ECE5022: Digital Signal Processing

Lecture 11: Homomorphic Signal Processing

May 31, 2022

Instructor: Professor Bowon Lee
Email: bowon.lee@inha.ac.kr | Web: dsp.inha.ac.kr
Department of Electrical and Computer Engineering
Inha University

1

Announcements

- Term Project open book No computer / tablet / phone
  - Due: 6/14
- Final Exam
  - 4:30PM – 5:30PM (60 Minutes) on Tuesday, 6/7
  - Place: Hi-Tech Center Room 228

(FFT 포함)
중간이후 전범위

- Grading
  - Exams: 30% (Midterm) + 30% (Final)
  - Project: 20%, Attendance: 10%, Homework: 10%

2

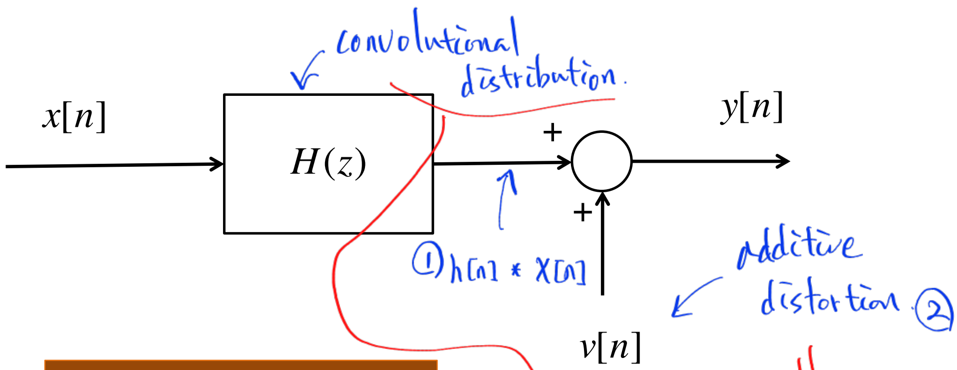
Outline

- Homomorphic Signal Processing
- Cepstrum Definitions (Analysis
  - Real Cepstrum
  - Complex Cepstrum
- Homomorphic Deconvolution
- Speech Processing Applications

Homomorphic Signal Processing 3/38

3

Signal Distortion Model



Convolutive Distortion

$$y[n] = \underbrace{h[n]*x[n]}_{①} + \underbrace{v[n]}_{②} \Rightarrow Y(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega}) + V(e^{j\omega})$$

Additive Distortion

Homomorphic Signal Processing 4/38

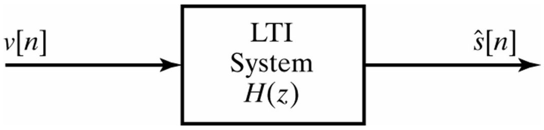
4

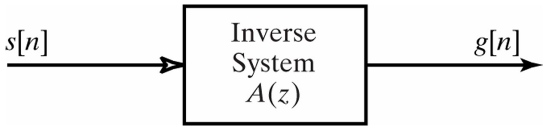
Two Types of Signal Distortion

- Additive Noise
  - Distortion caused by a noise signal ***added*** to the input
  - Example: Wiener Filtering for reducing additive noise
- Convolutive Noise
  - Distortion caused by ***convolving (filtering)*** the input with a noise signal

Removing Convolutive Noise

- Convolutive Noise
  - Need to find an inverse filter to remove it
  - Example: All-pole modeling gives an FIR inverse filter





$$H(z) = \frac{1}{1 - \sum_{k=1}^{p} a_k z^{-k}}$$

$$A(z) = \frac{1}{H(z)} = 1 - \sum_{k=1}^{p} a_k z^{-k}$$

Homomorphic Signal Processing

- Parametric signal modeling provides a simple inverse filter as far as the model is good enough
- Nonparametric approach may be required if the model is not good enough
- Convert the problem to an easier ~~solution~~ problem to solve.
  - Transform convolutive noise to additive noise

Homomorphic Signal Processing 7/38

7

Superposition Property

- Given a linear operator $L[\bullet]$

$$L\left[a_1x_1[n]+a_2x_2[n]\right]=a_1L\left[x_1[n]\right]+a_2L\left[x_2[n]\right]$$

z 변환

- Equivalently

$$L\left[a_1X_1(z)+a_2X_2(z)\right]=a_1L\left[X_1(z)\right]+a_2L\left[X_2(z)\right]$$

- Wiener filter can be used to remove the additive noise

Homomorphic Signal Processing 8/38

8

Generalized Superposition

- Given a nonlinear operator Φ that satisfies

$$\Phi[X_1(z)X_2(z)] = \Phi[X_1(z)] + \Phi[X_2(z)]$$

- If such a nonlinear operation exists, we can treat the convolutive noise as additive noise

problem

convolutive noise → additive noise

same shape

- Homomorphic signal processing
  - Convert the problem to an easier problem and solve
  - Transform the output signal back to the original domain

homomorphic signal process.

