ECE - 6255

JEET KIRAN PAWANI
GTid- 903397407

DIGITAL PROCESSING FOR SPEECH

$\boxed{1}$ 7.7

$$W(z) = \frac{\sum_{r=0}^{N_z} b_r z^{-r}}{1 - \sum_{k=1}^{N_p} a_k z^{-k}}$$

a) For $W(z)$ to be suitable for this application, $W(z)$ should be stable. All the poles must be inside the unit circle. It should be of the form of Low pass filter with frequency in frequency domain at $\omega = 0$.

b) Given $$W(z) = \frac{\sum_{r=0}^{N_z} b_r z^{-r}}{1 - \sum_{k=1}^{N_p} a_k z^{-k}}$$

$$X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} x[m]\, w[n-m]\, e^{-j\hat{\omega} m}$$

which can be written as

$$= x[m]\, e^{-j\hat{\omega} m} * w[n]$$

Now $$X_n(e^{j\hat{\omega}}) = x(n)\, e^{-j\hat{\omega} n} * w(n)$$

changing to z domain

$$X_n(e^{j\hat{\omega}}) = x(z)\, e^{-j\hat{\omega} n}\, W(z)$$

$$= X(z)\, e^{-j\omega n} \frac{\sum_{r=0}^{N_F} b_r z^{-r}}{1 - \sum_{r=1}^{N_p} a_k z^{-k}}$$

$$X_n(e^{j\hat{\omega}})\left[1-\sum_{k=1}^{N_p} a_k z^{-k}\right] = x(z]\, e^{-j\hat{\omega}n} \sum_{r=0}^{N_z} b_r z^{-n}$$

$$X_n(z)\left[1-\sum_{k=1}^{N_p} a_k z^{-k}\right] = X(z) e^{-j\omega n} \sum_{r=0}^{N_z} b_r z^{-n}$$

$$X_n(z) - \sum_{k=1}^{N_p} a_k X_n(z) z^{-k} = \sum_{r=0}^{N_z} X(z)\, e^{-j\omega n}\, b_r\, z^{-r}$$

reverting back to frequency

$$X_n(e^{j\hat{\omega}}) = \sum_{k=1}^{N_p} X_{n-k}(e^{j\omega})\, a_k + \sum_{r=0}^{N_z} x(n-r]\, b_r\, e^{-j\omega(n-r)}$$

$$\Rightarrow X_n(e^{j\hat{\omega}}) = \sum_{k=1}^{N_p} X_{n-k}(e^{j\omega})\, a_k + \sum_{r=0}^{N_z} x[n-r]\, e^{-j\hat{\omega}(n-r)}\, b_r$$

c) $W(z) = \frac{1}{1-az^{-1}}$ $|a|<1$ for stability and a should be real. (2)

$$W(e^{+j\omega}) = \frac{1}{1-ae^{-j\omega}} = \frac{1}{1-a(\cos(\omega) - j\sin(\omega))}$$

$$= \frac{1}{1-a\cos(\omega) + aj\sin(\omega)}$$

Lets take magnitude squared frequency response -

$$|W(e^{j\omega})|^2 = \frac{1}{(1-a\cos(\omega))^2 + a^2\sin^2(\omega)} = \frac{1}{1-2a\cos(\omega) + a^2\cos^2(\omega) + a^2\sin^2(\omega)}$$

$$= \frac{1}{1-2a\cos(\omega)+a^2}$$

we define the cutoff frequency of system for which:

$$|W(e^{j\omega_c})|^2 = \frac{1}{2}|W(e^{j0})|^2$$

Frequency resolution is 100 Hz on samplingrate of 10KHz.

Since cut off frequency is half of 100 Hz $= \frac{1}{2}100 = 50$ Hz

$$\omega_c = 2\pi f_c/f_s = 2\pi \frac{50}{10000} = \pi/100$$

$$\Rightarrow |W(e^{j\omega_c})|^2 = \frac{1}{1+a^2-2a\cos(\pi/100)}$$

$$\frac{1}{2}|W(e^{j0})|^2 = \frac{1}{2}\frac{1}{1+a^2-2a\cos(0)} = \frac{1}{2(1+a^2-2a)}$$

