ECE- 6255

JEET KIRAN PAWANI
GTid- 903397407

DIGITAL PROCESSING OF SPEECH SIGNALS ASSIGNMENT – 6

1) 9.1 Rabiner

$$\tilde{h}[n] = \sum_{k=1}^{p} \alpha_k \tilde{h}[n-k] + G\delta(n)$$

Autocorrelation func of $\tilde{h}(n)$

$$\tilde{R}(m) = \sum_{n=0}^{\infty} \tilde{h}[n]\tilde{h}(n+m)$$

a) To show $\tilde{R}[m] = \tilde{R}[-m]$

we know $\tilde{R}(m) = \sum_{n=0}^{\infty} \tilde{h}[n]\tilde{h}(n+m)$

$= h[m] * h(-m)$ convolution form

$m = -m$ in $\tilde{R}(m)$ eqn

$\Rightarrow \tilde{R}[m] = h(-m) * h(m)$

we know that convolution is commutative from properties of convolution.

$\Rightarrow \tilde{h}(-m) * \tilde{h}(m) = \tilde{h}(m) * \tilde{h}(-m)$

$$\Rightarrow \boxed{\tilde{R}(m) = \tilde{R}(-m)}$$

b) $\tilde{R}(-m) = \sum_{n=0}^{\infty} \tilde{h}(n)\tilde{h}(n-m)$

since $\tilde{R}(-m) = \tilde{R}(m)$

$\Rightarrow \tilde{R}(m) = \sum_{n=0}^{\infty} \tilde{h}(n)\tilde{h}(n-m)$

$$\tilde{h}(n) = \sum_{k=1}^{p} \alpha_k \, \hat{h}(n-k) + G\, d(n)$$

Substituting

$$\tilde{R}(m) = \sum_{n=0}^{\infty} \left[\left[\sum_{k=1}^{p} \alpha_k \hat{h}(n-k) + G\, d(n)\right] \left[\sum_{j=1}^{p} \alpha_j \tilde{h}(n-j-m) + G\, d(n-m)\right]\right]$$

different variables for both
as m is movement factor

$$= \sum_{n=0}^{\infty} \Big[\sum_k \sum_j \alpha_k \alpha_j \tilde{h}(n-k)\, \hat{h}(n-m-j)$$
$$+ \sum_{k=1}^{p} \alpha_k \tilde{h}(n-k)\, G\, d(n-m)$$
$$+ \sum_{j=1}^{p} \alpha_j \hat{h}(n-m-j)\, G\, d(n)$$
$$+ G^2 d(n)\, d(n-m) \Big]$$

There can be 2 possibilities for m either $m<0$ or $m \geq 0$

Assume $m \geq 0 \Rightarrow d(n)\, d(n-m) = 0$ as $d(n)$ is 1 at only $n=1$

$$\Rightarrow \tilde{R}(m) = \sum_{n=0} \sum_k \sum_j \alpha_k \alpha_j \hat{h}(n-k)\, \hat{h}(n-m-j)$$
$$+ \sum_n \sum_k \alpha_k \hat{h}(n-k)\, G\, d(n-m)$$
$$+ \sum_n \sum_j \alpha_j \tilde{h}(n-m-j)\, G\, d(n)$$

$$\boxed{\sum_n d(n) = 1}$$

$\hat{h}(n-j)\, d(n-m)$ can be written as $\hat{h}(n+m-j)\, d(n)$

$$\Rightarrow \tilde{R}(m) = \sum_n \sum_k \sum_j \alpha_k \alpha_j \tilde{h}(n-k)\tilde{h}(n-m-j)$$ (2)

$$+ \sum_k \alpha_k \tilde{h}(m-k)\, G$$

$$+ \sum_j \alpha_j \tilde{h}[-m-j]\, G$$

as $m \geq 0$, $\tilde{h}(-m-j) = 0$

$$\tilde{R}(m) = \sum_n \sum_k \sum_j \alpha_k \alpha_j \tilde{h}(n-k)\tilde{h}(n-m-j)$$

$$+ G \sum_n \sum_k \alpha_k \tilde{h}(m-k)\,\delta(n)$$

added $\sum_n \delta(n)$ here

$$= \sum_n \sum_k \alpha_k \tilde{h}(n-k) \underbrace{\left[\sum_j \alpha_j h(n-m-j)\right] + G\,\delta(n-m)}_{\tilde{h}(n-m)}$$

$$= \sum_n \sum_k \alpha_k \tilde{h}(n-k)\tilde{h}(n-m)$$

Substitute $n = n^{\#} + m$

$$\Rightarrow \tilde{R}(m) = \sum_k \alpha_k \sum_{n^{\#}=-m}^{\infty} \tilde{h}(n^{\#}+m-k)\tilde{h}(n^{\#})$$

we know $\tilde{h}(n^{\#})$ causal $\rightarrow$ $n^{\#}$ ranges from 0 to ∞

$$\tilde{R}(m) = \sum_k \alpha_k \underbrace{\sum_{n^{\#}=0}^{\infty} \tilde{h}(n^{\#}+m-k)\tilde{h}(n^{\#})}_{\tilde{R}(m-k)}$$

$$\Rightarrow \tilde{R}(m) = \sum_k \alpha_k \tilde{R}(m-k)$$

Case when $m \leq 0$

we will get $\tilde{R}(m) = \sum_{k=1}^{p} \alpha_k \tilde{R}(k+m)$

$\Rightarrow \tilde{R}(m)$ can be written as $\sum_k \alpha_k \tilde{R}[|m-k|]$ $m=1 \dots p$

$$\Rightarrow \boxed{\tilde{R}(m) = \sum_{k=1}^{p} \alpha_k \tilde{R}[|m-k|] \quad m=1 \dots p}$$

2] 9.6 Rabiner

$F_s = 8000$ samples/sec

Segment size = 300 samples

linear pred error filters

$$A^{(i)}(z) = 1 - \sum_{k=1}^{i} \alpha_k^{(i)} z^{-k} \quad i=1 \text{ to } i=4$$

Table is formed by Levinson - Durbin

a) $A^{(4)}(z)$ determine. We can get this from the table in row 4.

$$\boxed{A^{(4)}(z) = 1 - 0.8047z^{-1} + 0.0414z^{-2} - 0.4940z^{-3} + 0.4337z^{-4}}$$

Plot on matlab

b) k-parameters of 4th order prediction error lattice filter

This can be obtained by last coeff on each row.

$\Rightarrow k_1 = 0.8328$ (Row 1)

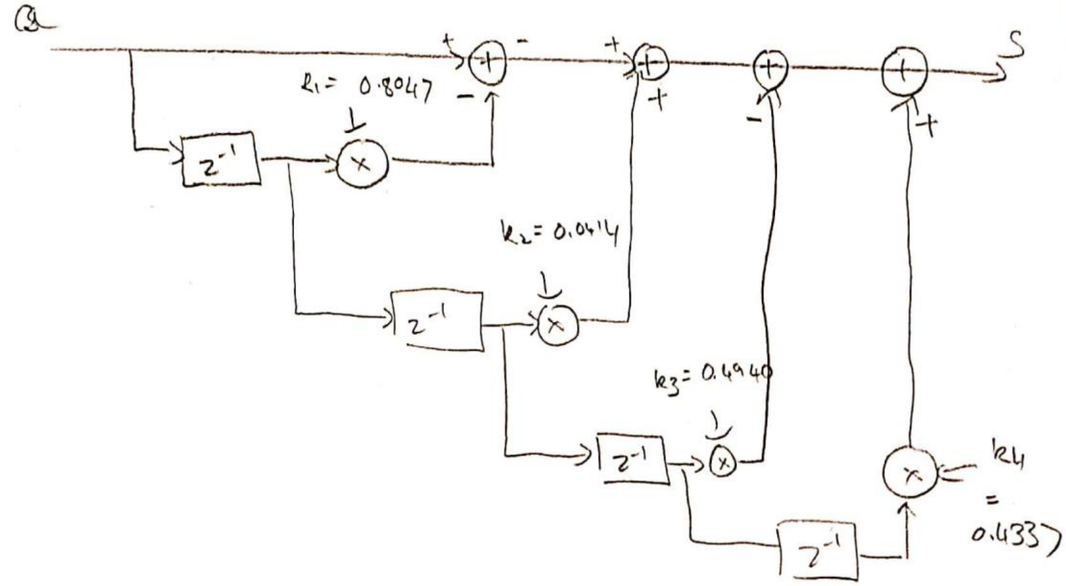
$k_2 = 0.1044$ (Row 2 last element)

$k_3 = 0.1786$ (Row 3 last element)