Complex Logarithm I/II

- We can easily see that

$$\log_e[X_1(z)X_2(z)] = \log_e[X_1(z)] + \log_e[X_2(z)]$$

- For the DTFT (complex valued)

$X(e^{jw}) = |X(e^{jw})| e^{j\angle(e^{jw})}$

$$\log_e[X(e^{j\omega})] = \log_e\left[|X(e^{j\omega})| e^{j\angle X(e^{j\omega})}\right]$$
$$= \log_e\left[|X(e^{j\omega})|\right] + \log_e\left[e^{j\angle X(e^{j\omega})}\right]$$
$$= \log_e\left[|X(e^{j\omega})|\right] + j\angle X(e^{j\omega})$$

log of the magnitude of j times phase?

Complex Logarithm II/II

- Suppose

$$X(e^{j\omega}) = X_1(e^{j\omega})X_2(e^{j\omega})$$

- Then

실수(+짝수) even 홀수(odd) magnitude phase

$$\log\left[X(e^{j\omega})\right] = \log\left[\left(\left|X_1(e^{j\omega})\right| e^{j\angle X_1(e^{j\omega})}\right)\left(\left|X_2(e^{j\omega})\right| e^{j\angle X_2(e^{j\omega})}\right)\right]$$

$$= \log\left[\left|X_1(e^{j\omega})\right|\left|X_2(e^{j\omega})\right| e^{j\left(\angle X_1(e^{j\omega}) + \angle X_2(e^{j\omega})\right)}\right]$$

$$= \log\left|X_1(e^{j\omega})\right| + \log\left|X_2(e^{j\omega})\right| + j\left(\angle X_1(e^{j\omega}) + \angle X_2(e^{j\omega})\right)$$

Outline

- Homomorphic Signal Processing
- Cepstrum Definitions 켑스트럼.
  - Real Cepstrum
  - Complex Cepstrum
- Homomorphic Deconvolution
- Speech Processing Applications

Real Cepstrum

- For a real sequence $x[n]$
  - $|X(e^{j\omega})|$ is ***even*** and $\angle X(e^{j\omega})$ is ***odd*** and both are ***real***
  - $\log|X(e^{j\omega})|$ is ***real*** and ***even*** and $j\angle X(e^{j\omega})$ is ***imaginary*** and ***odd***
- If we define

$$C_x(e^{j\omega}) = \log\left|X(e^{j\omega})\right|$$

whose inverse DTFT is defined as the Real Cepstrum

$$c_x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} C_x(e^{j\omega})e^{j\omega n}\,d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left|X(e^{j\omega})\right| e^{j\omega n}\,d\omega$$

Real Cepstrum

Homomorphic Signal Processing 13/38

13

되돌릴수있음!

각스텝은 invertable 함.

Complex Cepstrum

- If we define

convolutive noise in time domain

$X[n] \xrightarrow{①} \boxed{DTFT} \rightarrow \boxed{\log} \xrightarrow{③} \boxed{IDTFT} \rightarrow \hat{x}[n]$

$\boxed{IDTFT} \leftarrow \boxed{\exp} \leftarrow \boxed{DTFT}$

②

additive noise in freq. domain

$$\hat{X}(e^{j\omega}) = \log\left[X(e^{j\omega})\right]$$

- The complex cepstrum is defined as

항상 real

inverse dft 결과

$$\hat{x}[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{X}(e^{j\omega})e^{j\omega n}\,d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left[X(e^{j\omega})\right]e^{j\omega n}\,d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\log\left|X(e^{j\omega})\right| + j\angle X(e^{j\omega})\right]e^{j\omega n}\,d\omega$$

Even — real part

imaginary part — Odd

Homomorphic Signal Processing 14/38

14

Invertibility of Complex Cepstrum

- The inverse DTFT is invertible, i.e., DTFT
- The inverse of logarithm is invertible, i.e., exponential
- The complex cepstrum is therefore invertible
- The cepstrum of $X(e^{j\omega}) = X_1(e^{j\omega})X_2(e^{j\omega})$

$$\hat{X}(e^{j\omega}) = \hat{X}_1(e^{j\omega}) + \hat{X}_2(e^{j\omega})$$
$$= \log\left|X(e^{j\omega})\right| + j\angle X(e^{j\omega})$$

- The cepstrum is guaranteed to be real for real $x[n]$

Some Terminologies

- Quefrency
  - The time index n for the cepstrum

$x[n] \rightarrow \square \rightarrow \square \rightarrow \square \rightarrow \hat{x}[n]$

time index

~~time~~ ~~freq~~

quefrency domain

$$\hat{x}[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left[X(e^{j\omega})\right] e^{j\omega n}\, d\omega$$

  - Reverse of frequency -> quefrency 큐프런시.