we know $|W(e^{j\omega_c})|^2 = \frac{1}{2}|W(e^{j0})|^2$

$$\frac{1}{1+a^2-2a\cos(\pi/100)} = \frac{1}{2(1+a^2-2a)}$$

$$2+2a^2-4a = 1+a^2-2a\cos(\pi/100)$$

$$1+a^2-4a+2a\cos(\pi/100) = 0$$

$$1+a^2-2a\left(2-\cos(\pi/100)\right) = 0$$

$$1+a^2-2a(2-0.9995) = 0$$

$$1+a^2-2a(1.00049) = 0$$

$$1+a^2-2.00098\,a = 0$$

$a = 1.031798$ and $a = \underline{0.96918}$

since $|a|<1$ we choose $\boxed{a = 0.96918}$

2) $$X_n(e^{jw}) = \sum_{m=-\infty}^{\infty} x(m)\,\underline{w(n-m]}\,e^{-jmw}$$

$$w(n) = n\beta^n\,\underline{u[n]}$$

a) $X_{50}(e^{jw})$ basically will use $x(m)$ with $m$ less than or equal to 49 else there will be aliasing

$X_{100}(e^{jw})$ will use $x(m)$ with $m \le 100$ else there will be aliasing

b) $X_n(e^{j\omega}) = \sum_{m=-\infty}^{\infty} w(n-m)\, x(m)\, e^{-j\omega m}$ 

Let $y(n) = X_n(e^{j\omega})$

$v(n) = x[n]\, e^{-j\hat{\omega} n}$

$y(n) = v(n) * w(n)$

$$v(n) = x(n)e^{-j\hat{\omega} m} \longrightarrow \boxed{W(z)} \xrightarrow{\;y(n) = X_n(e^{j\omega})\;}$$

$w(n) = n\beta^n u(n)$

$W(z) = -z \dfrac{d}{dz}\left(\dfrac{1}{1-\beta z^{-1}}\right)$

$= -z\left[\dfrac{-1\;(-\beta z^{-2})}{(1-\beta z^{-1})^2}\right]$

$$W(z) = \frac{\beta z^{-1}}{1 - 2\beta z^{-1} + \beta^2 z^{-2}}$$

We know $Y(z) = V(z)\, W(z)$

$\rightarrow W(z) = \dfrac{Y(z)}{V(z)}$

$$\frac{\beta z^{-1}}{1 - 2\beta z^{-1} + \beta^2 z^{-2}} = \frac{Y(z)}{V(z)}$$

$\beta z^{-1}(V(z)) = (1 - 2\beta z^{-1} + \beta^2 z^{-2})(Y(z))$

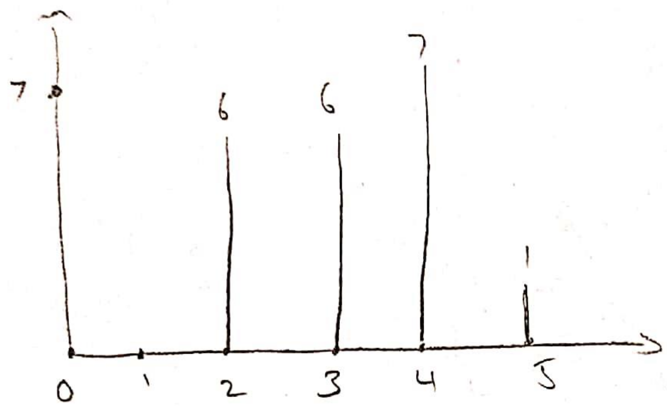
$\cancel{\beta z^{-1}(v(n))}$

$\beta\, v(n-1) = y(n) - 2\beta y(n-1) + \beta^2 y(n-2)$

$y(n) = X(e^{j\omega}) \qquad v(n) = x[n]\, e^{-j\omega m}$

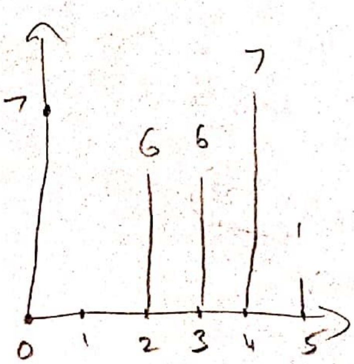
$$X_n(e^{j\omega}) = 2\beta X_{n-1}(e^{j\omega}) - \beta^2 X_{n-2}(e^{j\omega}) + \beta\, x(n-1)\, e^{j\omega(n-1)}$$