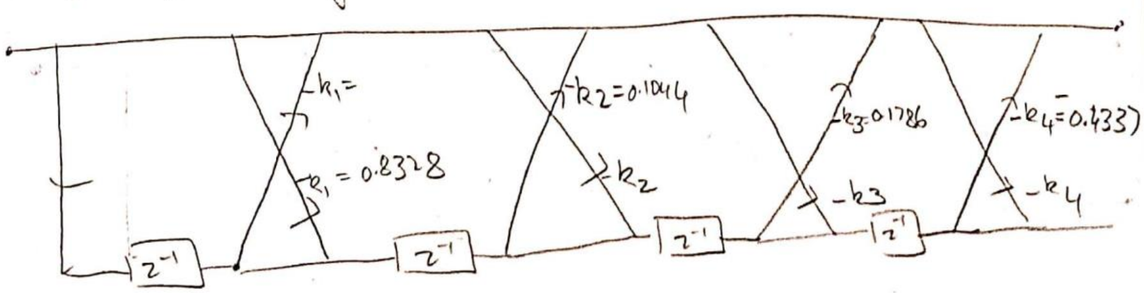
$k_4 = -0.4337$ (Row 4 last element)

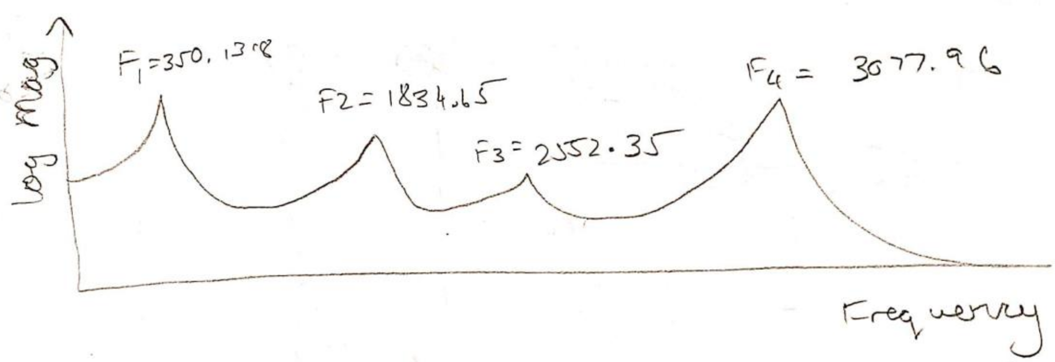
Jeet Kiran Pawani || Assignment -6 || ECE – 6255 || GTid - 903397407

2a) Direct form



2b) Lattice form





c) Prove

$A^{(4)}(z) = A^{(3)}z - k_4 z^{-4} A^{(3)}(z^{-1})$ ③

From table,

$A^{(3)}(z) = 1 - 0.7273z^{-1} + 0.0289z^{-2} - 0.1786z^{-3}$ → ①

$A^{(3)}(+z^{-1}) = 1 - 0.7273z^{+1} + 0.0289z^{2} - 0.1786z^{4}$ → ④

$k_4 = -0.4337$ from ⓑ part

$-k_4 z^{-4} = +0.4337 z^{-4}$

$-k_4 z^{-4} A^{3}(z^{-1}) = 0.4337z^{-4}(1 - 0.7273z^{+1} + 0.0289z^{2} - 0.1786z^{4})$

$= 0.4337z^{-4} - 0.3154z^{-3} + 0.0125z^{-2} - 0.0774z^{-1}$ → ②

$A^{4}(z) = A^{(3)}(z) - k_4 z^{-4} A^{3}(z^{-1}) = ① + ②$

$= 1 - 0.7273z^{-1} + 0.0289z^{-2} - 0.1786z^{-3}$
$-0.0774z^{-1} + 0.0125z^{-2} - 0.3154z^{-3} + 0.4337z^{-4}$

$$\boxed{A^{4}(z) = 1 - 0.8047z^{-1} + 0.0414z^{-2} - 0.494z^{-3} + 0.4337z^{-4}}$$

This is the same that we got in part ⓐ

d) $E^{(2)} = 0.5803$

to find $E^{(4)}$ $R[0]$ and $R[1]$

From Levinson durbin method

$$E^{(i)} = (1 - k_i^2) E^{(i-1)}$$

$$\Rightarrow E^{(4)} = (1 - k_4^2) E^3$$

$$= (1 - k_4^2)(1 - k_3^2) E^{(2)}$$

from part (b)

$k_4 = -0.4337 \quad k_3 = 0.1786$

$$= (1 - (-0.4337)^2)(1 - (0.1786)^2)(0.5803)$$

$$= 0.8119 \times 0.9681 \times 0.5803$$

$$\boxed{E^{(4)} = 0.4561}$$

We know $E^{(0)} = R[0]$

$$E^{(2)} = (1 - k_2^2) E^{(1)} = (1 - k_2^2)(1 - k_1^2) E^{(0)}$$

$$= (1 - k_2^2)(1 - k_1^2) R(0)$$

$k_2 = 0.1044 \quad k_1 = 0.8328$

part (b)

$$R(0) = \frac{E^{(2)}}{(1 - k_2^2)(1 - k_1^2)} = \frac{0.5803}{(0.9891)(0.3064)}$$

$$\boxed{R(0) = 1.91452}$$

(4)

$R(1) = k_1 R(0)$

$= 0.8328 \times 1.91452$

$\boxed{R(1) = 1.59441}$

e) we know $E^{(i)} = (1 - k_i^2) E^{(i-1)}$

we see a big reduction in E only when k is sufficiently near 1

from equation when $k \approx 1$ then $k^2 \approx 1$

the next order energy is almost close to 0

So, depending value of ki we can say E changes

f) 11 pole filter.

Given five poles are complex with real values means other 5 poles are its real complex conjugates.

All these poles are within the unit circle seen from $|z_i|$ value.

Thus for the 11^{th} pole, we can assume that the pole is real, within unit circle and we can say anything of the sign or where it will lie within the circle.

g) we know that $G^2 = E$
here it is a 11 pole filter

$\Rightarrow G^2 = E^{(11)}$

$E^{(11)} = (1 - k_{11}^2) E^{10}$

$= \prod_{i=1}^{11} (1 - k_i)^2 E^0$ $\quad E^{(0)} = R(0)$

$= \prod_{i=1}^{11} (1 - k_i)^2 R(0)$

$= (1-k_1^2)(1-k_2^2)(1-k_3^2)(1-k_4^2)(1-k_5^2)(1-k_6^2)(1-k_7^2)(1-k_8^2)(1-k_9^2)(1-k_{10}^2)(1-k_{11}^2) R(0)$

From table we can find k_i values

$k_1 = 0.8328$ $k_2 = 0.1044$ $k_3 = 0.1786$ $k_4 = -0.4337$

$k_5 = -0.0978$ $k_6 = -0.7505$ $k_7 = -0.0067$ $k_8 = -0.0713$

$k_9 = 0.5605$ $k_{10} = 0.1323$ $k_{11} = 0.0371$

$E^{(11)} = (0.3064)(0.9891)(0.9681)(0.8119)(0.9904)(0.4367)(0.99995)(0.9949)(0.6858)(0.9824)(0.998)(1.91452)$

$= 0.13194$

$G^2 = 0.13194$

$G = 0.3632$

1) Formant frequency $= \frac{(\angle z)\, F_s}{(2\pi)}$

(5)

$\lvert z_i \rvert$	$\angle z_i$	F(in Hz)	Formant
0.2567	2.0677	2634.0	no (away from circle)
0.9681	1.4402	1834.65	F2
0.9850	0.2750	350.318	F1
0.8647	2.0036	2552.35	F3
0.9590	2.4162	3077.96	F4

The formants are approximately closest to unit circle As seen in the table above F1 is given by pole 3 Formant 2 is given by pole 2, F3 is given by pole 4 and last formant is given by pole 5

i) Plot on matlab

3] 9.18 Rabiner

$$\mathcal{E}(\beta, N_p) = \sum_m \left(s(m) - \beta s(m-N_p)\right)^2$$

β = predictor coefficients

N_p = candidate pitch period.

a) $$\mathcal{E}(\beta, N_p) = \sum_m \left(s(m) - \beta\, s[m-N_p]\right)^2$$

$$= \sum_m \left(s^2(m) - 2\beta s(m)\, s(m-N_p) + \beta^2 s^2(m-N_p)\right)$$

$$= \left[\sum_m s^2(m) - 2\beta \sum_m s(m)\, s(m-N_p) + \beta^2 \sum_m s^2(m-N_p)\right]$$

we know $\frac{\partial E(\beta, N_p)}{\partial \beta} = 0$ for minimum

$$\Rightarrow \frac{\partial E(\beta, N_p)}{\partial \beta} = 0 - 2\sum_m s(m)\, s(m-N_p) + 2\beta \sum_m s^2(m-N_p)$$

$$0 = -2\sum_m s(m)\, s(m-N_p) + 2\beta \sum_m s^2(m-N_p)$$

$$\beta = \frac{\sum_m s(m)\, s(m-N_p)}{\sum_m s^2(m-N_p)}$$

here N_p is fixed

b) To solve for optimal N_p, we substitute the value of β and find max. of P

$$N_p^{opt} = \max_{N_p} \beta$$

$$\boxed{N_p^{opt} = \max_{N_p} \frac{\sum_m s[m]\, s[m-N_p]}{\sum_m s^2(m-N_p)}}$$

c) More computation is required in this method but this provides higher accuracy of pitch (N_p value) for higher pitches. In case of smaller peaks, it is difficult to differentiate the pitch peak in auto correlation func. from peaks due to formant structure