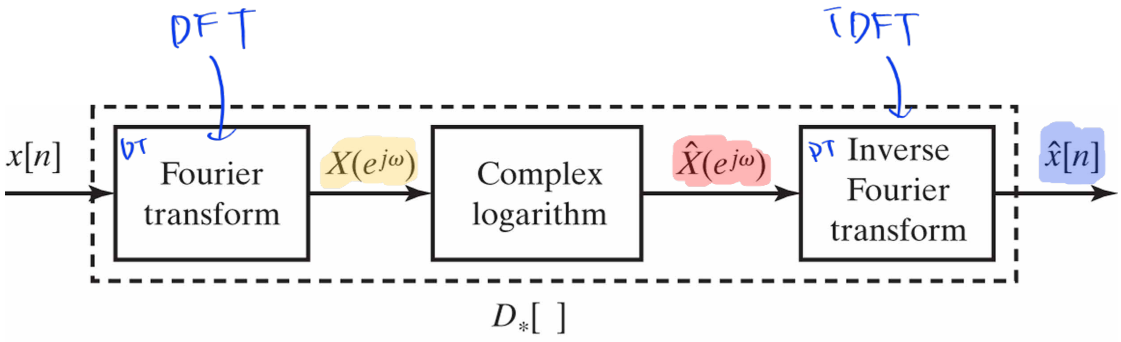
filtering

- Liftering 리프터링
  - Filtering in the cepstrum domain
  - Reverse of filtering -> liftering

시간축

↳ 새로운 도메인으로 봐야!!

Complex Cepstrum Diagram



Homomorphic Signal Processing 17/38

17

Computation of Complex Cepstrum

- The computation is represented as

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}$$ **DTFT**

$$\hat{X}(e^{j\omega}) = \log\left[X(e^{j\omega})\right]$$ **Complex Logarithm**

$$\hat{x}[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \hat{X}(e^{j\omega})e^{j\omega n}\,d\omega$$ **Inverse DTFT**

Homomorphic Signal Processing 18/38

18

Complex Cepstrum in DFT

- Sampling the DTFT

$$X[k] = X(e^{j\omega})\Big|_{\omega=\frac{2\pi k}{N}} = \sum_{n=-\infty}^{\infty} x[n]e^{-j\frac{2\pi k}{N}n}$$

DFT

FFT 누구는 FFT해도 됨.

유사해서 쓰긴함.

$$\hat{X}[k] = \log\left[X(e^{j\omega})\right]\Big|_{\omega=\frac{2\pi k}{N}}$$

Complex Logarithm

$$\hat{x}_p[n] = \frac{1}{N}\sum_{k=0}^{N-1}\hat{X}[k]e^{-j\frac{2\pi n}{N}k}$$

Inverse DFT

FFT

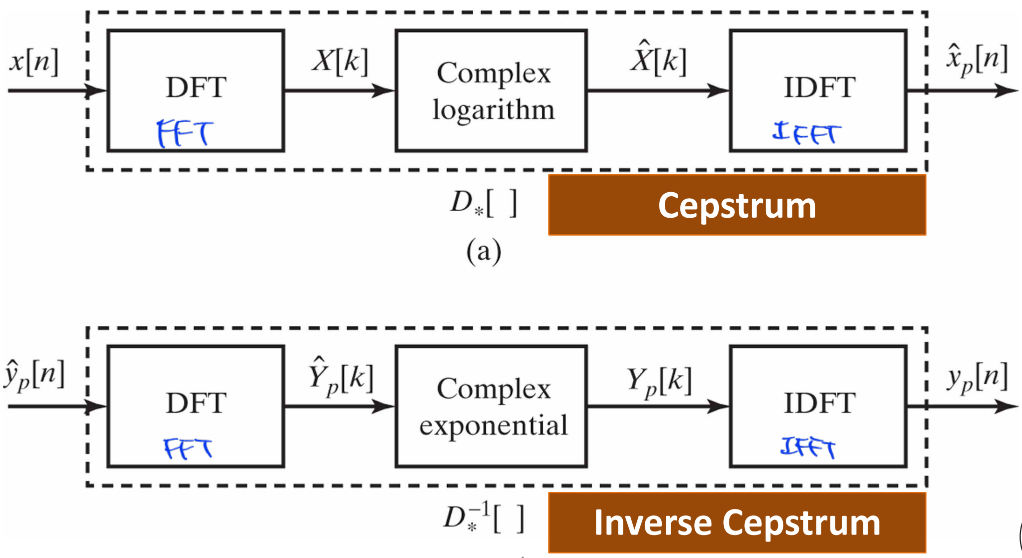
incorrect!!

엄밀히는 틀림.