$\boxed{3}$ $x[n] = 7\delta(n) + 6\delta(n-2) + 6\delta(n-3) + 7\delta(n-4) + \delta(n-5)$

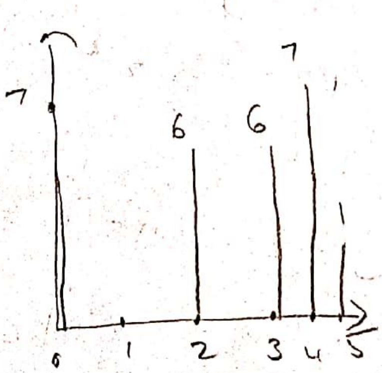



we can see that the sample is 6 points from 0 to 5. So, when we sample the dft at N points we alias at N-th point. If $N > 5$ in our case there will be no aliasing and response will be the same as original sequence.

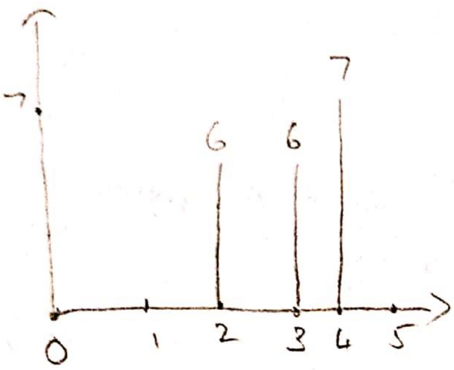
So, for 40 pt and 10 pt DFT there is no aliasing. For $N=5$ aliasing happens at point 5. In case of $N=3$, points 3-5 get removed.



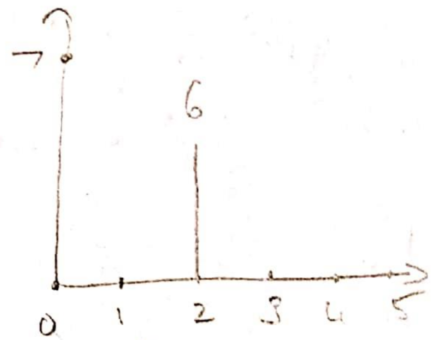

N = 40




N = 10



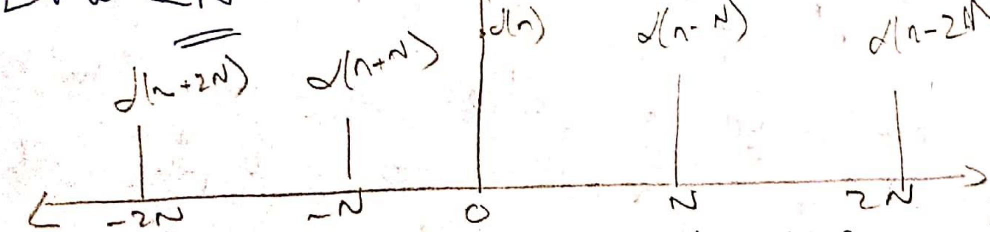
N = 5.
$5^{th}$ point
~~cut off~~



N = 3
points 3, 4, 5 gets aliased

**4** a) $y(n) = x(n) \otimes w(n)\left[\frac{1}{N}\sum_{r=-\infty}^{\infty} N\,\delta(n-rN)\right]$

$L = N_w < N$



Impulse train.

window is multiplied with impulse train.

It is given that $N_w$ (window) length is less than the IDFT samples.

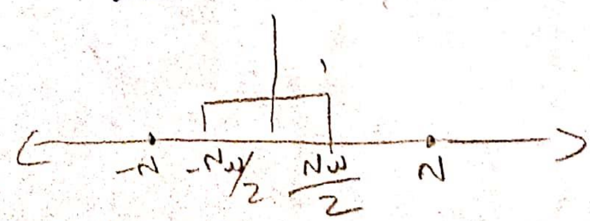
$w(n)\,\frac{1}{N}\sum_{r=-\infty}^{\infty} N\,\delta(n-rN)$ can be written as

$$\sum_{r=-\infty}^{\infty} w[rN]\,\delta(n-rN)$$

Now we know that $N_w < L$ so, the only values which are included are that at $r=0$

$$\Rightarrow \sum_{r=0} w[0]\,\delta(n)$$

$$= \boxed{w[0]\,\delta(n)}$$



Now,

$$y(n) = x[n] \circledast w(n) \frac{1}{N} \sum N \delta(n - rN)$$

$$y(n) = x(n) \circledast w[0] \delta(n)$$

This is 1 only at one position
Convolution changes to multiplication.

