Southern University of Science and Technology

Speech Signal Processing

Lab 6 Report

11510478 郭锦岳

Question 1

In question 1, the task is to resynthesize the signal using autocorrelation method with linear prediction coefficients. This need frame by frame calculation of the coefficients and error signal. Then, according the linear assumption of the vocal system, and the assumption of autocorrelation method that s(m) is only non-zero in 0<m<L-1.

Code:

```
function [ error,synth ] = lab6q1( input_file, frame_size,
frame_shift, order )
%LAB6Q1 use LPC
signal = audioread(input_file);
error = zeros(1,length(signal));
synth = zeros(1,length(signal));
for n_hat = 1:frame_shift:length(signal)-frame_size
    frame = signal(n_hat:n_hat+frame_size-1);
    frame = frame.*hamming(frame_size);% windowing
    %calculate a_k
    a = lpc(frame,order);% p-th order lpc
    a = a(2:end);
    %calulate e(n)
    error_frame = filter([1,a],1,frame);% frame error
    error(n_hat:n_hat+frame_size-
1)=error(n_hat:n_hat+frame_size-1)+error_frame';% adding to
total error
    synth_frame = filter(1,[1,a],error_frame);
    synth(n_hat:n_hat+frame_size-
1)=synth(n_hat:n_hat+frame_size-1)+synth_frame';% adding to
total signal
end

subplot(3,1,1);plot(signal);title('original signal s[n]');
subplot(3,1,2);plot(error);title('error signal e[n]');
subplot(3,1,3);plot(synth);title('resynthesized signal
s_{hat}[n]');
end
```

After generating the frame error and resynthesized signal, we need to use overlap addition to add it on the whole error and resynthesized signal.

Another point to mention is the hamming window of the frame signal. This is to reduce the side effect of the error signal, since the error is larger near m = 0 and m = L.

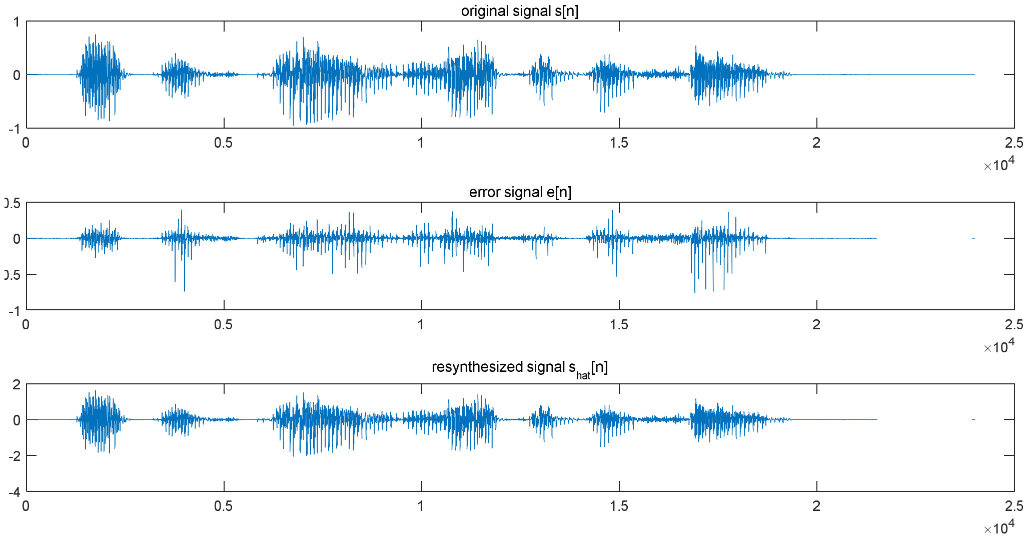

Figure 1, frame size = 320, frame shift = 80, p = 12

The result shows a pretty good similarity between the resynthesized signal and the original one. When listening to the resynthesized one, the words can be clearly recognized, and the tones and voices are remained. The only difference is a little bit saw could be heard.

An interesting fact is that the error signal seems to have more similarity, and when listening to the error signal, it is even clearer than resynthesized signal.

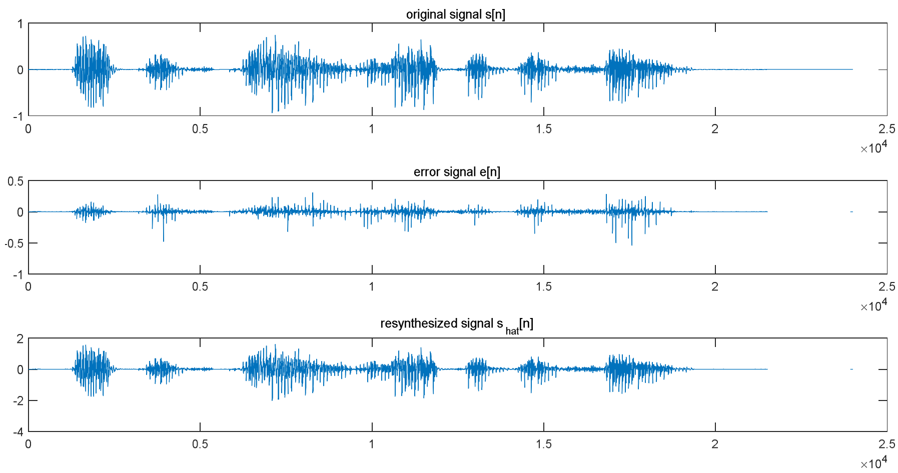

Figure 2 frame size = 320, frame shift = 80, p = 128

But when p increases, the error signal shows less similarity, and the resynthesized signal becomes better.

