

# Speech Enhancement



# Noise Sources

- Distant Microphone
- Background environment (indoor, outdoor, factory, ...etc..)
- Recording device (transducer)
- Transmission channel
- Reverberation
- Interfering speakers

# Objectives of Speech Enhancement

- Listener's Perspective
  - ✓ Intelligibility
  - ✓ Quality (Naturalness)
  - ✓ Auditory Fatigue (listening effort)
- In view of Development of speech systems
  - ✓ Feature normalization/compensation

# Applications of Speech Enhancement

- SE on devices using speech (mobiles, telephone, ...)
- Speech generated from noisy environments
- Military applications (Battle fields)
- Aircraft cockpits
- Development of speech systems
- Perceptual speech enhancement
- SE for parameter normalization

# Evaluation

- Objective Evaluation
  - ✓ SNR, Seg-SNR, PESQ, ...
- Subjective Evaluation
  - ✓ Listening tests
    - Intelligibility
    - Naturalness (MOS)
- Quality : Voiced & Periodic (Harmonic) segments
- Intelligibility : Consonant regions
- Fricatives : Highly affected by noise

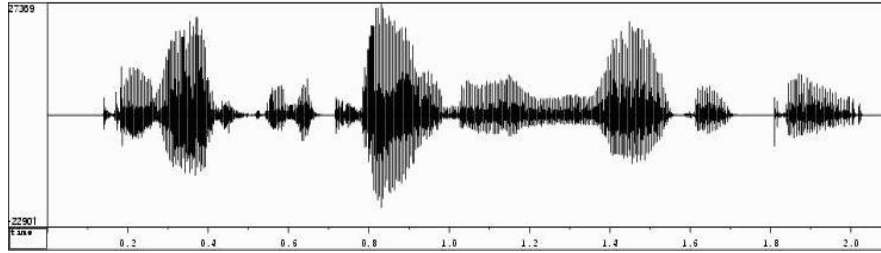
# Types of Noises : SE Techniques

- Broadband noise
- Impulsive noise
- Band-limited noise
- Steady vs time-varying noise
- Techniques for noise removal
  - ✓ Additive noise : Noise subtraction or Noise filtering
  - ✓ Multiplicative noise : Log + Subtraction
  - ✓ Convolutional noise : Frequency transform + Log + Subtraction
- SE Techniques
  - ✓ Spectral subtraction
  - ✓ Harmonic filtering
  - ✓ Parametric resynthesis
  - ✓ Spectral vs Excitation source modification
  - ✓ Enhance speech regions and deemphasize non-speech/noisy regions  
(Perceptual enhancement)

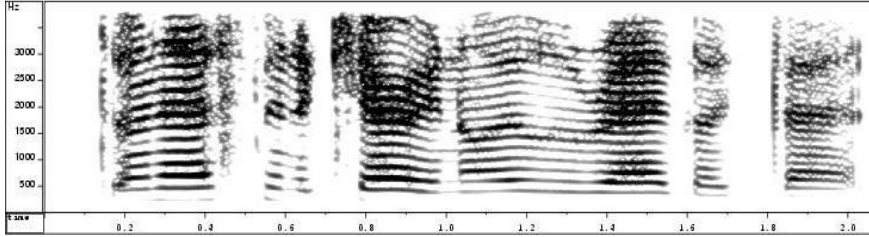
# Speech Signals and Spectrograms

## CLEAN SPEECH

File: F:\Org\k\war Page: 1 of 1 Printed: Mon Mar 01 10:54:22

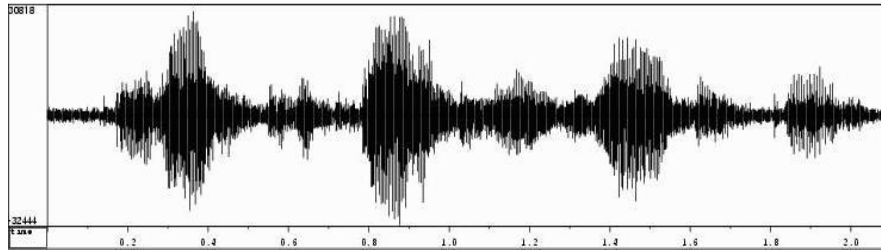


File: home\stareocharu\k\war\F Org\k\war Page: 1 of 1 Printed: Mon Mar 01 10:55:12