4] 9.19 Rabiner

6

$P(z)$ and $Q(z)$ are derived from p^{th} optimum prediction error filter $A(z)$:-

$$P(z) = A(z) + z^{-(p+1)} A(z^{-1})$$

$$Q(z) = A(z) - z^{-(p+1)} A(z^{-1})$$

$$A(z) = 1 - \sum_{k=1}^{p} \alpha_k z^{-k}$$

a) p is even, $P(z)$ is 0 at $z = -1$

$$P(-1) = A(-1) + (-1)^{(p+1)} A(-1)$$

$$A(-1) = 1 - \sum_{k=1}^{p} \alpha_k (-1)^{-k}$$

$$P(-1) = \cancel{1 - \sum_{k=1}^{p} \alpha_k (-1)^{-k}} - \left[\cancel{1 - \sum \alpha_k (-1)^{-k}}\right]$$

$(-1)^{p+1} = -1$ as p is even

$$P(-1) = 0$$

$$Q(1) = A(1) \mp (+1)^{-(p+1)} A(+1)$$

$$= A(1) - A(+1) = 0$$

$Q(1) = 0$

$(+1)^{(p+1)} = 1$ for anything

b) To show, if $A(z)$ is system function of an optimum prediction filter derived by autocorrelation func, then zeros of $P(z)$ $Q(z)$ are on unit circle

zeros of $P(z)$

$$A(z) + z^{-(p+1)} A(z^{-1}) = 0$$

$$A(z) = -z^{-(p+1)} A(z^{-1})$$

$$\frac{-z^{-(p+1)} A(z^{-1})}{A(z)} = \pm 1$$

zeros of $Q(z)$

$$A(z) - z^{-(p+1)} A(z^{-1}) = 0$$

$$\frac{z^{-(p+1)} A(z^{-1})}{A(z)} = \pm 1$$

=> we can see all zeros are within unit circle

On the unit circle we have $|z| = 1$

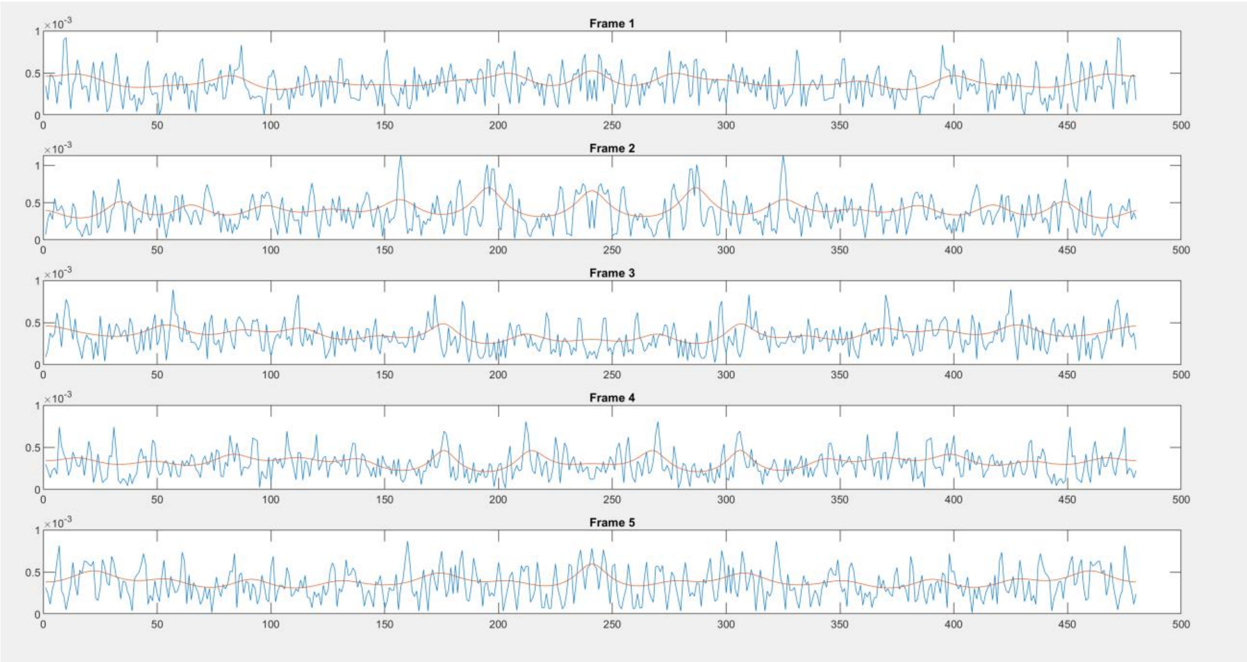
=> $\frac{A(z^{-1})}{A(z)} = \pm 1$ in both cases when $P(z)$ and $Q(z)$ are ~~Else the zeros will be unit circle~~ multiples of π. For any other values of $|z|$, ~~the unit circle~~ value will not be 1, => root of equations have to lie on unit circle

Question 5)

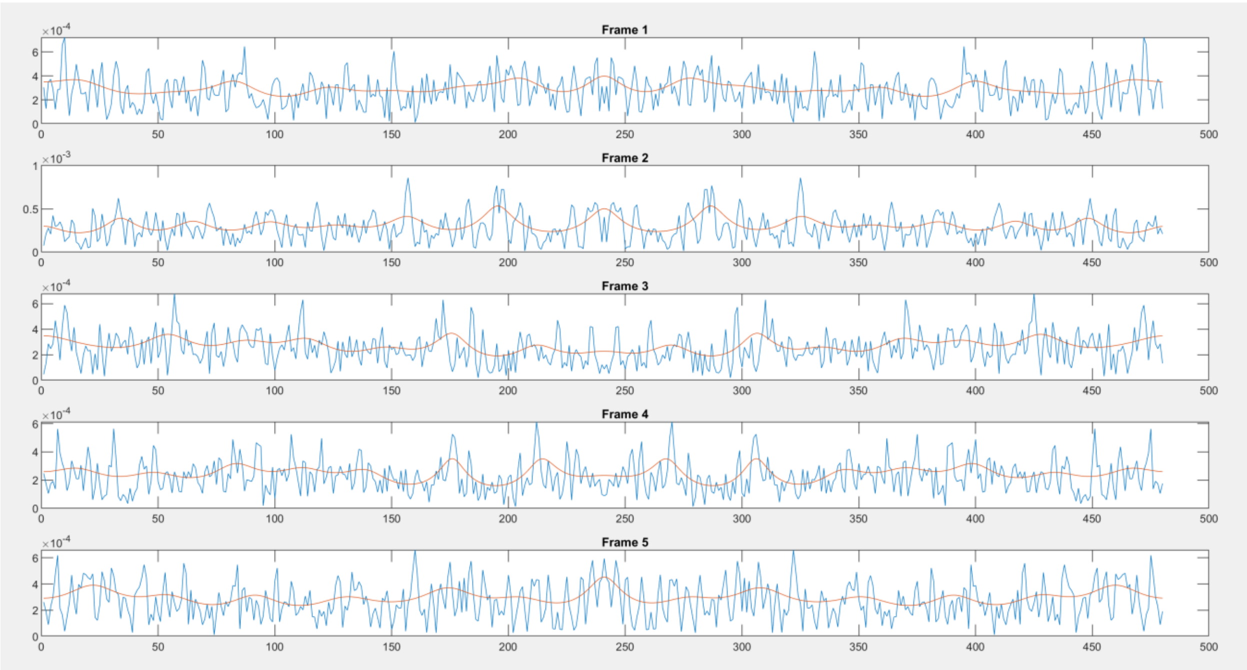
Jeet Kiran Pawani || Assignment -6 || ECE – 6255 || GTid - 903397407

After generating the autocorrelation coefficients, the output of short time speech and autocorrelation LPC coefficients is –