$$\Rightarrow y(n) = x[n] \cancel{\circledast} w(0)$$

$$y(n) = x(n) w[0)$$

$$\boxed{x(n) = \frac{y(n)}{w(0)}}$$

b) Now we know that $w(n) = 0$ at $n = \pm N, \pm 2N \ldots$
This implies that $w(n)$ is non zero only at $w[0)$
=> Any window size no matter of what length will have only one non-zero sample that is at $n=0$, i.e., $w(0)$.

=> Since only one term occurs at $n=0$, so any window length will not cause any aliasing and we can construct the perfect signal back again.

c) $w(n)\left[\sum_{r=-\infty}^{\infty} d(n-rN)\right] = w(0)d(n)$ (5)

From a part.

Taking DTFT on both sides

$$LHS = \sum_{n=-\infty}^{\infty} w(n)\left[\sum_{-\infty}^{\infty}(n-rN)\right] e^{-j\omega n}$$

$$\sum_{n=-\infty}^{\infty} w(n)e^{-j\omega n} = W(\omega) \text{ given}$$

$$= W(\omega) \sum_{r=-\infty}^{\infty} d(n-rN)$$

we can write $\sum_{r=-\infty}^{\infty} d(n-rN)$ as $\frac{1}{N}\sum_{k=0}^{N-1} e^{+j\frac{2\pi kn}{N}}$

$$= \sum_{n=-\infty}^{\infty} w(n)e^{-j\omega n} \frac{1}{N}\sum_{k=0}^{N-1} e^{+j\frac{2\pi kn}{N}}$$

$$= \sum_{k=0}^{N-1}\sum_{n=-\infty}^{\infty} w(n)e^{-j\left(\omega - \frac{2\pi k}{N}\right)n}\frac{1}{N}$$

we know $\sum_{n=-\infty}^{\infty} w(n)e^{-j\omega n} = W(\omega)$

$$\Rightarrow \sum_{n=-\infty}^{\infty} w(n)e^{-j\left(\omega - \frac{2\pi k}{N}\right)n} = W\left(\omega - \frac{2\pi k}{N}\right)$$

$$\Rightarrow LHS = \frac{1}{N}\sum_{k=0}^{N-1} W\left(\omega - \frac{2\pi k}{N}\right)$$

$$RHS = \sum_{n=-\infty}^{n=\infty} w(0)\, d(n)\, e^{-j\omega n}$$

we know $\sum_{n=-\infty}^{\infty} d(n)e^{-j\omega n} = 1$

$$RHS = w(0).1 = w(0)$$

Now LHS = RHS

$$\Rightarrow \frac{1}{N}\sum_{k=0}^{N-1} W\left(\omega - \frac{2\pi k}{N}\right) = \underline{W(0)} \quad \text{for any } \omega$$

$\boxed{5}$ a) $\hat{w}^{m}(r) = \underbrace{w(rR-n)}_{\text{time shifted version of } w(n)}$ downsampled version of $w(n)$

we know $\hat{W}(e^{j\omega}) = \frac{1}{R}\sum_{r=0}^{R-1} W\left(e^{j(\omega-2r\pi)/R}\right)$

This is basically sampling around the unit circle in frequency domain.

if we keep $\omega = 0$

$$\Rightarrow \hat{W}(e^{j0}) = \frac{1}{R}\sum_{r=0}^{R-1} W\left(e^{j2\pi r/R}\right)$$

There is a case of possible time domain aliasing which can happen. This needs to be avoided.

From the textbook by Rabiner,

time aliasing occurs when window length is way greater than sample FFT length.

This is the case where aliasing occurs in time domain where certain samples are lost. So we know to avoid this window length must be length samples

$\Rightarrow N_w$ or $L < N$ $\Rightarrow$ $\boxed{R \leq L \leq N}$ from text book

$R \leq L$ should be the case for undersampling the first time but as it is nothing but frequency domain sampling.