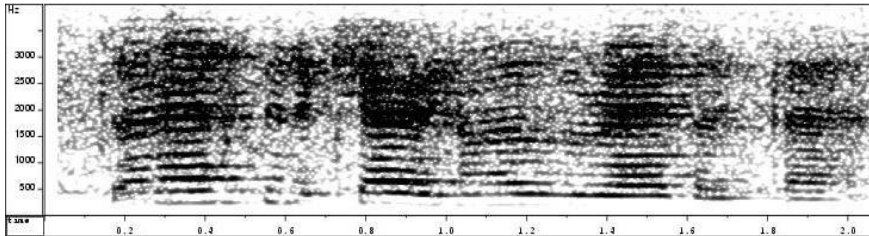


## DEGRADED SPEECH

File: home\stareocharu\k\war\F Deg\k\war Page: 1 of 1 Printed: Mon Mar 01 10:55:52

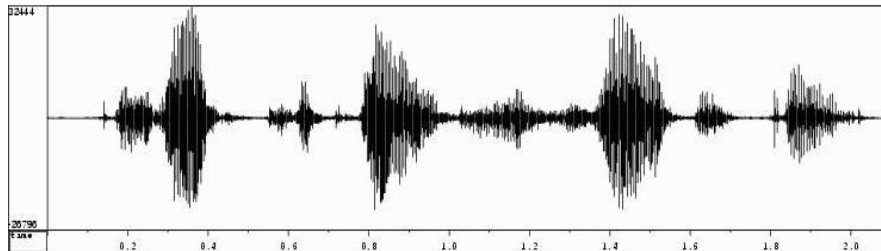


File: home\stareocharu\k\war\F Deg\k\war Page: 1 of 1 Printed: Mon Mar 01 10:56:34