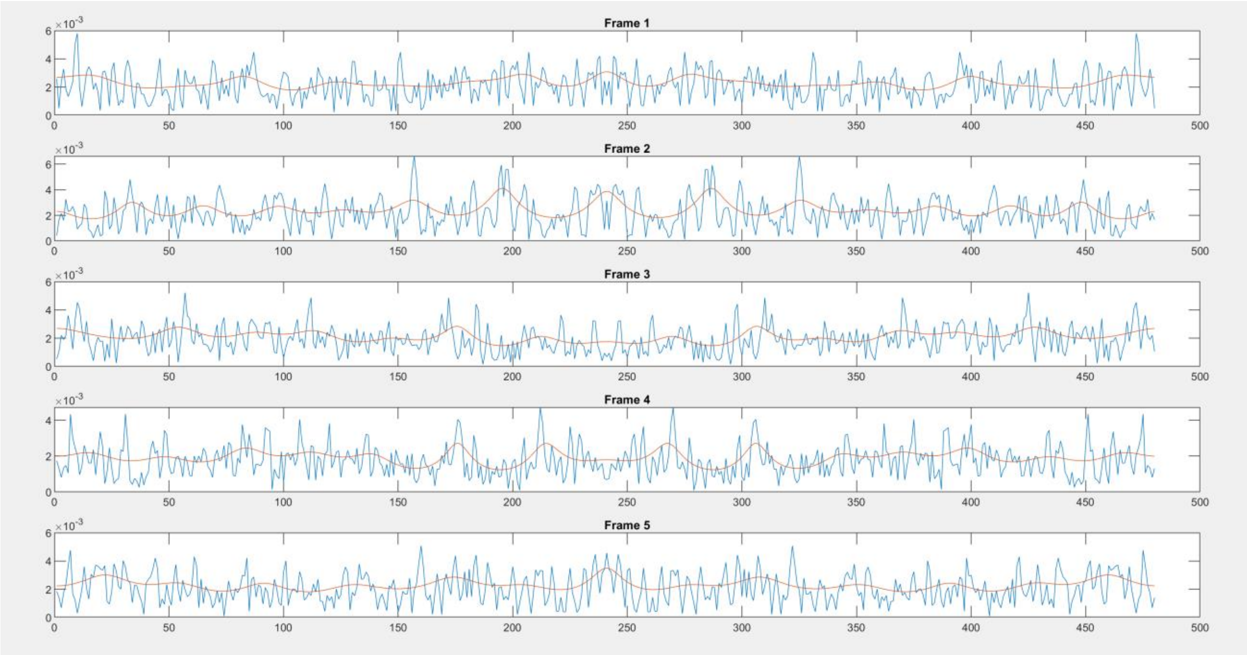
IY Vowel



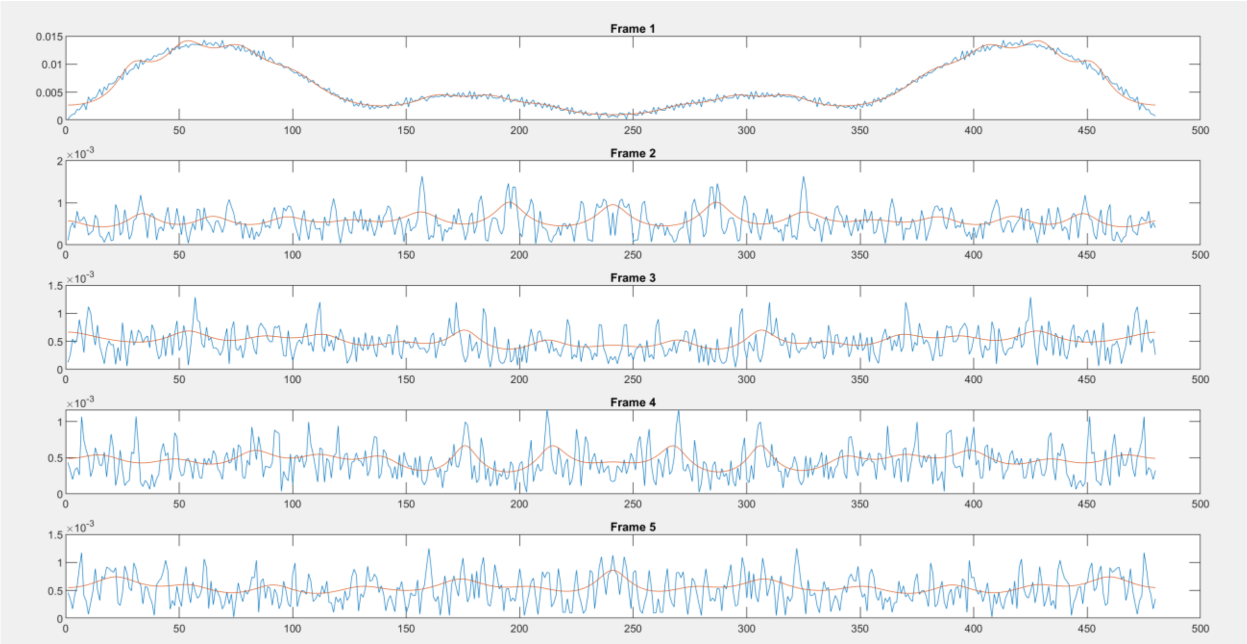
AA Vowel



UW Vovel



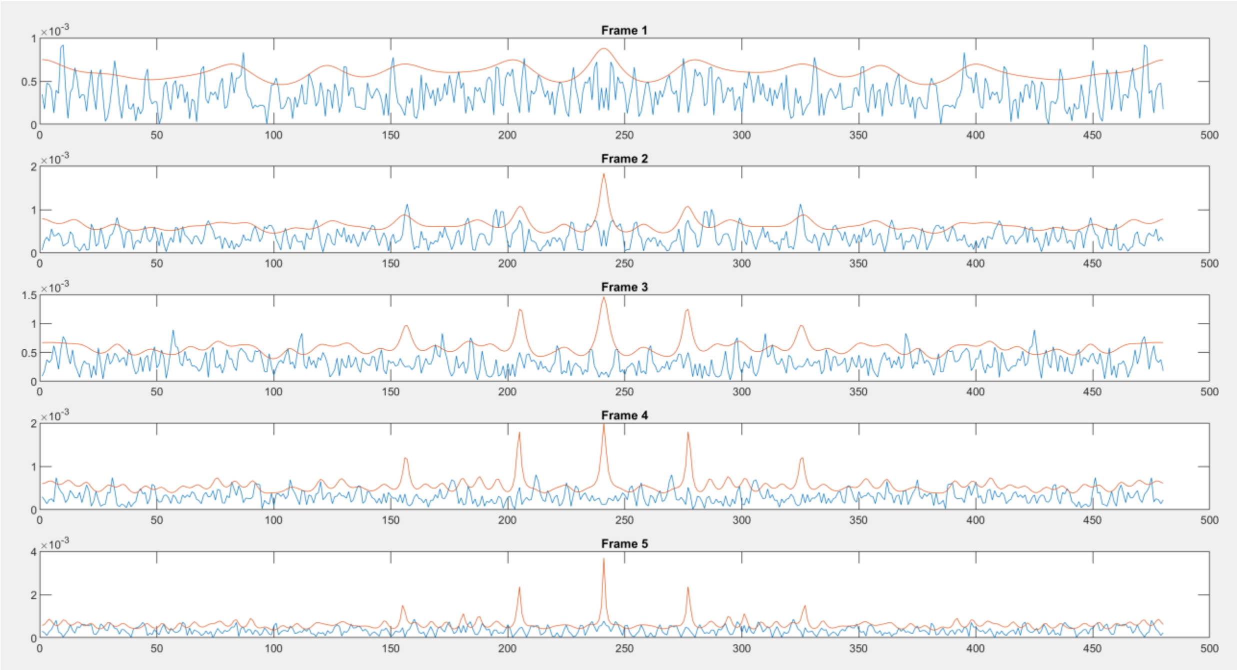
ZH Fricative



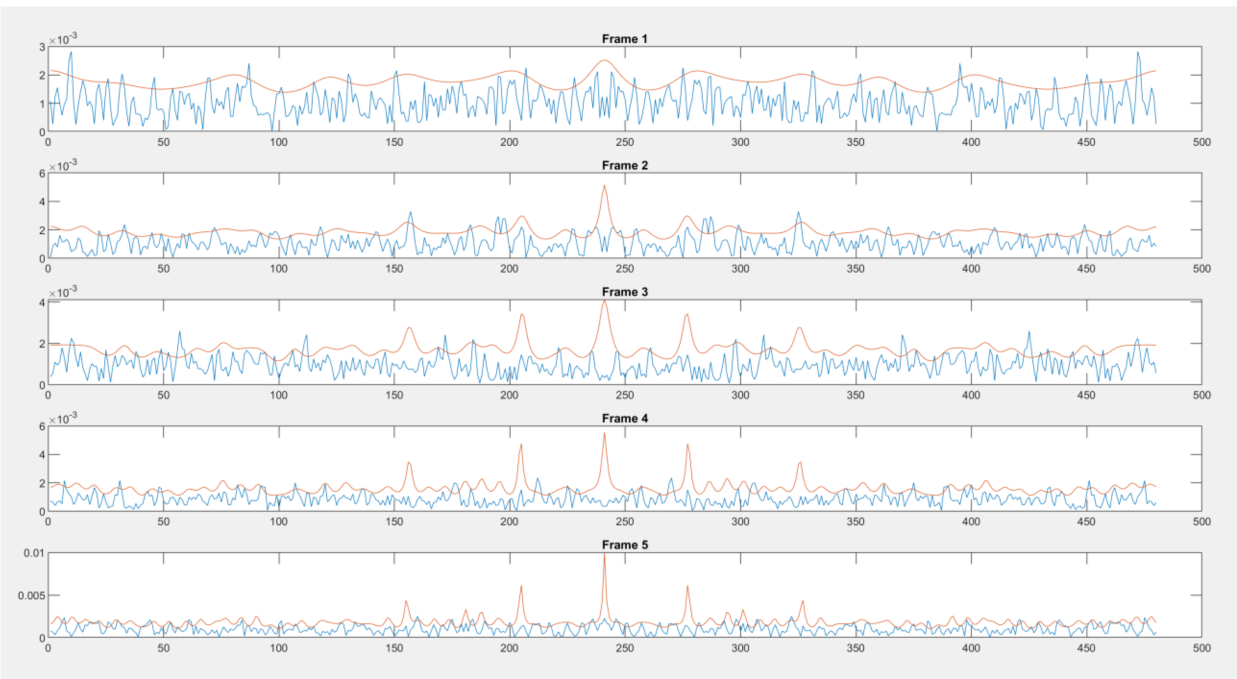
After generating the Covariance coefficients, the output of short time speech and covariance LPC coefficients is –

