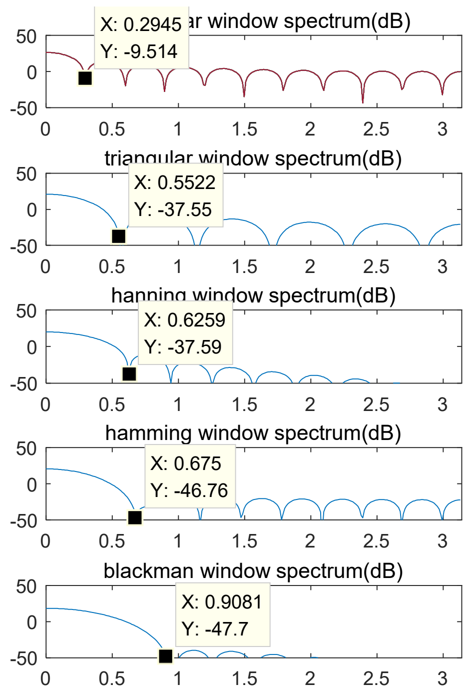


Figure 2

As we can see, the rectangular window has the smallest main-lobe width, while Blackman window has the biggest width. But the hamming window has a better performance of the flatness of main-lobe. Therefore, each window has its own pros and cons, and we need to choose the better window according to our own case.

Question 2

Code:

```
function [  ] = audanalyze( input_file, window_type,
window_shift, window_size)
%AUDANALYZE This function is to calculate and plot figures for
lab3,q2.
signal = audioread(input_file);
R = window_shift;
L = window_size;
figure;
subplot(4,1,1);plot(signal);title('waveform');xlim([0,length(s
ignal)]);
% Use the attribute window type to determine the window
if window_type == 'R'
    window = ones(1,L)';
elseif window_type == 'H'
    window = hamming(L);
end

% short time energy
window_e = window.^2;% effective window for short time energy
stenergy = zeros(1, round((length(signal)-L)/R));
for i = 1:length(stenergy)
    map_i = (i-1)*R+1;% map_i is n hat, beginning of each
small piece
    stenergy(i) = sum((signal(map_i:map_i+L-1).^2).*window_e);
end
subplot(4,1,2);plot(stenergy);title('short-time
energy');xlim([0,length(stenergy)]);

% short time magnitude, almost the same with stenergy
stmagnitude = zeros(1, round((length(signal)-L)/R));
for i = 1:length(stmagnitude)
    map_i = (i-1)*R+1;
    stmagnitude(i)  =  sum(abs(signal(map_i:map_i+L-1)).*window);
end
subplot(4,1,3);plot(stmagnitude);title('short-time
magnitude');xlim([0,length(stmagnitude)]);

% short time zero-crossing rate, using the object to calculate
stzeroc = zeros(1, round((length(signal)-L)/R));
Hzerocross = dsp.ZeroCrossingDetector;
for i = 1:length(stzeroc)
    map_i = (i-1)*R+1;
    stzeroc(i) = step(Hzerocross,signal(map_i:map_i+L-1));
end
subplot(4,1,4);plot(stzeroc);title('short-time zero-
crossing');xlim([0,length(stzeroc)]);

end
```