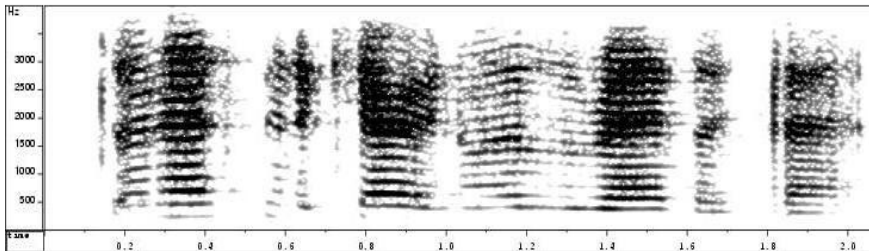


## ENHANCED SPEECH

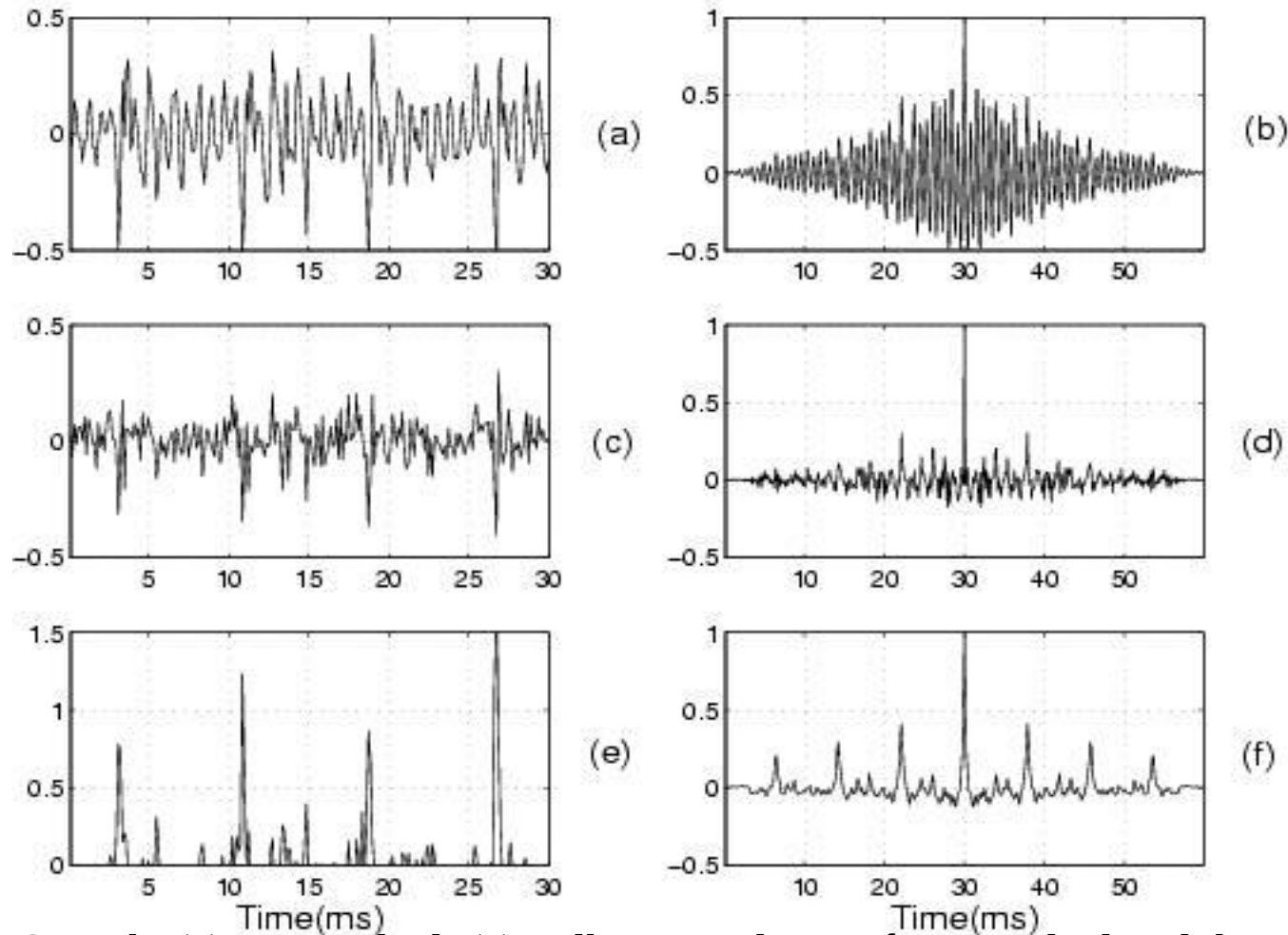
File: home\stareocharu\k\war\F Enh\k\war Page: 1 of 1 Printed: Mon Mar 01 10:58:25



File: home\stareocharu\k\war\F Enh\k\war Page: 1 of 1 Printed: Mon Mar 01 10:59:26



# Basis for exploiting the Hilbert Envelope of the LP residual for speech enhancement



(a) Speech, (c) LP residual, (e) Hilbert envelope of LP residual and their autocorrelation sequences ((b), (d) and (f), respectively)