IY Vowel

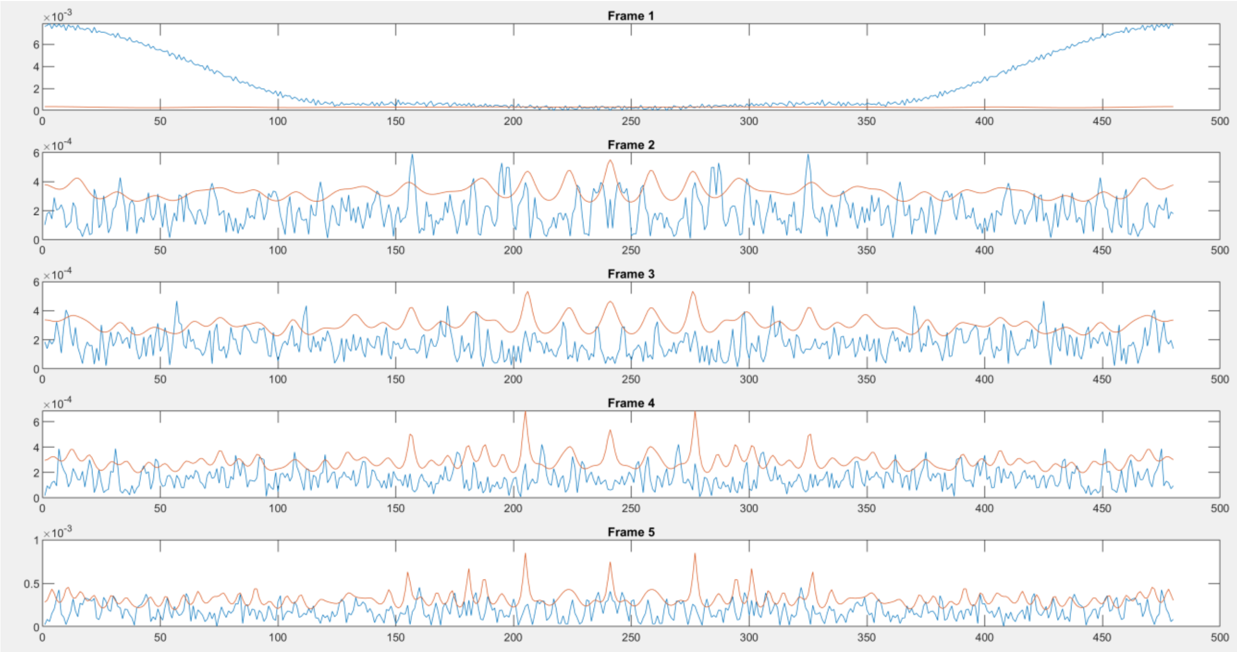
Jeet Kiran Pawani || Assignment -6 || ECE – 6255 || GTid - 903397407