$\omega_c = \frac{2\pi}{L}$ is the condition, if we consider the case of down sampling.

$\Rightarrow$ For $|\omega| = \frac{2\pi r}{R}$ we know that

From the constraint condition

$|\omega| \geq \omega_c$

$\Rightarrow \frac{2\pi r}{R} \geq \frac{2\pi}{L}$

$\boxed{R \leq L}$

(b) $Y_r(k) = Y_r(e^{j\omega_k}) = X_{rR}(e^{j2\pi kn/N})$ (6)

$y_r(m) = x(m)\ w(rR - m)$

we have $y(n) = \sum_{r=-\infty}^{\infty} y_r(n) = \sum_{r=-\infty}^{\infty} x(n)\, w(rR-n)$

$x$ is independent of $R$

$$\sum_{r=-\infty}^{\infty} y_r(n) = x(n) \sum_{r=-\infty}^{\infty} w(rR-n)$$

given $\sum_{r=-\infty}^{\infty} w(rR-n) = \frac{w(0)}{R}$

$$\Rightarrow \sum_{r=-\infty}^{\infty} y_r(n) = x[n]\ \frac{w(0)}{R} \rightarrow (1)$$

Given $y_r(n) = \frac{1}{N} \sum_{k=0}^{N-1} Y_r(k)\, e^{j\frac{2\pi kn}{N}}$

from above

$$= \frac{1}{N} \sum_{k=0}^{N-1} X_{rR}(e^{j2\pi kn/N})\, e^{j\frac{2\pi kn}{N}}$$

$$\approx \boxed{y_r(n) = \frac{1}{N} \sum_{k=0}^{N-1} X_{rR}(k)\, e^{j\frac{2\pi kn}{N}}} \rightarrow (2)$$

From (1)

$$\sum_{r=-\infty}^{\infty} y_r(n) = x(n)\ \frac{w(0)}{R}$$

$$x(n) = \frac{R}{w(0)} \sum_{r=-\infty}^{\infty} y_r(n)$$

Sub (2) here

$$\boxed{x(n) = \frac{R}{w(0)} \sum_{r=-\infty}^{\infty} \left[ \frac{1}{N} \sum_{k=0}^{N-1} X_{rR}(k)\, e^{j\frac{2\pi k}{N}} \right]}$$

c) $W(e^{j\omega}) = 0$ at $\omega = 2\pi k/R$, k is given

we know $\hat{W}(e^{j\omega}) = \frac{1}{R} \sum_{r=0}^{R-1} W\left(e^{j(\omega - 2r\pi)/R}\right)$

→ from 5a)

Let $\omega = 0$

$$\hat{W}(e^{j0}) = \frac{1}{R} \sum_{r=0}^{R-1} W\left(e^{j(-2\pi r)/R}\right)$$

here given
$W(e^{j\omega}) = 0$ at $\omega = 2\pi k/R$

in this case lets take r

Let us assume $R \leq L \leq N$ from 5a)