Homomorphic Signal Processing 19/38

19

Cepstrum and Inverse Cepstrum in DFT



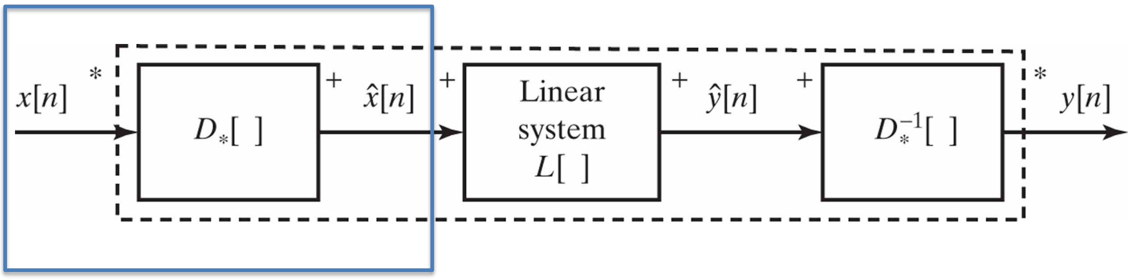
Homomorphic Signal Processing 20/38

20

Outline

- Homomorphic Signal Processing
- Cepstrum Definitions
  - Real Cepstrum
  - Complex Cepstrum
- Homomorphic Deconvolution
- Speech Processing Applications

Homomorphic Signal Processing I/II

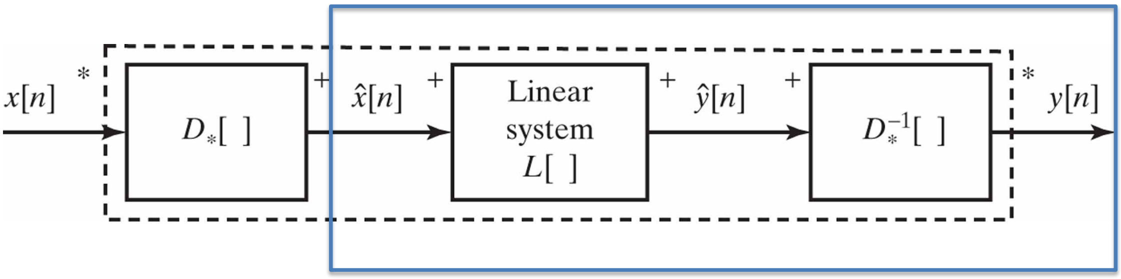


$$x[n] = x_1[n] * x_2[n] \Rightarrow X(z) = X_1(z) \cdot X_2(z)$$

$$\hat{X}(z) = \log\left[X(z)\right] = \log\left[X_1(z)\right] + \log\left[X_2(z)\right] = \hat{X}_1(z) + \hat{X}_2(z)$$

$$\hat{x}[n] = D_*\left[x_1[n] * x_2[n]\right] = \hat{x}_1[n] + \hat{x}_2[n]$$

Homomorphic Signal Processing II/II



$$\hat{x}[n] = \hat{x}_1[n] + \hat{x}_2[n]$$

$$\hat{y}[n] = L\left[\hat{x}_1[n] + \hat{x}_2[n]\right] = L\left[\hat{x}_1[n]\right] + L\left[\hat{x}_2[n]\right] = \hat{y}_1[n] + \hat{y}_2[n]$$

$$\hat{y}[n] = D_*^{-1}\left[\hat{y}_1[n] + \hat{y}_2[n]\right] = y_1[n] * y_2[n]$$

Homomorphic Signal Processing 23/38

23

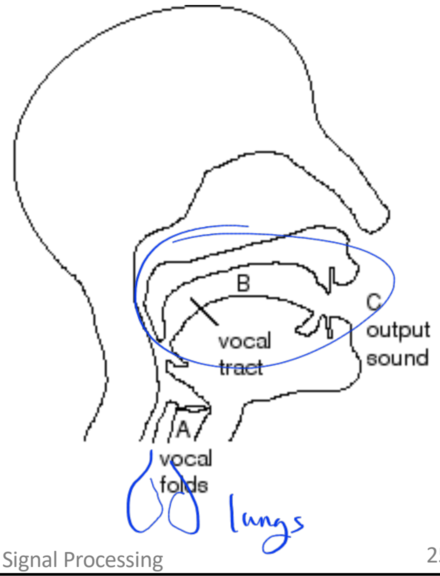
Outline

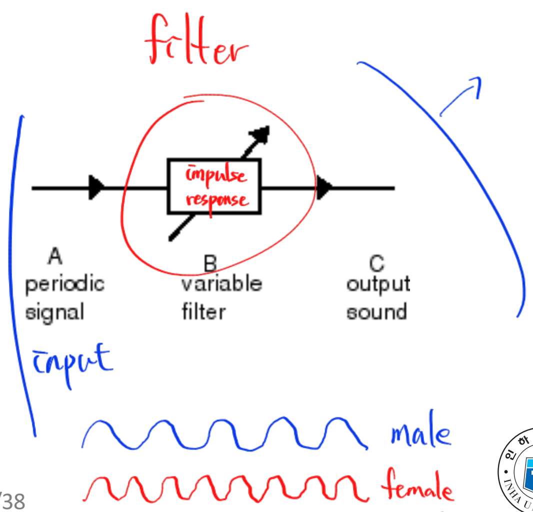
- Homomorphic Signal Processing
- Cepstrum Definitions
  - Real Cepstrum
  - Complex Cepstrum
- Homomorphic Deconvolution
- Speech Processing Applications