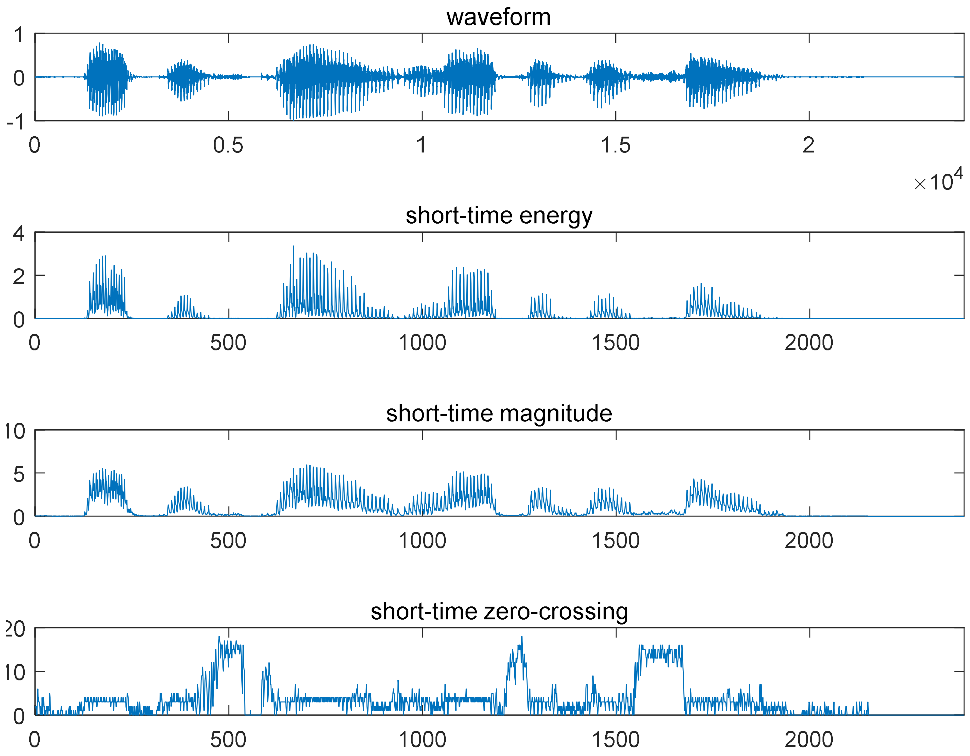


Figure 3 Hamming, L = 21, R = 10

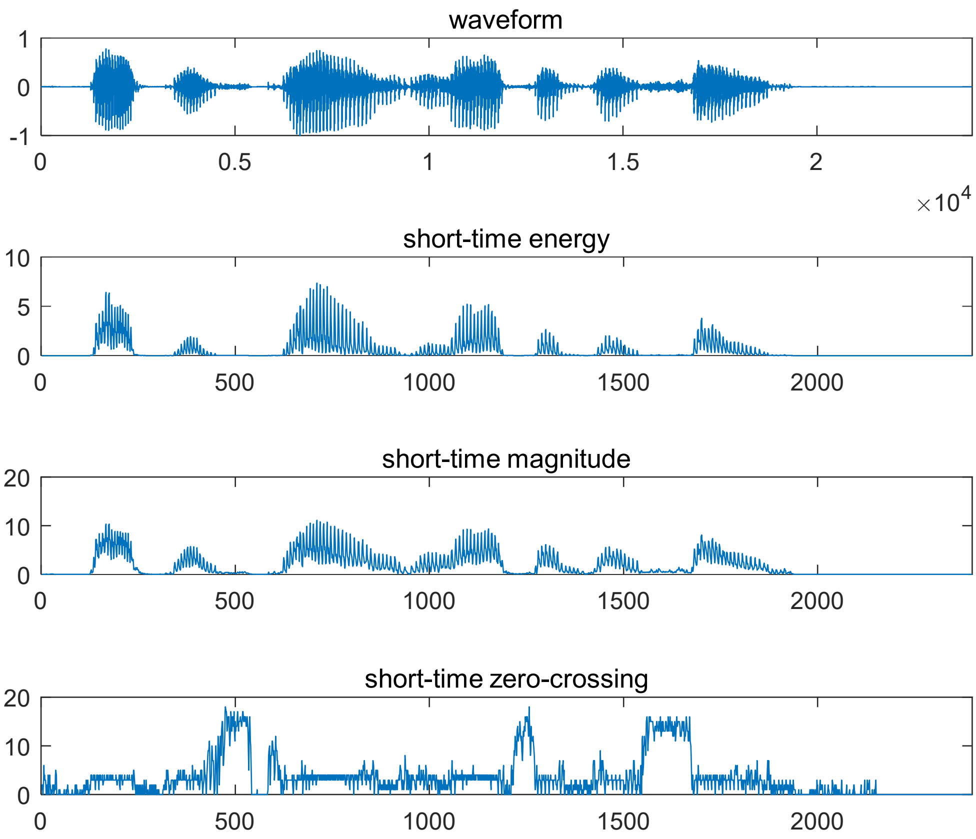
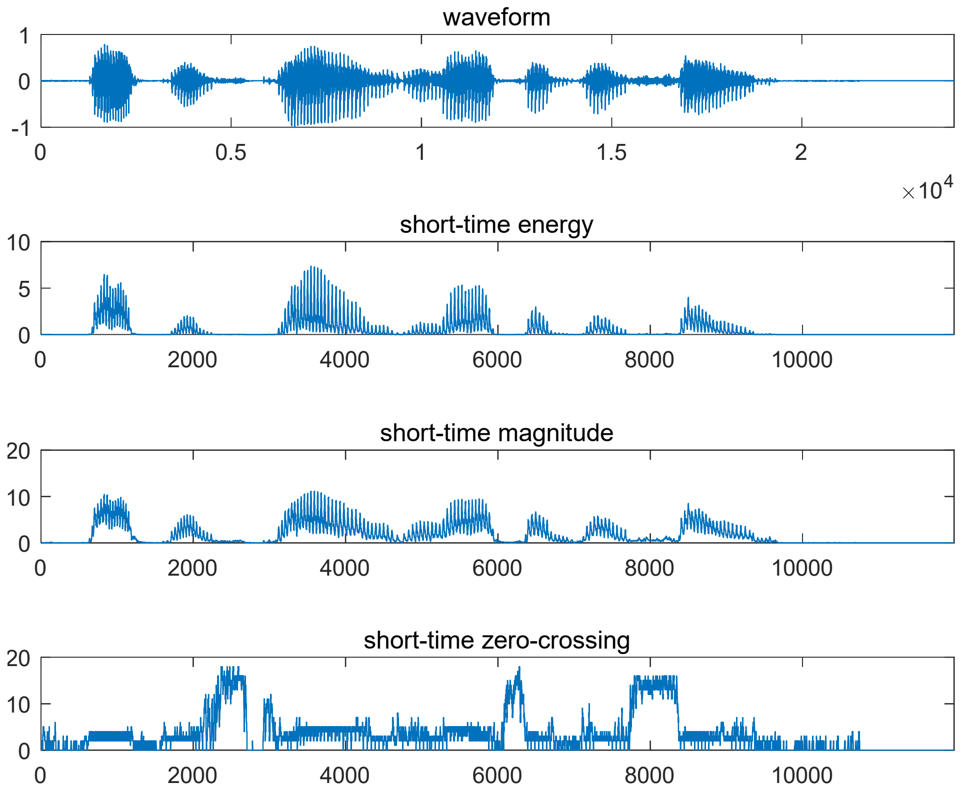