$\Rightarrow \hat{W}(e^{j0}) = \frac{W(0)}{R}$ this has no constraints on R

$\Rightarrow \boxed{r > 0}$ or $\boxed{k > 0}$

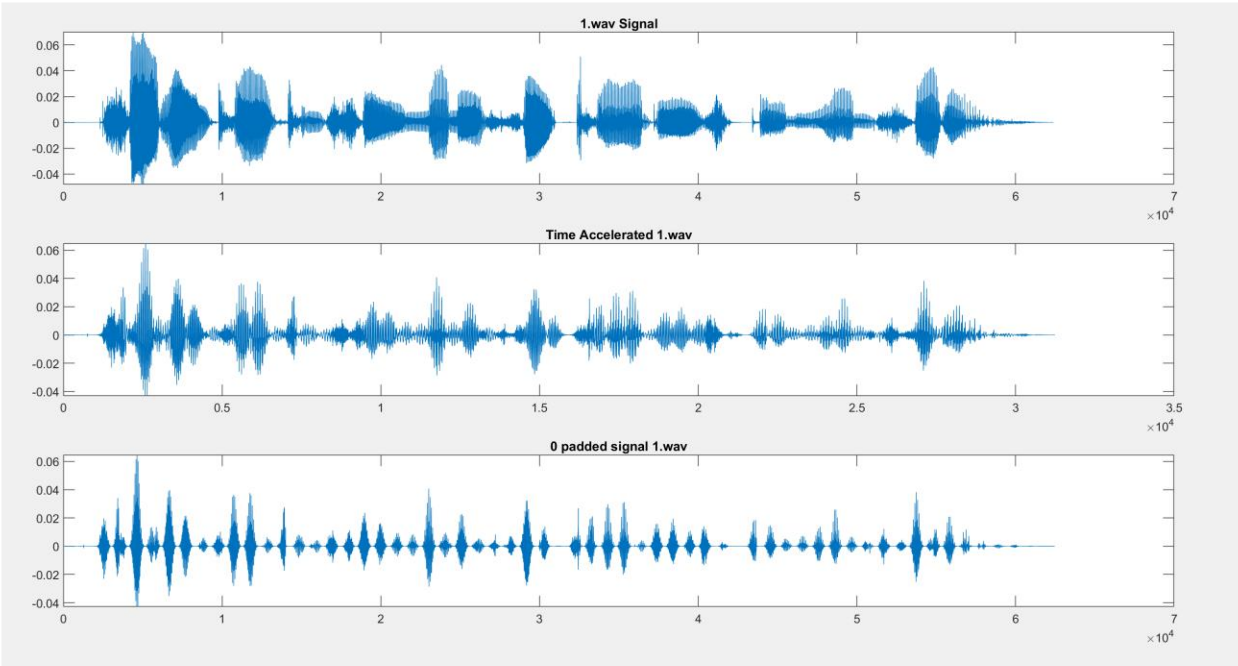
if $r \geq 0$ we get $\hat{W}(e^{j0}) = 0$ as $0^{th}$ sample position sample is also 0.

$\Rightarrow r \geq 0$ is feasible to get no boundary constraints for $W(e^{j\omega}) = 0$ $\omega = \frac{2\pi r}{R}$

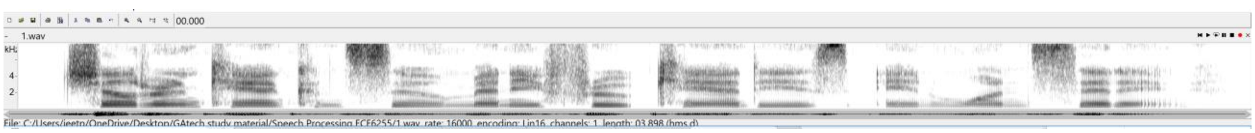
6 I have used a 1024 pt FFT and window length of 1024. On downsampling and performing ifft, we can hear a lot gaps in time due to which we hear staggered sound. Sound is heard and is understandable as we remove only alternate samples and most of the data still remains.. Now, in case of interpolation with zero padding, we add 0 using overlap add method. Here ~~is~~ the high frequency components are replaced by 0 which means pitch is being ~~suppre~~ ~~sup~~ suppressed. If the frequency was ~~initi~~ initially at $2\pi$ we downsampled it to $\pi$. So, frequency suppression happens and lots of low frequency components come up. The voice file is heard but we can still understand but the voice is robotic sound. We didnt remove all the data but only halfed the data. So the time index is relatively short. Code and outputs are stored and uploaded with submission.

(7)

