ECE 438 Digital Signal Processing
Week 13: Speech Processing (Lab 9b)

Date 4/22/2020

Section 2

Name		Sign	Time spent outside lab
David Dang	[%]	David Dang	15
Benedict Lee	[%]	Benedict Lee	24

Grading Rubric (Spring 2020)

	below expectations	lacks in some respect	meets all expectations
Completeness of the report			
Organization of the report *One-sided, with cover sheet, answers are in the same order as questions in the lab, copies of the questions*			
Quality of figures *Correctly labeled with title, x-axis, y-axis, and name(s)*			
Understanding of linear predictive coding (55 pts) *Synthesis of voiced speech: Matlab plots, questions* *Coefficient computation: Matlab code (mylpc)*			
Understanding of speech coding and synthesis (45 pts) *Matlab plots and code, questions*			

ECE 438 Lab 9b

David Dang & Benedict Lee

1.2 Synthesis of Voiced Speech

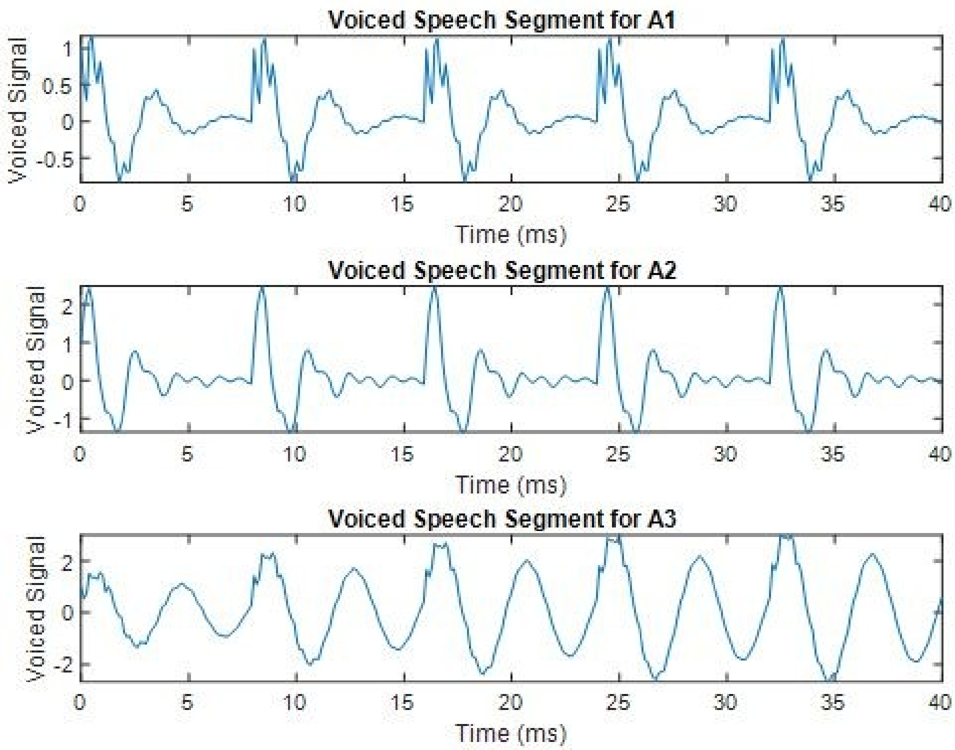
INLAB REPORT: Hand in the following:

• A figure containing the three time-domain plots of the voiced signals.