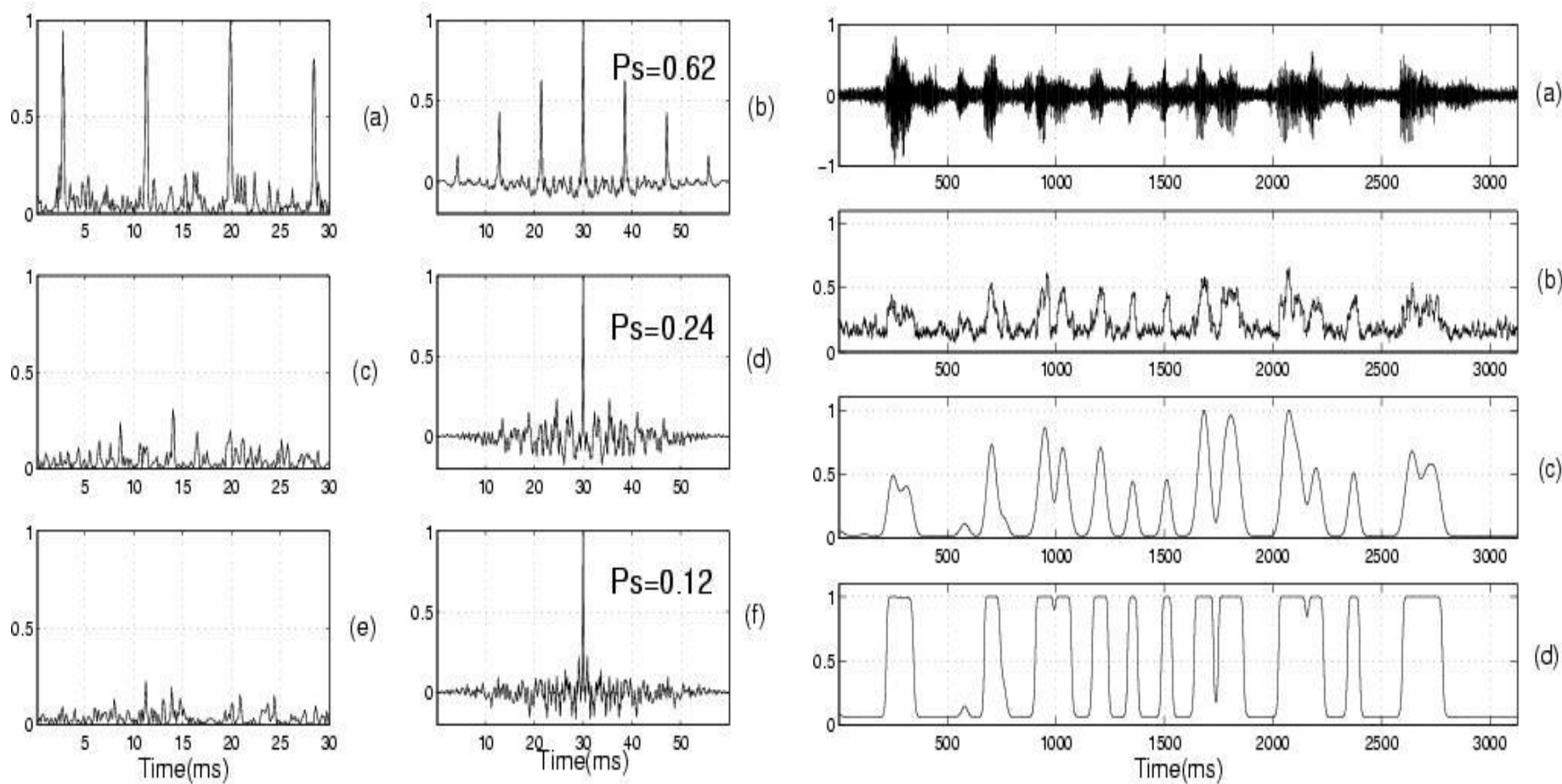
# Enhancement of Speech in Single Channel Case

**Objective:** To enhance the speech components

**Existing Methods:** Estimate characteristics of degradation and reduce the same from degraded speech

**Limitation:** Estimating degradation characteristics is a difficult task

# Enhancement of Speech in Single Channel Case: Basis

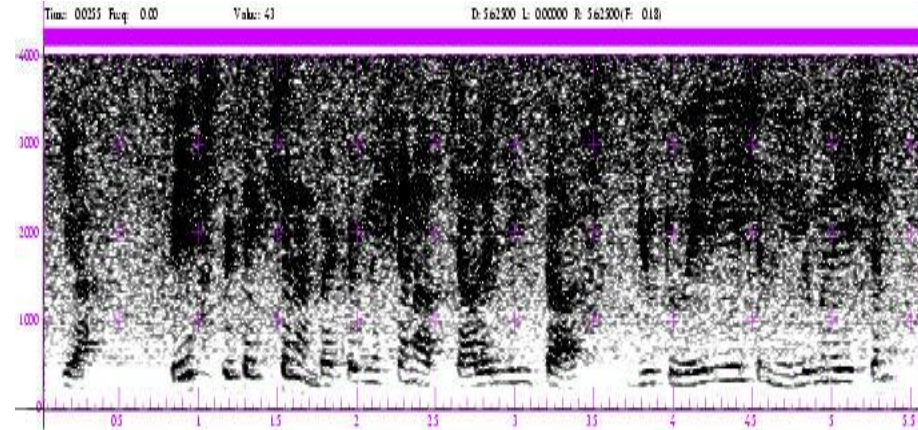
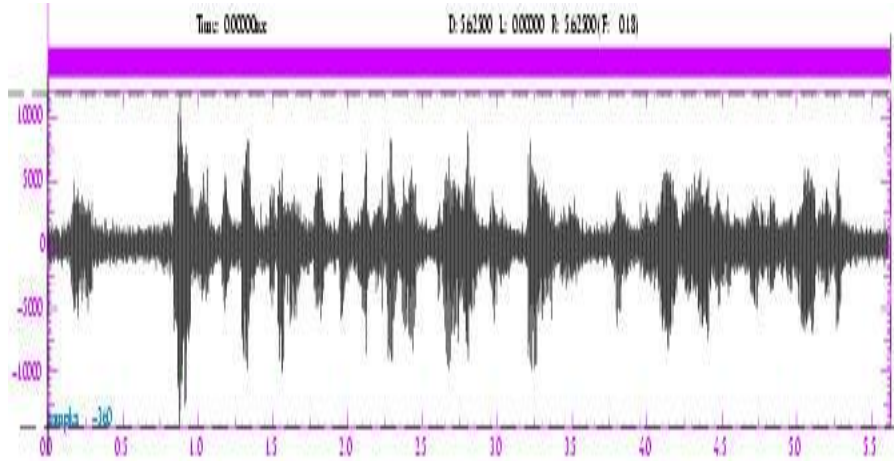