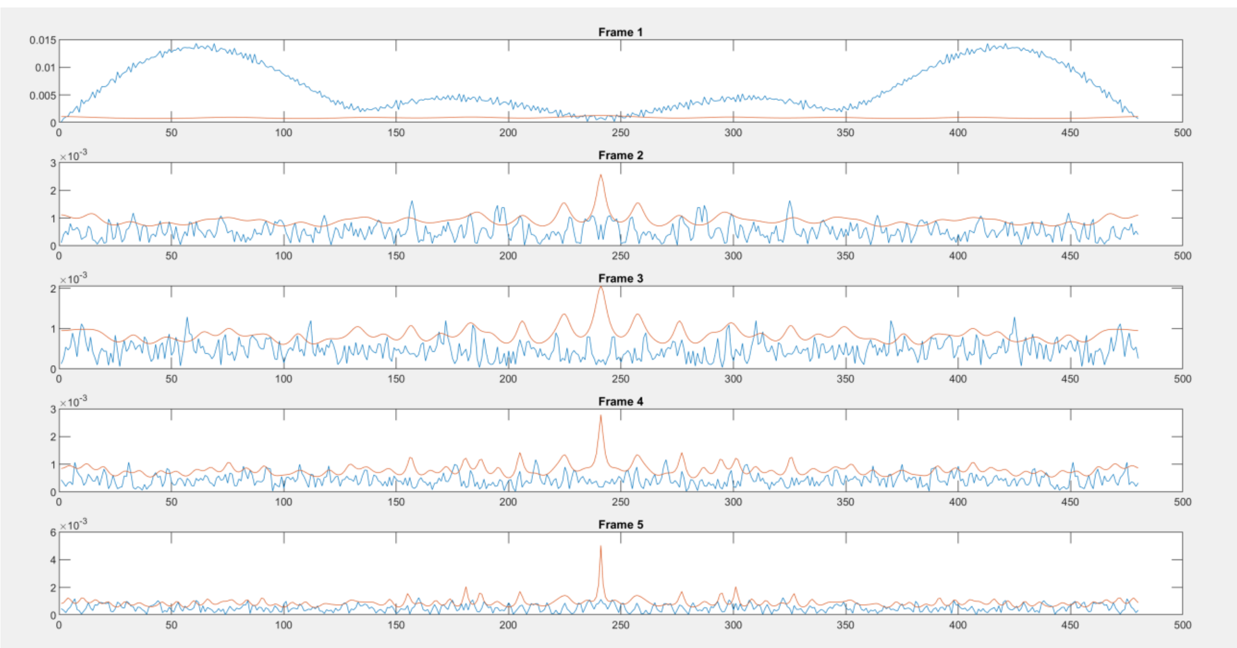
UW Vowel



AA Vowel

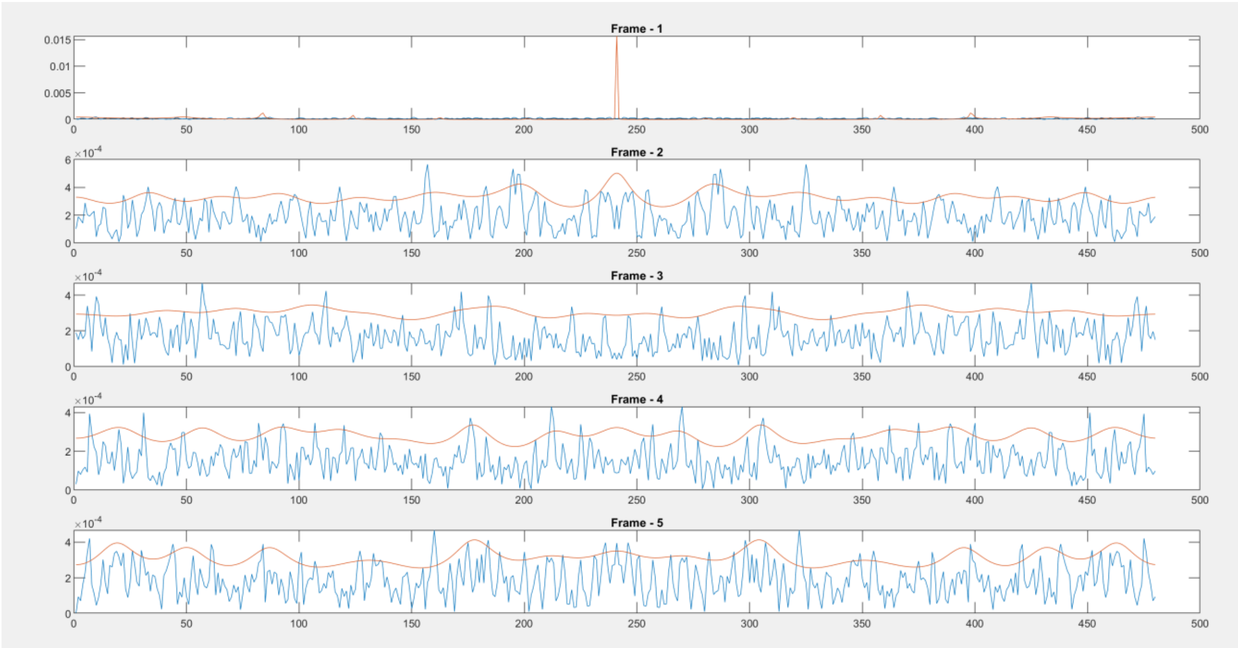


ZH Fricative

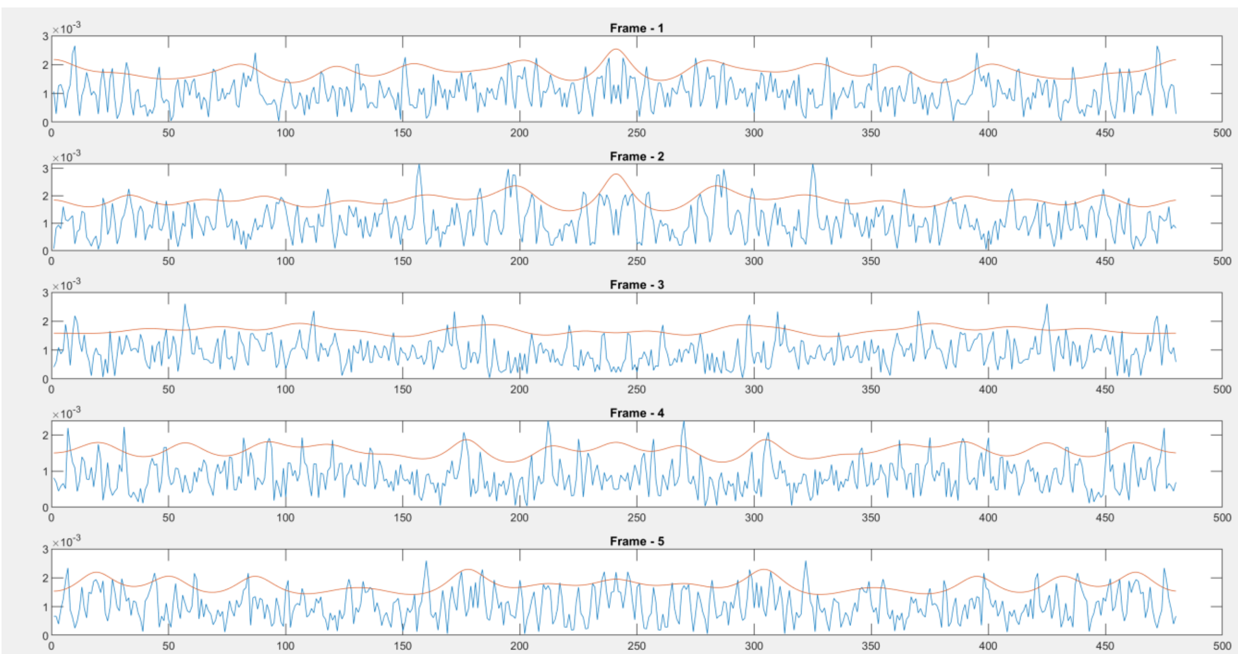


After generating the Lattice coefficients, the output of short time speech and lattice LPC coefficients is –