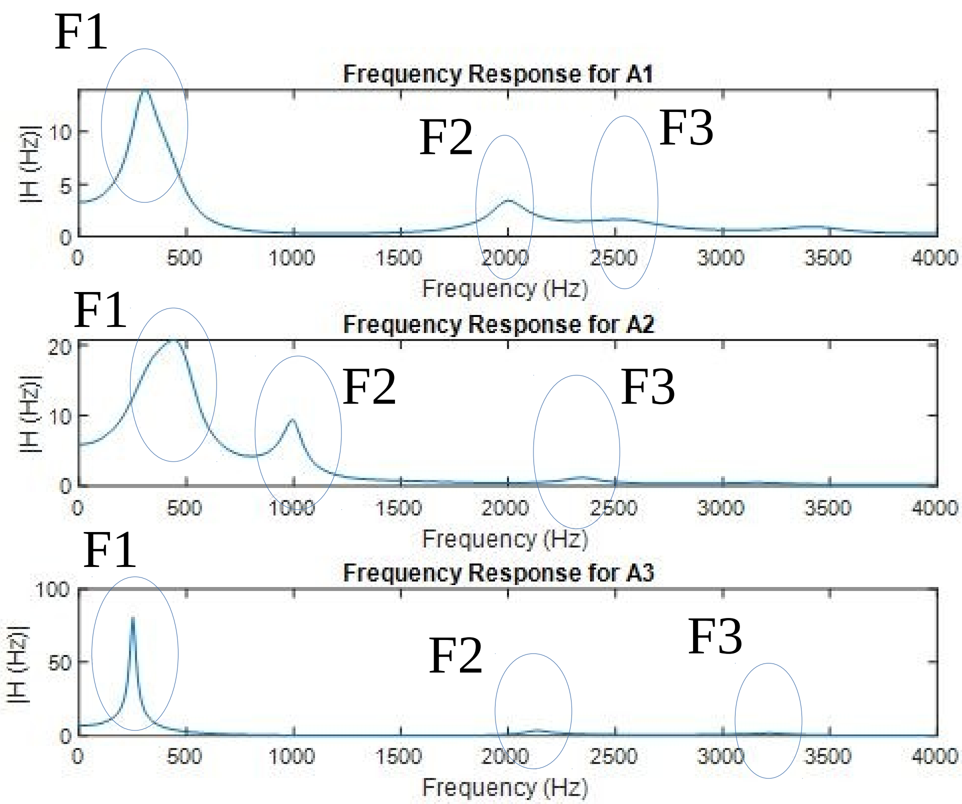
• Plots of the frequency responses for the three filters. Make sure to label the frequency axis in units of Hertz.

• For each of the three filters, list the approximate center frequency of the first three formant peaks.

• Comment on the audio quality of the synthesized signals.



- The synthesized signals all sound across as buzzing/humming sounds, with that for A1 sounding the most high-pitched, followed by that for A2, then lastly that for A3 which is lowest-pitched. This is corroborated by the 3 frequency response plots above, whereby the synthesized signal for A1 has the most high-frequency components, followed by that for A2, then lastly that for A3 which has the least high-frequency components. Comparing the center frequencies of the first three formant peaks with the average formant frequencies for the vowels in the formant table, A1 corresponds to the vowel sound ‘I’, A2 corresponds to ‘U’, and A3 corresponds to ‘IY’.



	F1 Formant Peak Center Frequency (Hz)	F2 Formant Peak Center Frequency (Hz)	F3 Formant Peak Center Frequency (Hz)
A1 Filter Coefficient Set	304	2000	2520
A2 Filter Coefficient Set	440	992	2340
A3 Filter Coefficient Set	250	2140	3225

2.3 LPC Exercise

INLAB REPORT: Hand in your mylpc function.

```
function coef = mylpc(x, P)
rhat = xcorr(x, P);
rhat = rhat(P+1:end);
rs = rhat(2:end);
Rs = toeplitz(rhat(1:P));
coef = rs/Rs;
```