Segment of (a) High voiced speech, (c) Low voiced speech and (e) Nonspeech and their autocorrelation sequences ((b), (d) and (f), respectively)

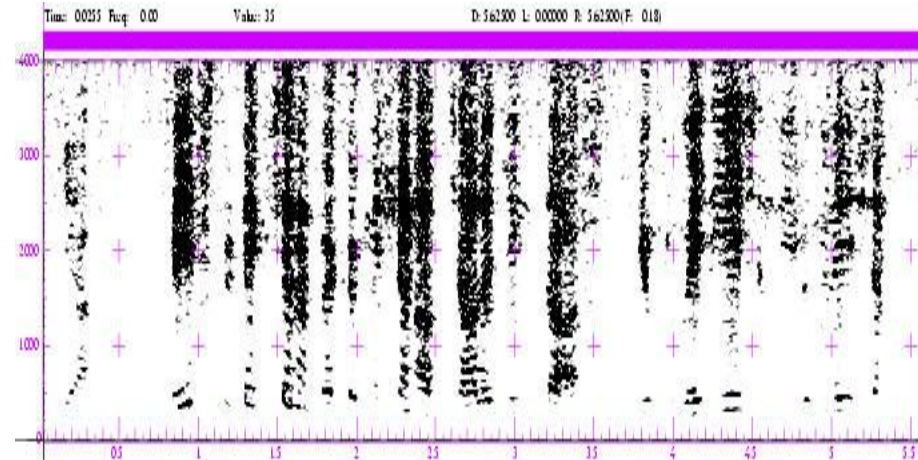
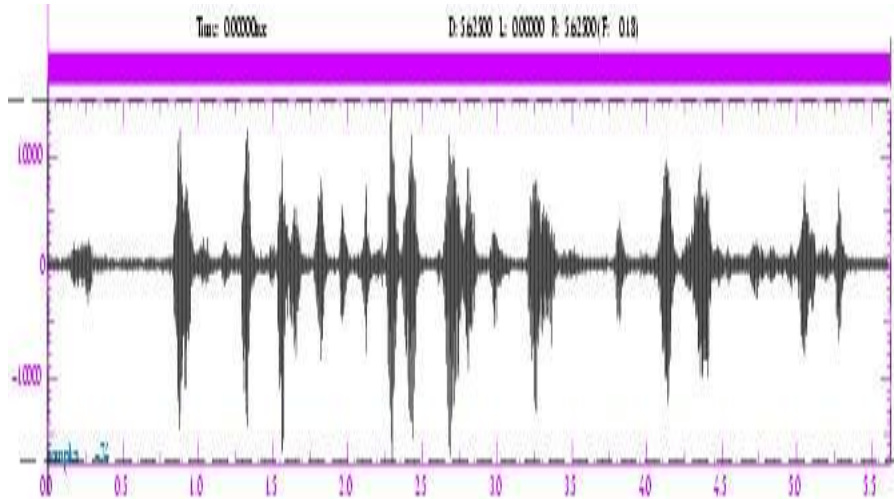
(a) Degraded speech signal, (b) normalized peak strengths, (c) smoothed peak strengths and (d) weight function to enhance speech components