Homomorphic Signal Processing 24/38

24

Source-Filter Model for Speech Production





음성인식

Cepstrum 사용

deconvolution

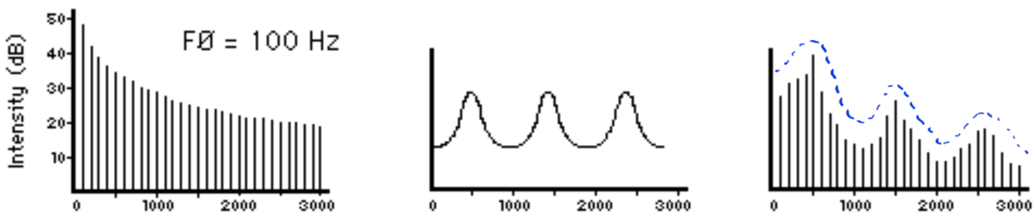
핵심

고주파!

Homomorphic Signal Processing 25/38

25

Source-Filter Model for Voiced Speech



Excitation	Vocal Tract	Output
$E(e^{j\omega})$	$V(e^{j\omega})$	$X(e^{j\omega}) = E(e^{j\omega})V(e^{j\omega})$

Homomorphic Signal Processing 26/38

26

Logarithm of the Fourier Transform

- The output signal $x[n]$ is the convolution of
  - The excitation signal $e[n]$ and the vocal tract filter $v[n]$
- In the frequency domain

$$X(e^{j\omega}) = E(e^{j\omega})V(e^{j\omega})$$

- If we take the magnitude, then

$$|X(e^{j\omega})| = |E(e^{j\omega})||V(e^{j\omega})|$$

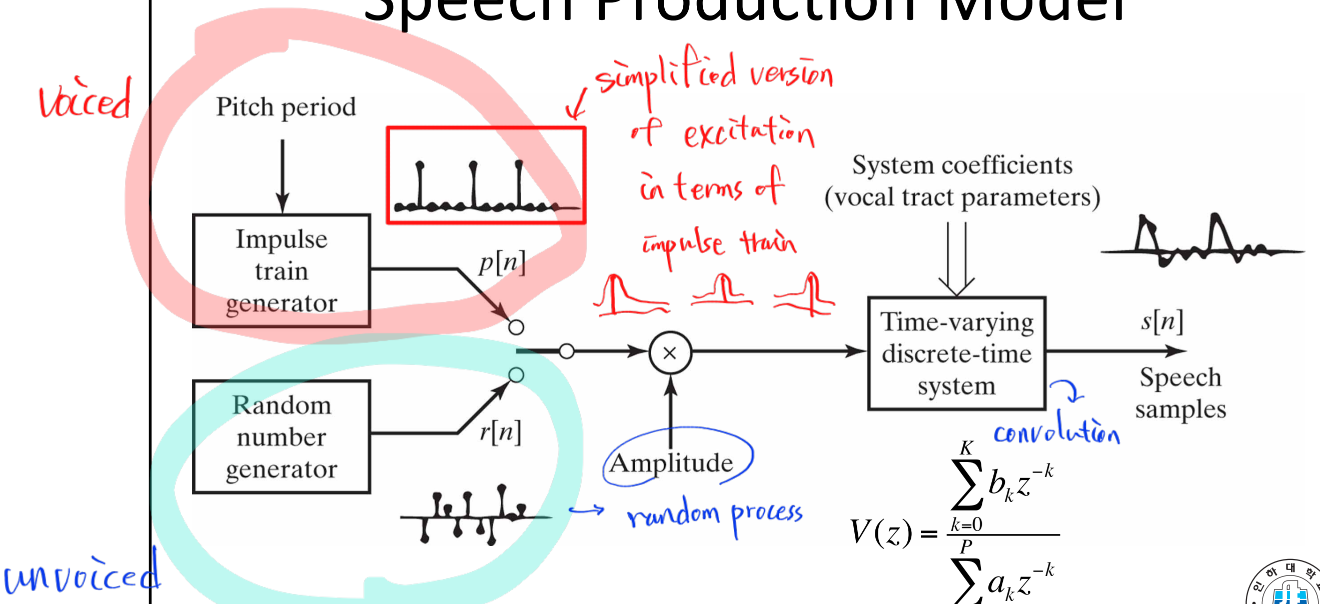
- If we take the logarithm, then

$$\log|X(e^{j\omega})| = \log|E(e^{j\omega})| + \log|V(e^{j\omega})|$$

Homomorphic Signal Processing 27/38

27

Speech Production Model



$$V(z) = \frac{\sum_{k=0}^{K} b_k z^{-k}}{\sum_{k=0}^{P} a_k z^{-k}}$$

Homomorphic Signal Processing 28/38

28

* noise ? => unwanted signal.