3 Speech Coding and Synthesis

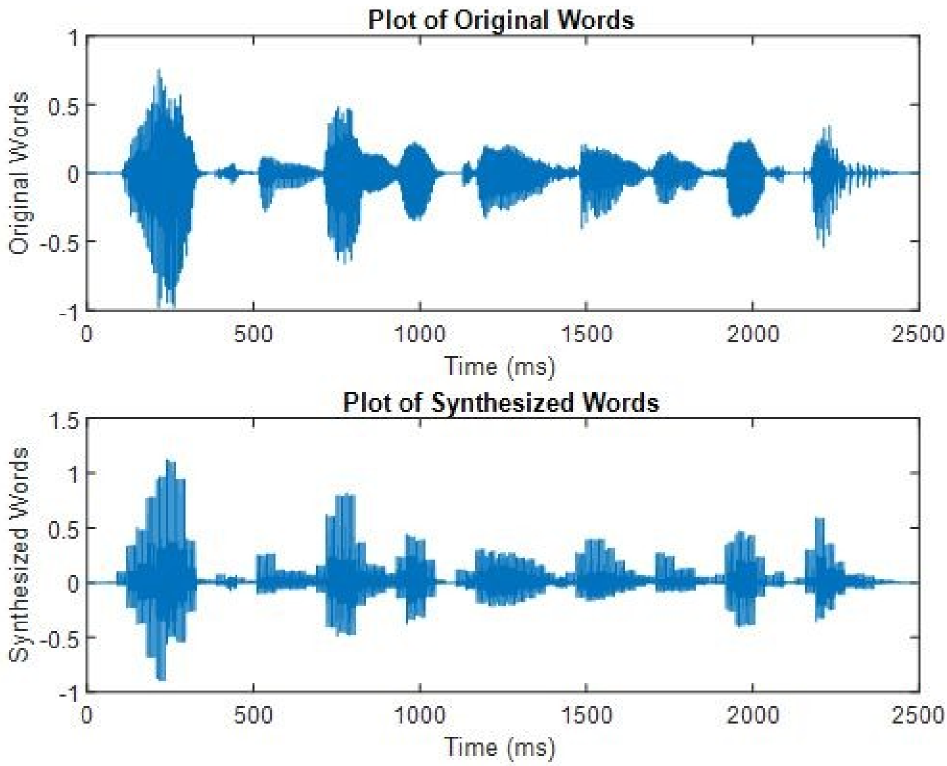
INLAB REPORT: Hand in the following:

• Your analysis and synthesis code.

• The compression ratio.

• Plots of the original and synthesized words.

• Comment on the quality of your synthesized signal. How might the quality be improved?



- The synthesized signal sounds robotic as opposed to the relatively human-sounding original signal. The quality might be improved by shortening the timespan of the non-overlapping frames so that the filter coefficients are more unchanging, as well as increasing the order of the LPC coefficients for each frame.

- compression_ratio = 0.0708

analysis and synthesis code

```
orig_vec = audioread('phrase.au');
rows = 30*10^(-3)*8000;
cols = floor(length(orig_vec)/rows);
phrase_vec = reshape(orig_vec(1:rows*cols), rows, cols);
energy = zeros(1, cols);
VU = zeros(1, cols);
A = zeros(15, cols);
for i=1:cols
    energy(i) = sum(phrase_vec(1:end, i).^2);
    VU(i) = (zero_cross(phrase_vec(1:end, i)) < (rows/2));
    A(1:end, i) = mylpc(transpose(phrase_vec(1:end, i)), 15);
end
A(isnan(A))=0;

Awhos = whos('A');
VUwhos = whos('VU');
energywhos = whos('energy');
phrasewhos = whos('phrase_vec');
compression_ratio = (Awhos.bytes+VUwhos.bytes+energywhos.bytes)/phrasewhos.bytes;

out_vec = zeros(rows, cols);
vexcite = exciteV(rows, 7.5/0.125);
uexcite = exciteUV(rows);
for i=1:cols
    if VU(i) == 1
        temp = filter(1, [1 -transpose(A(1:end, i))], vexcite);
    else
        temp = filter(1, [1 -transpose(A(1:end, i))], uexcite);
    end
    energizer = sum(temp.^2);
    temp = transpose(temp * sqrt(energy(i)/energizer));
    out_vec(1:end, i) = temp;
end
out_vec = reshape(out_vec, rows*cols, 1);
temp = (length(orig_vec)-1)*0.125;
norig = linspace(0, temp, temp/0.125+1);
temp = (length(out_vec)-1)*0.125;
nout = linspace(0, temp, temp/0.125+1);
figure(1)
subplot(2,1,1)
plot(norig, orig_vec)
title('Plot of Original Words')
xlabel('Time (ms)')
ylabel('Original Words')
subplot(2,1,2)
plot(nout, out_vec)
title('Plot of Synthesized Words')
xlabel('Time (ms)')
ylabel('Synthesized Words')
%soundsc(orig_vec);
%soundsc(out_vec);
```