# Enhancement of Speech in Single Channel Case

## DEGRADED SPEECH



## ENHANCED SPEECH

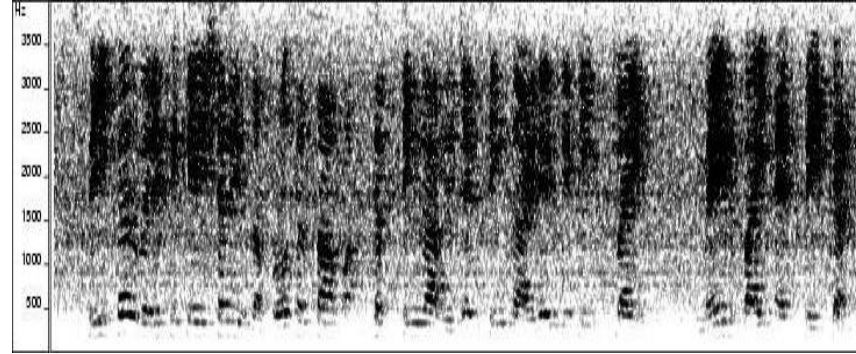
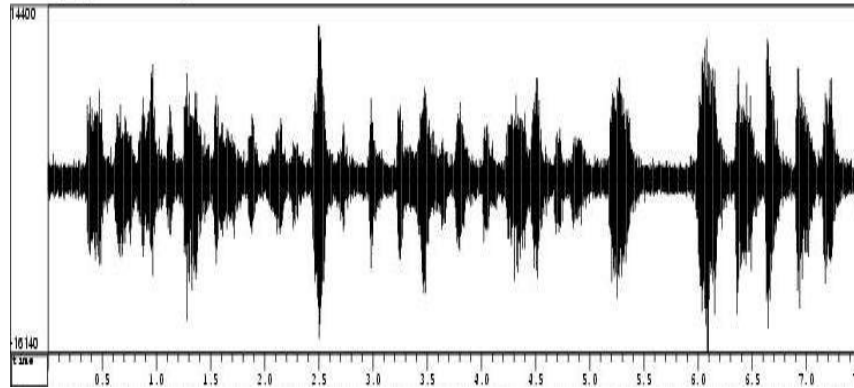




# Enhancement of Speech in Single Channel Case

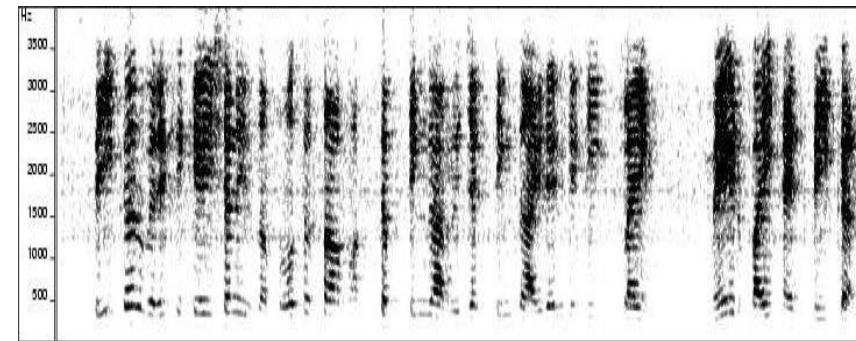
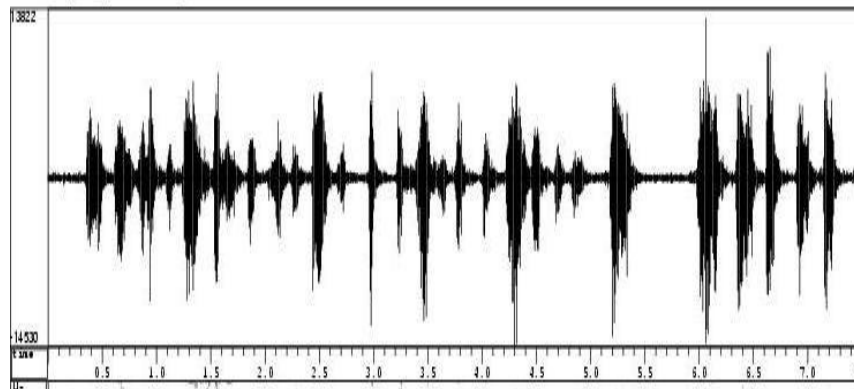
## DEGRADED SPEECH