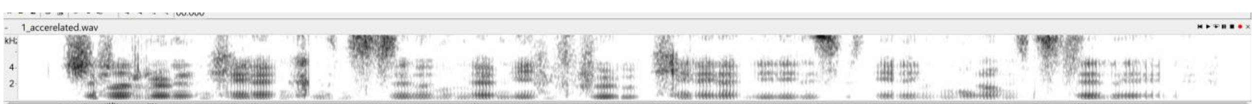
1.wav



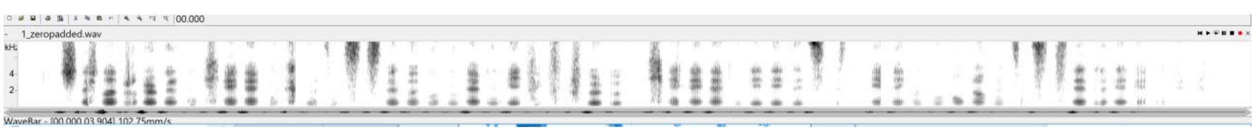

Original signal



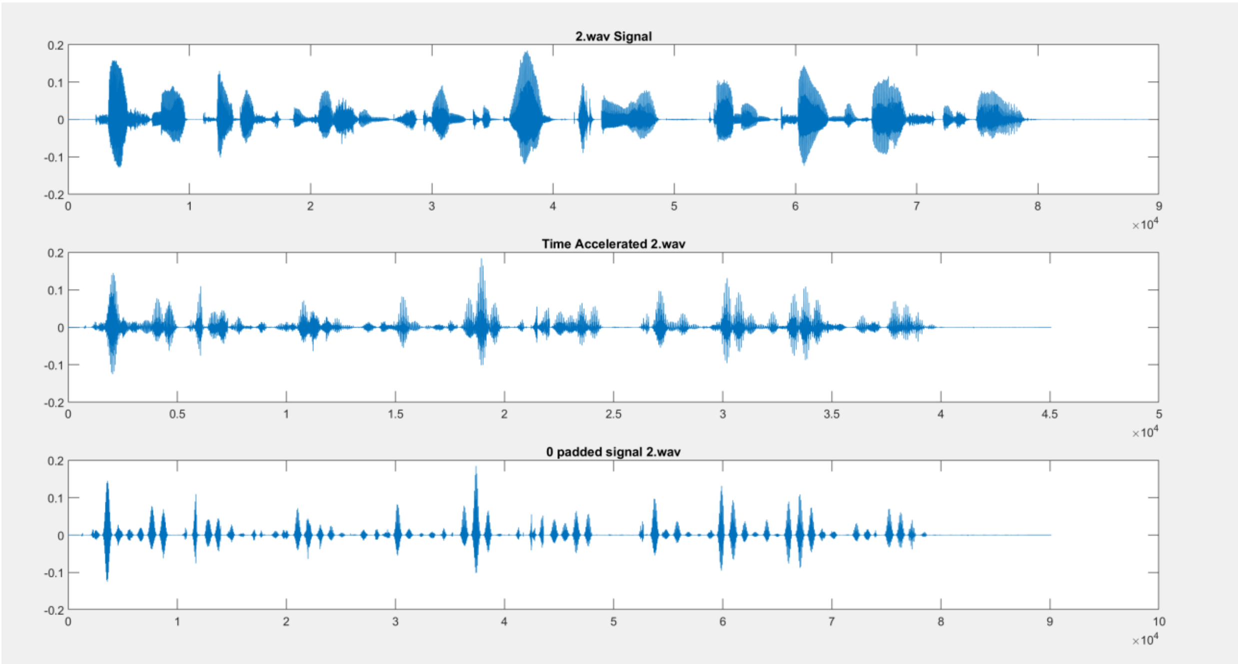

Time accelerated signal



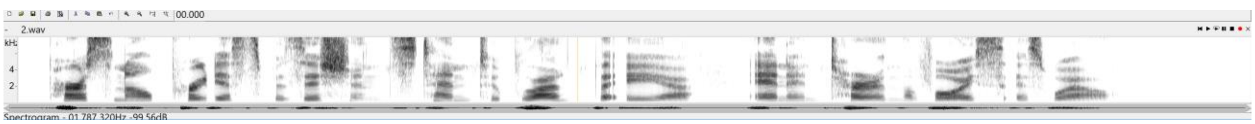

Time interpolated signal



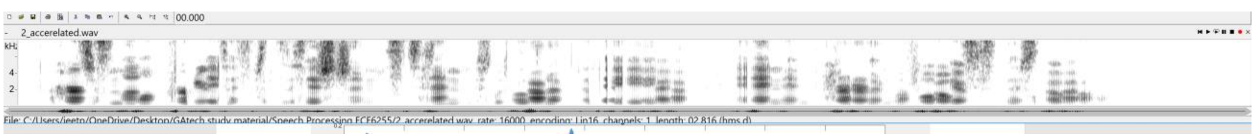
2.wav