Figure 4 Rectangular, L = 21, R = 10



--

Figure 5 ,Rectangular, L = 21, R = 2

We can see these figures are almost the same, regardless of changing window type or shift length. The only difference for R = 2 is the length of the signal has increased, since the sample rate is higher than R = 10.

Therefore, I suggest that shift length should be larger, for the consideration of saving calculation resource.

Question 3

Code:

```
[signal, Fs] = audioread('test_16k.wav');
R = 5; % window_shift;
L = [51, 101, 201, 401]; %window_size;
figure;

% short-time energy
figure;
for j = 1:4
    stenergy = zeros(1, round((length(signal)-L(j))/R));
    window = hamming(L(j));
    window_e = window.^2;
    for i = 1:length(stenergy)
        map_i = (i-1)*R+1;
        stenergy(i) = sum((signal(map_i:(map_i+L(j)-
1)).^2).*window_e);
    end
    subplot(4,1,j);plot(stenergy);title(sprintf('short-time
energy when L=%d', L(j)));xlim([0,length(stenergy)]);
end

% short time magnitude
for j = 1:4
    stmagnitude = zeros(1, round((length(signal)-L(j))/R));
    window = hamming(L(j));
    for i = 1:length(stmagnitude)
        map_i = (i-1)*R+1;
        stmagnitude(i) = sum(abs(signal(map_i:map_i+L(j)-
1)).*window);
    end
    subplot(4,1,j);plot(stmagnitude);title(sprintf('short-time
magnitude when L=%d',L(j)));xlim([0,length(stmagnitude)]);
end

% short time zero crossing
for j = 1:4
    stzeroc = zeros(1, round((length(signal)-L(j))/R));
    Hzerocross = dsp.ZeroCrossingDetector;
    for i = 1:length(stzeroc)
        map_i = (i-1)*R+1;
        stzeroc(i) = step(Hzerocross,signal(map_i:map_i+L(j)-
1));
    end
    subplot(4,1,j);plot(stzeroc);title(sprintf('short-time
zero-crossing when L=%d',L(j)));xlim([0,length(stzeroc)]);

end
```