AA Vowel

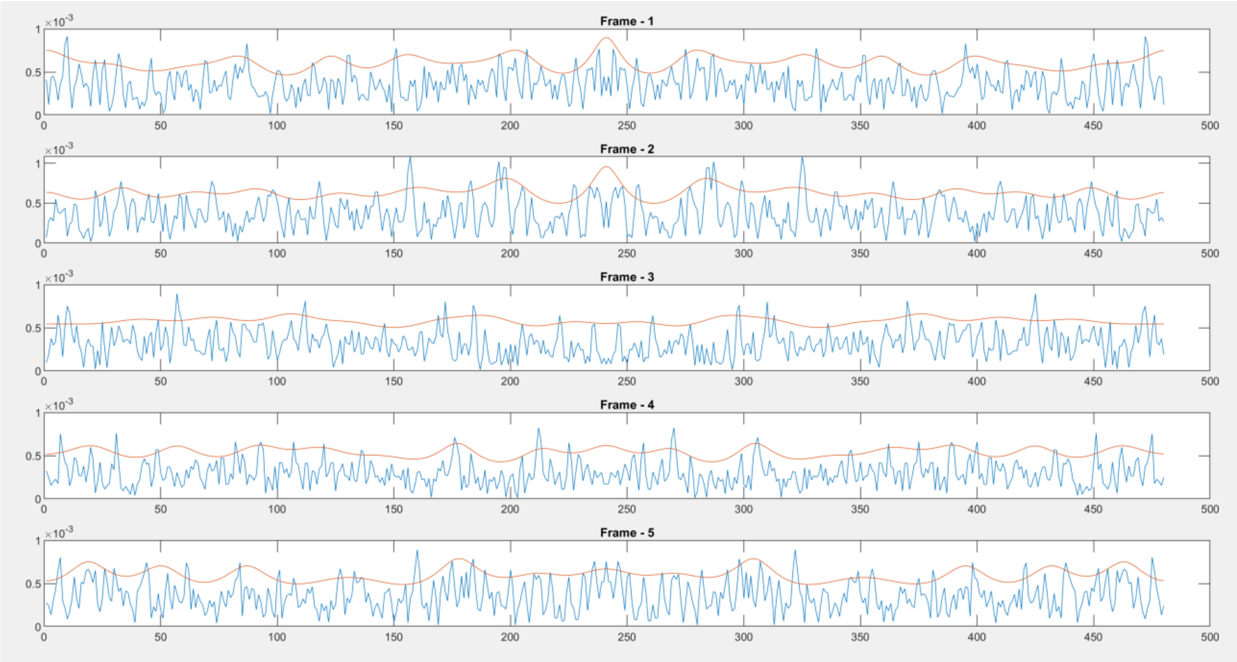


UW Vowel

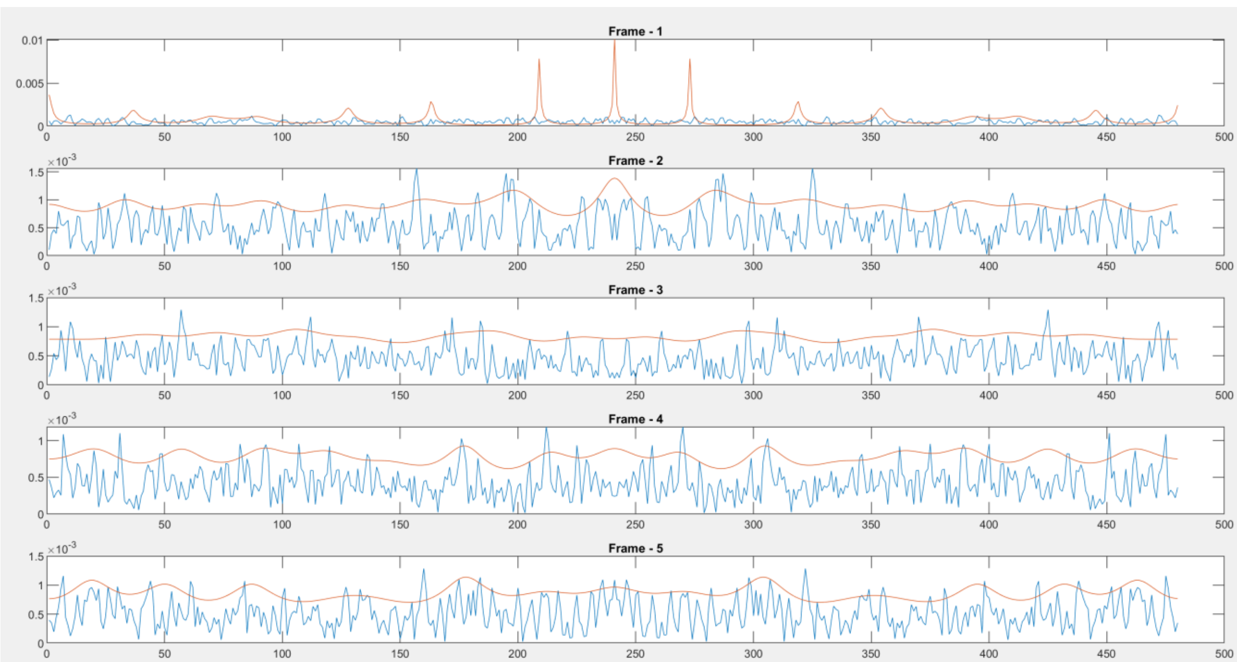


IY Vowel

Jeet Kiran Pawani || Assignment -6 || ECE – 6255 || GTid - 903397407

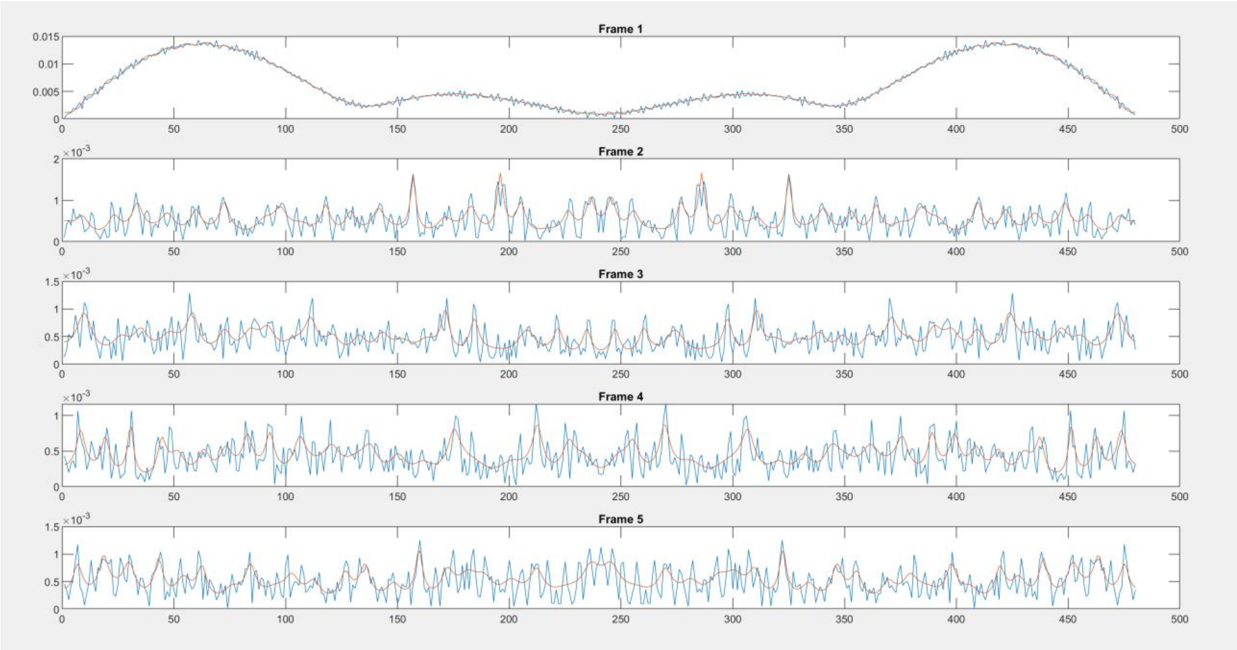


ZH Fricative



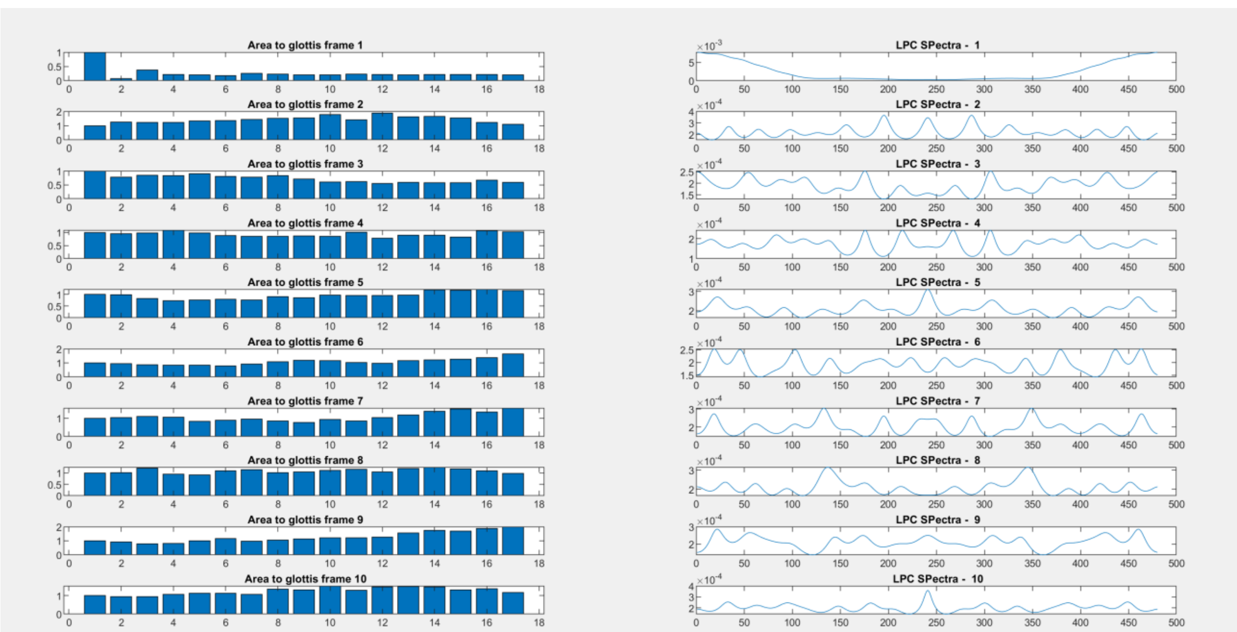
P=50

ZH fricative

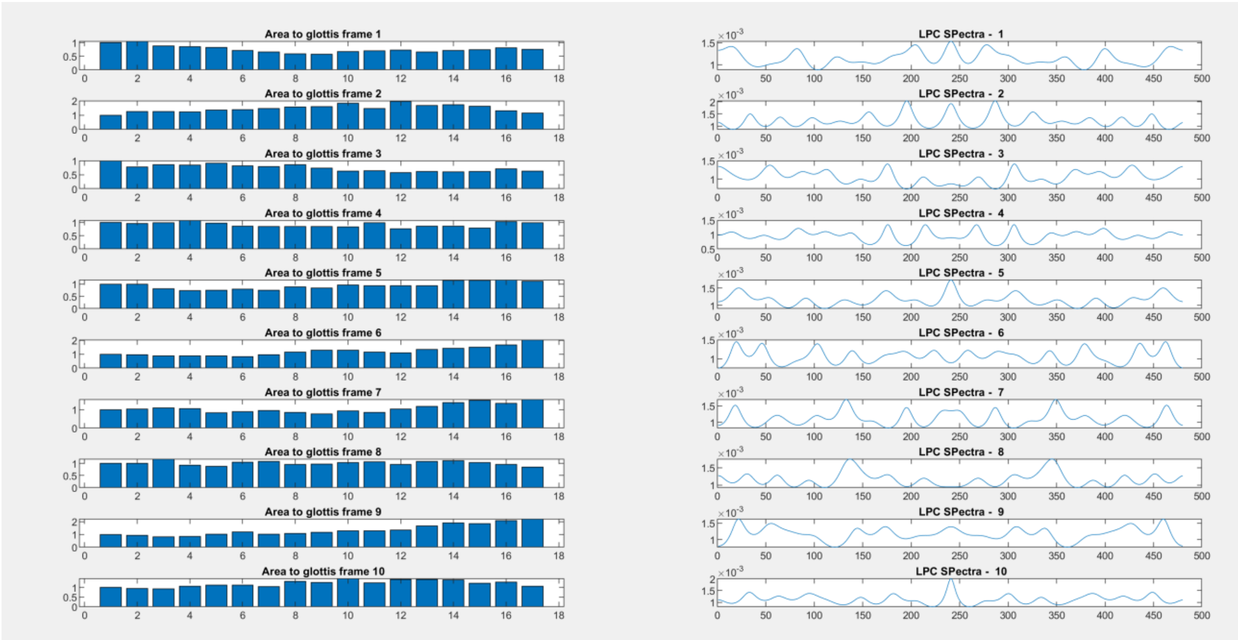


Question 6)