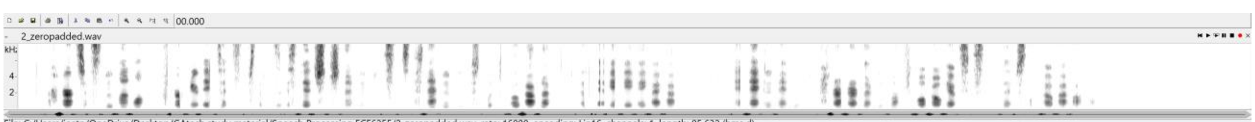

Original signal



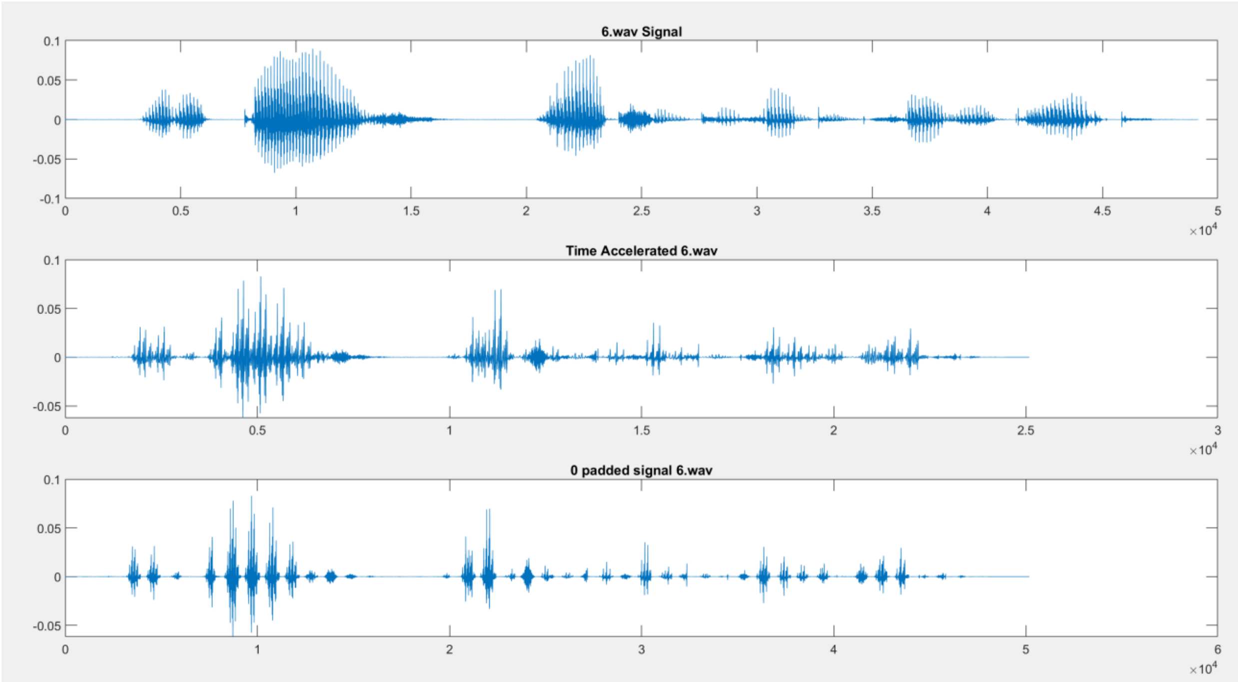

Time accelerated signal



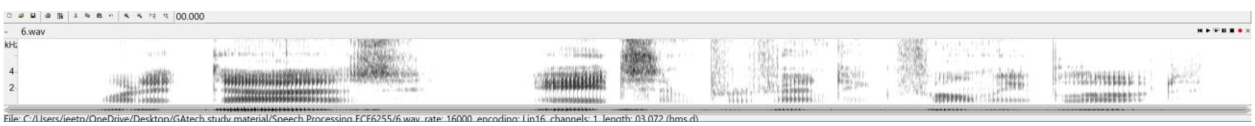

Time interpolated signal



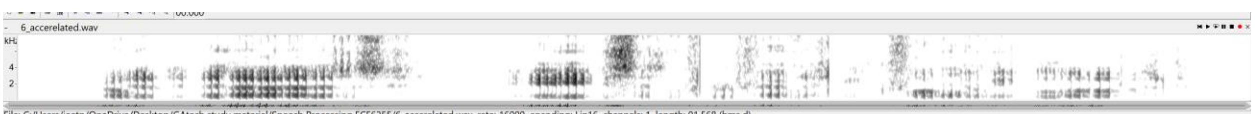
6.wav




Original signal




Time accelerated signal




Time interpolated signal