Question 2

In question2, the task is to compare the spectrum similarity of autocorrelation method. According to page 7 in slide 7, H(z) is the reconstruction of spectrum without excitation signal. Therefore, it should show the same formants with the spectrum obtained by DFT.

Code:

```
function [  ] = lab6q2( input_file, frame_start, frame_size,
order )
%LAB6Q2
signal = audioread(input_file);
frame = signal(frame_start:frame_start+frame_size-1);
%calculate a_k
a = lpc(frame,order);% p-th order lpc
a = a(2:end);
%calulate e(n)
error_frame = filter([1,-a],1,frame);% frame error
spec_frame = 10*log(abs(fft(frame,1024))); % frame spectrum
spec_frame = spec_frame(1:end/2); % get a half
lpc_frame = 10*log(abs(freqz(1,[1,a],1024)));
error_spec = 10*log(abs(fft(error_frame,1024)));
error_spec = error_spec(1:end/2); % adjusting axis
```

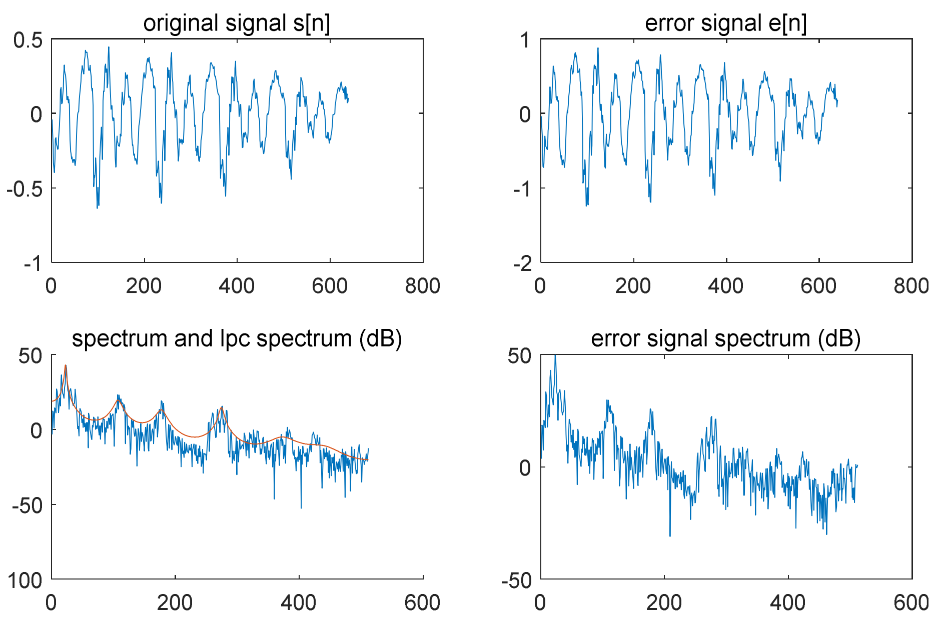

Figure 3 frame = 6000, frame size = 640, p = 12

In the third picture, the blue line is the spectrum of fft, and the orange line is the spectrum of lpc when p = 12. We can see that the formants are in the same position, showing that both method can obtain the same correct answer.

Since the order of the lpc method directly determines the number of poles, which is the same as the number of peaks, increasing the order can have a better result. As shown below, the curves has more same values when the order comes to 64.

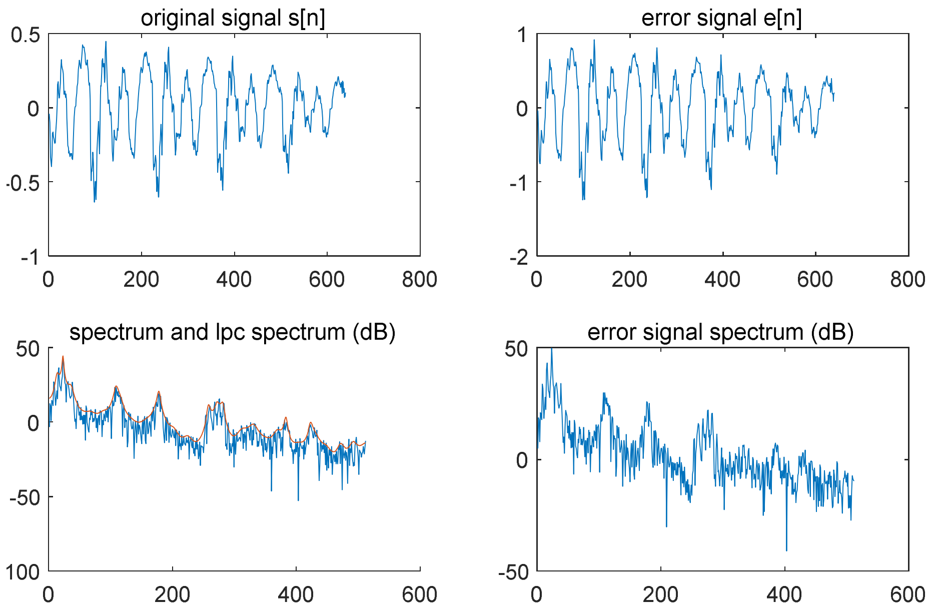

Figure 4 frame = 6000, frame size = 640, p = 64