즉, 목적에 따라 noise는 달라짐.

Speech Deconvolution

- For voiced speech

$$s[n] = v[n] * p[n]$$

- After windowing

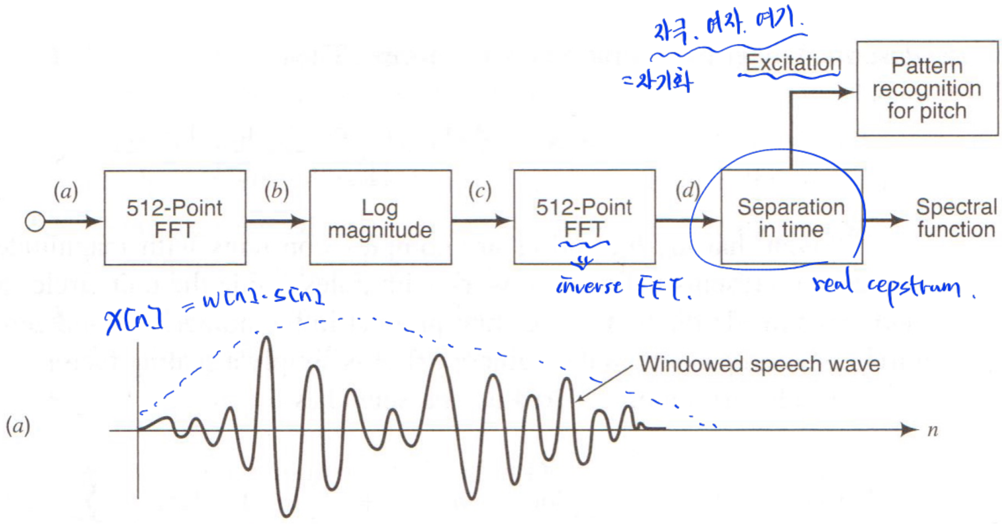
$$\underline{w[n]}s[n] = v[n] * \underline{p_w[n]}, \quad p_w[n] = w[n]p[n]$$

- The complex cepstrum

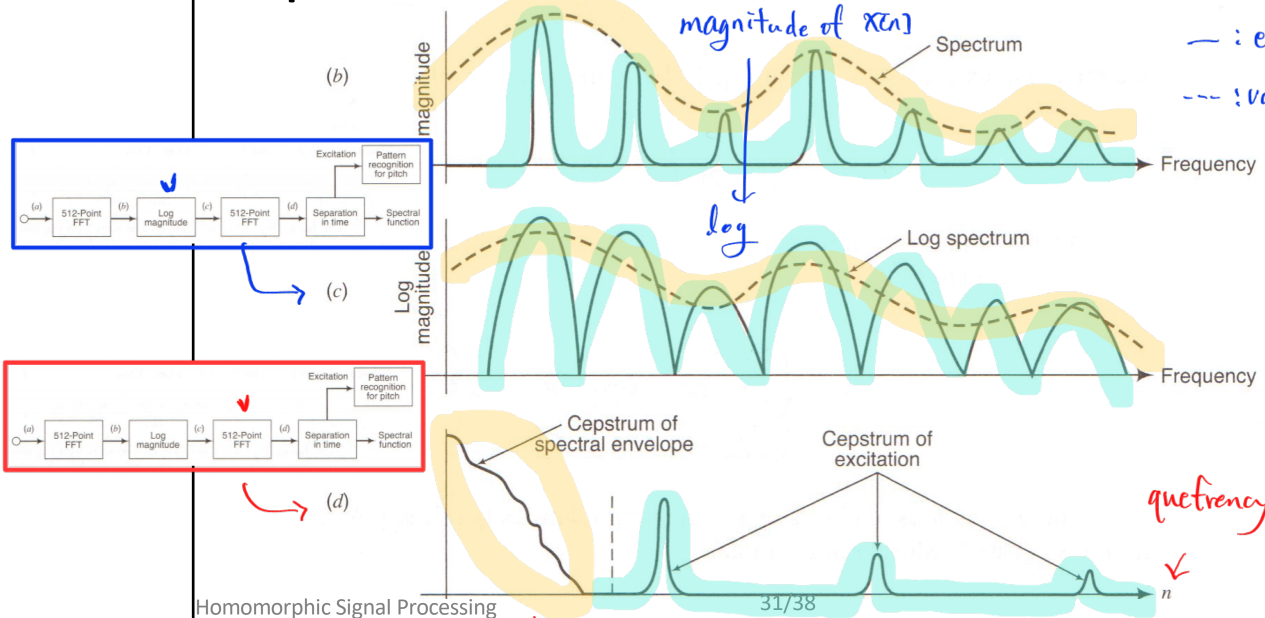
$$\hat{x}[n] = \hat{v}[n] + \hat{p}_w[n]$$

windowing. to preserve this part.