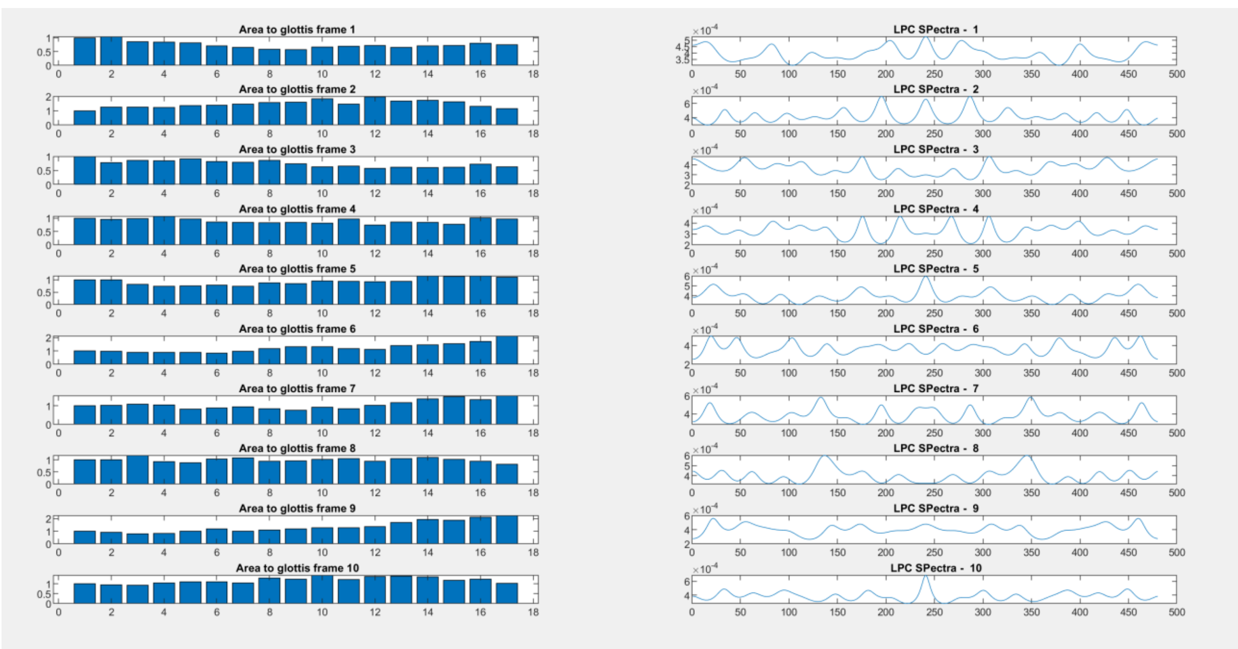
AA_Vowel



UW Vowel

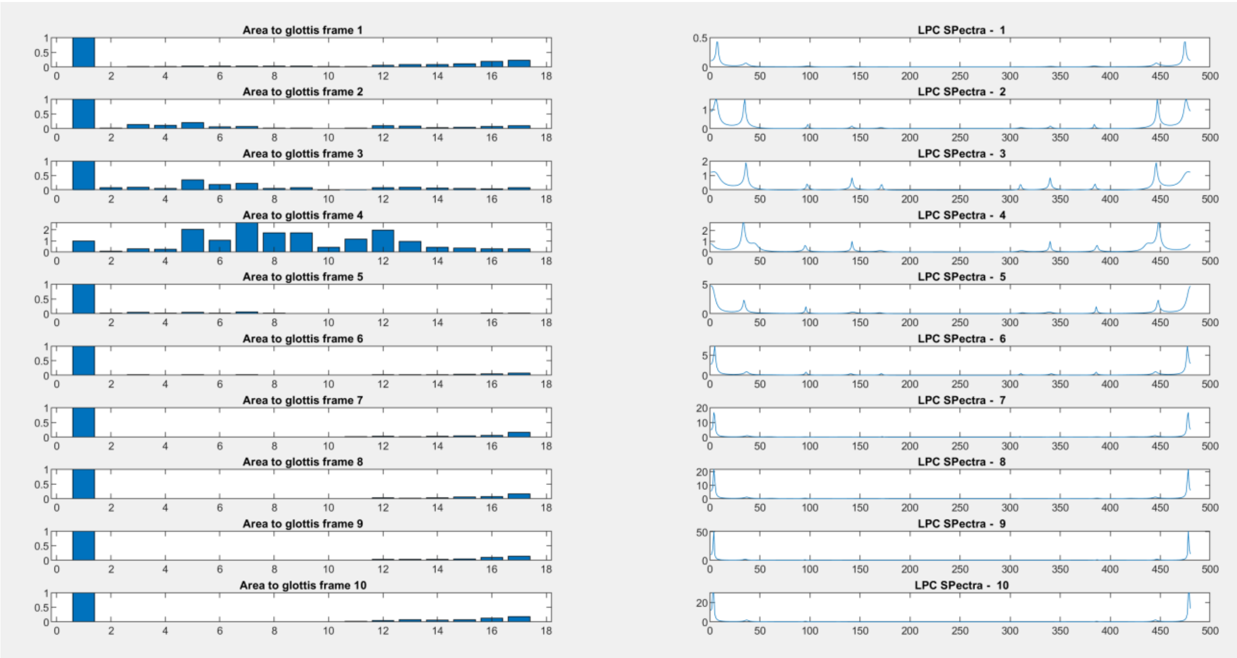


IY Vowel

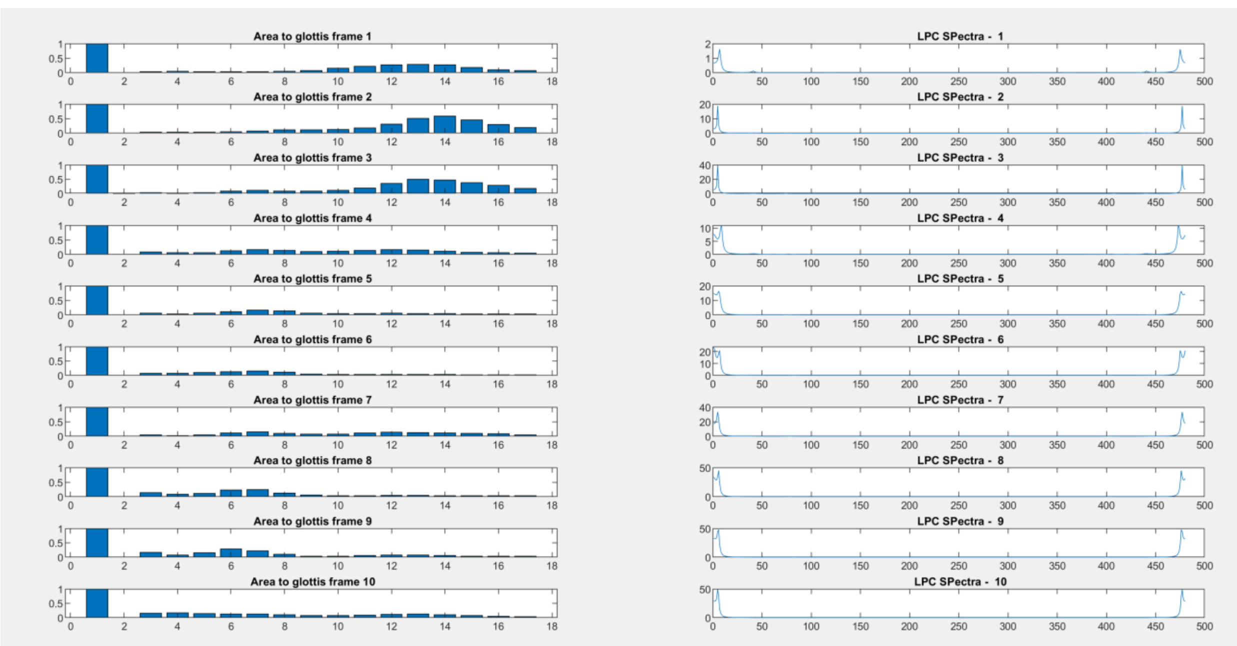


AE Vowel

Jeet Kiran Pawani || Assignment -6 || ECE – 6255 || GTid - 903397407



UH Vowel



Question 7)

Distance from NxM to 0

	Male1	Male2	Male3	Female1	Female2	Female3
Male1	0/50.6539	50.6539	51.0049	50.9145	50.9145	51.2371
Male2	29.011	0/29.011	30.6955	30.0958	30.0958	30.9466
Male3	49.3071	49.3071	0/51.5813	50.6565	50.6565	52.7661
Female1	38.5237	38.5237	38.7980	0/38.7201	38.7201	39.1508
Female2	34.39	34.39	34.8148	34.7758	0/34.7758	35.4233
Female3	49.0059	49.0059	50.3994	50.2096	50.2096	0/50.7883

Distance from 0 to NxM

	Male1	Male2	Male3	Female1	Female2	Female3
Male1	0/44.0372	44.0372	44.8656	44.1419	44.1419	45.553
Male2	32.9234	0/32.9234	34.2926	33.8987	33.8987	34.4317
Male3	50.3750	50.3750	0/52.4856	51.9429	51.9429	53.7414
Female1	35.0876	35.0876	35.9823	0/35.7057	35.7057	36.2652
Female2	34.0971	34.0971	34.7413	34.4350	0/34.4350	35.3543
Female3	35.7719	35.7719	36.7037	36.3087	36.3087	0/38.1725

In this question the DTW has been implemented and the above are the results I obtained for comparisons of all Speeches given.

I got different answers for both Dynamic programming means starting at 0,0 and starting from N,M where N and M are number of frames in each sequences.

I have attached both the codes in with this assignment.

B) It is very clearly seen that on comparison with male to male speeches the distances are lesser. WE can say that the coefficient values are similar for the spoken words and it’s the similar with female to female comparison as well confirming lesser distance.

But when we compare between male and female speaker the distances seem to increase. There is a wide difference between both confirming that both have there were spoken by different speakers.

There is a place where we can see that distance from female speaker to male speaker is less.

Checking 0 to NxM distance matrix -
The average speaker distance male 1 to other male speakers is 44.5
Male 1 to female speakers is – 44.9
The average speaker distance male 2 to other male speakers is 33.5
Male 2 to female speakers is – 33.7
The average speaker distance male 3 to other male speakers is 51.2
Male 3 to female speakers is 52.7
The average speaker distance female 1 to other male speakers 35.5
female 1 to female speakers is – 36
The average speaker distance female 2 to other male speakers is 34.3
Female 2 to female speakers is – 34.8
The average speaker distance female 3 to other male speakers is 36
female 3 to female speakers is – 37.5