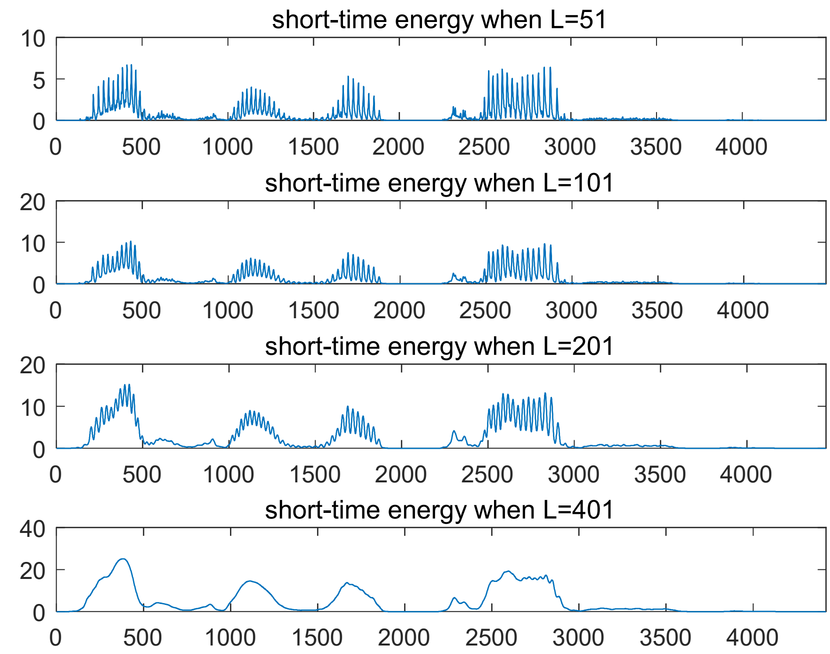


Figure 6

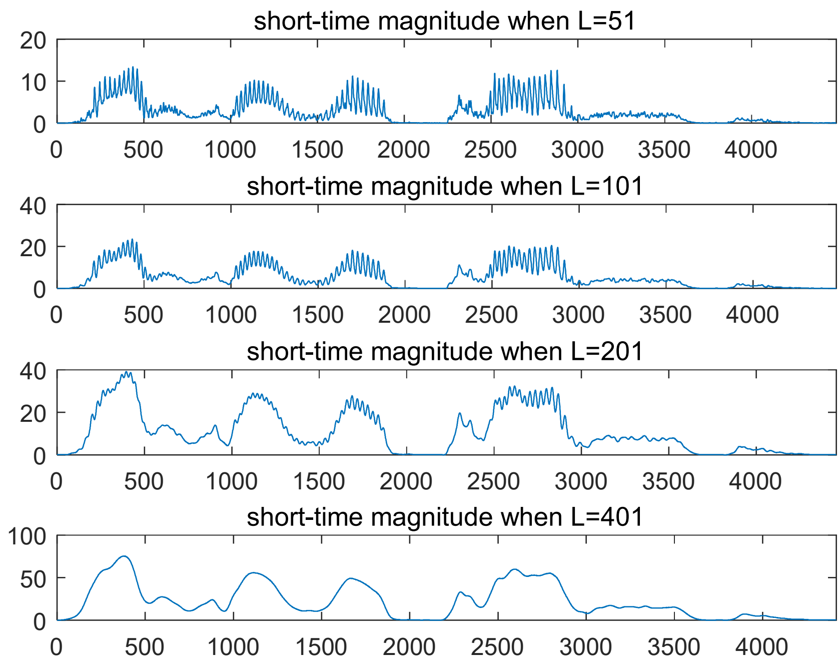


Figure 7

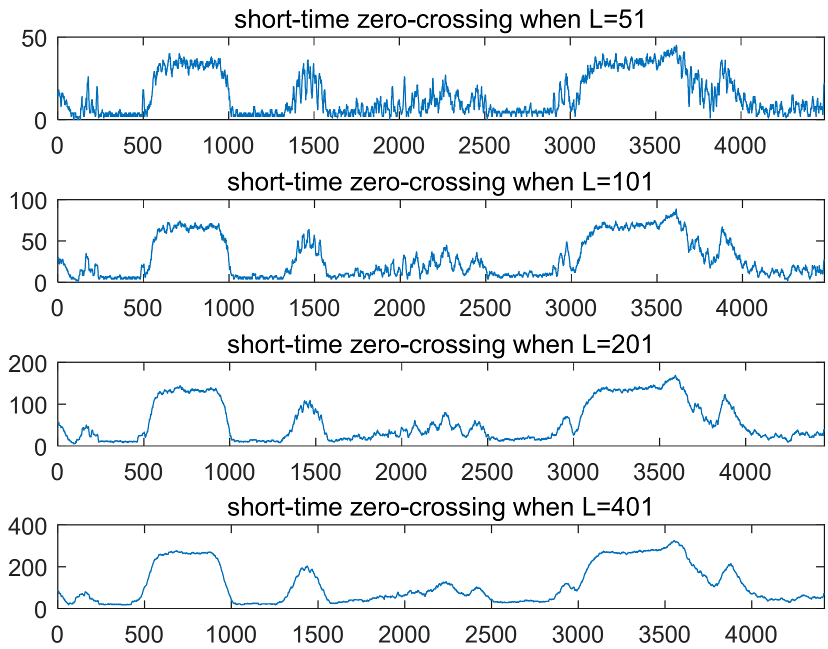


Figure 8

We can see that, different from window type or shift length in Q2, the window length L can significantly change the figure. When L gets bigger, it is equivalent with a low-pass filter with a lower cut-off frequency. Therefore, the high-frequency components are reduced when L increases.

This is either a pro or con, depending on your demand with the operation. If your only wish is to get the envelope, then bigger L is better, since you will not need another low-pass filter to get the envelope. But if the details of the signal is important for you, the smaller window length is the better choice, most of the high-frequency component is remained.