Speech Deconvolution Overview I/II



Speech Deconvolution Overview II/II



— : excitation (P[n])

--- : vocal tract (V[n])

excitation 고정

vocal 바꾸면

lower part 바뀜

반대로

excitation 바꾸면

higher part 바뀜.

31

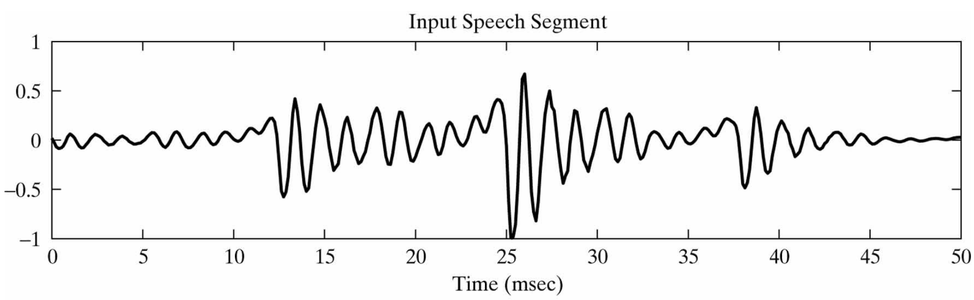
inverse cepstrum 하면

vocal part 정보 겟.

이쪽 inverse cepstrum 하면

excitation part 겟.

Speech Deconvolution Example I/V

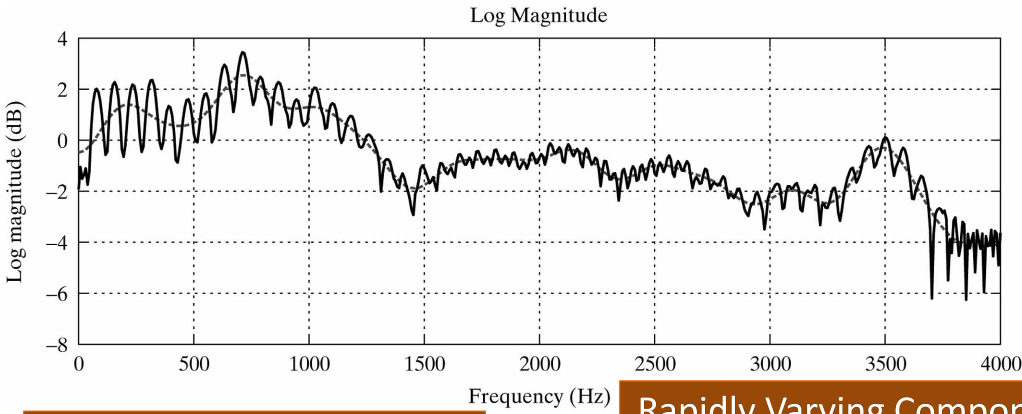


- Input speech signal after windowing

$$x[n] = w[n]s[n] = v[n] * p_w[n], \quad p_w[n] = w[n]p[n]$$

32

Speech Deconvolution Example II/V

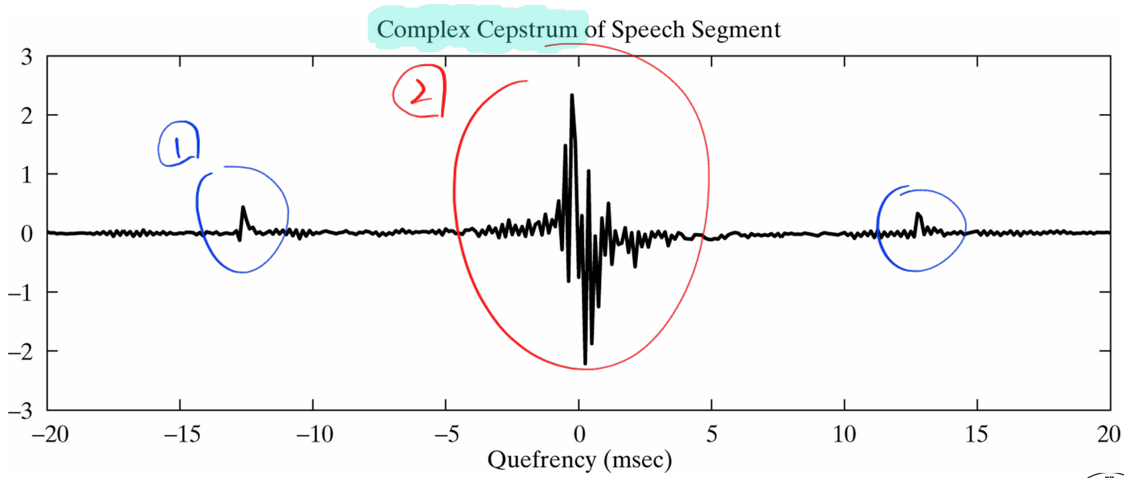


Slowly Varying Component

Rapidly Varying Component

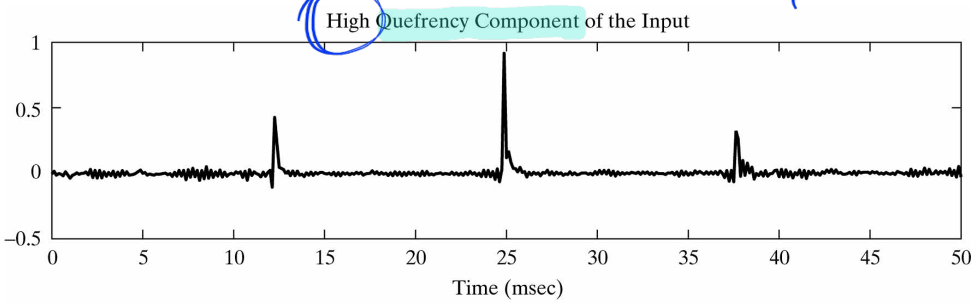
$$x[n] = v[n] * p_w[n] \Rightarrow V(e^{j\omega}) P_w(e^{j\omega})$$

Speech Deconvolution Example III/V



Speech Deconvolution Example IV/V

invers cepstrum of (1)

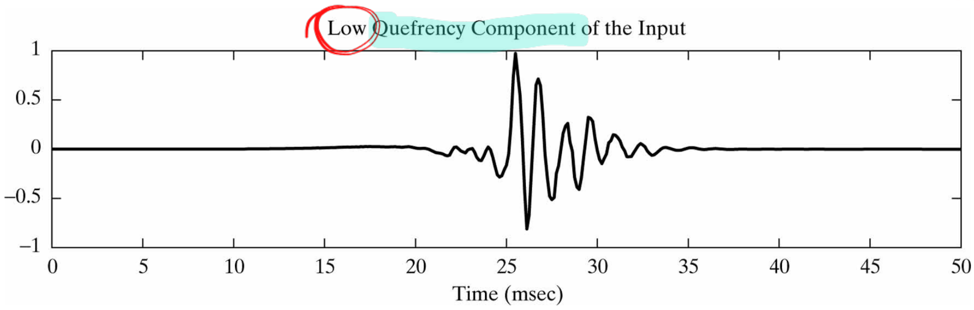


- High quefrency component

$$p_w[n] = w[n]p[n]$$

Speech Deconvolution Example V/V

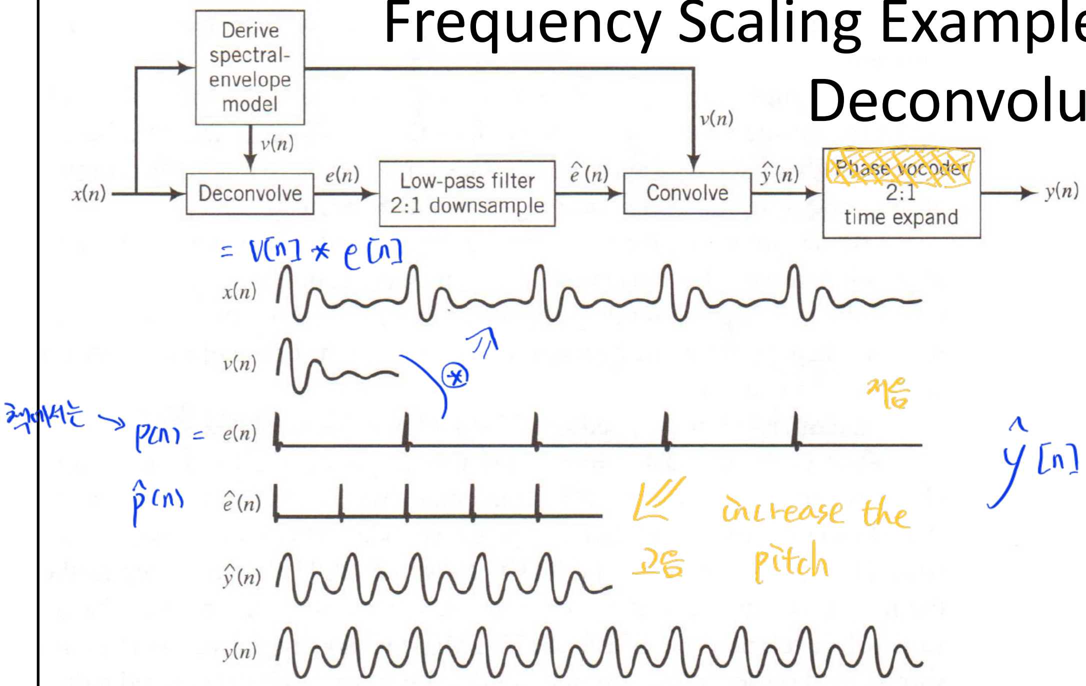
inverse cepstrum of (2)



- Low quefrency component

$$v[n]$$

Frequency Scaling Example via Deconvolution



$\hat{y}$[n] has higher pitch than x[n]

37

feature extraction → knocker 가능.

Questions?

38