File: OrigDegSpchFANI.wav Page: 1 of 1 Printed: Thu Nov 20 23:39:56



## ENHANCED SPEECH

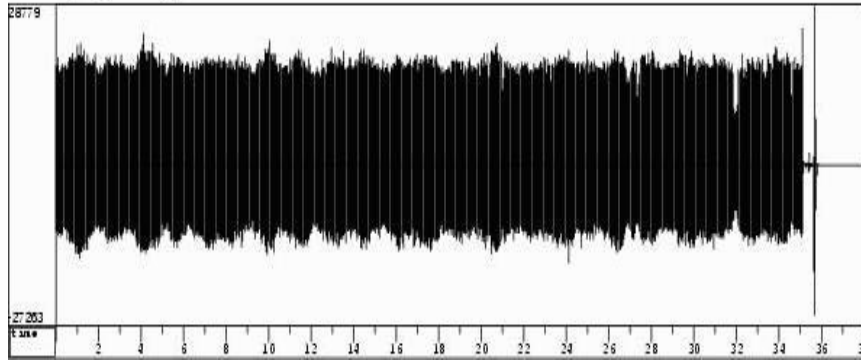
File: EnhSpchPropFANI.wav Page: 1 of 1 Printed: Thu Nov 20 23:42:21



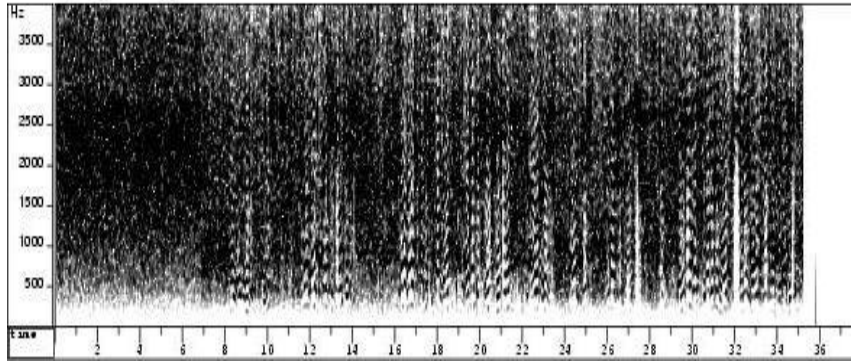
# Enhancement of Speech in Single Channel Case

## DEGRADED SPEECH

File: DegLong6.wav Page: 1 of 1 Printed: Fri Mar 19 17:58:33

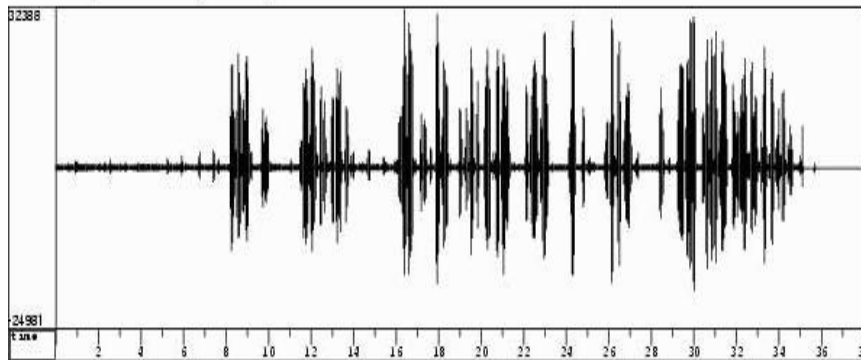


File: DegLong6.wav Page: 1 of 1 Printed: Fri Mar 19 18:02:46

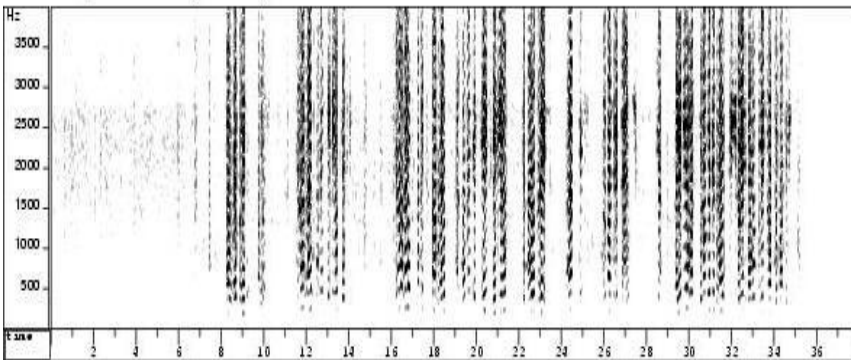


## ENHANCED SPEECH

File: /home/prasanna/EnhLong6.wav Page: 1 of 1 Printed: Fri Mar 19 18:03:29



File: /home/prasanna/EnhLong6.wav Page: 1 of 1 Printed: Fri Mar 19 18:04:14



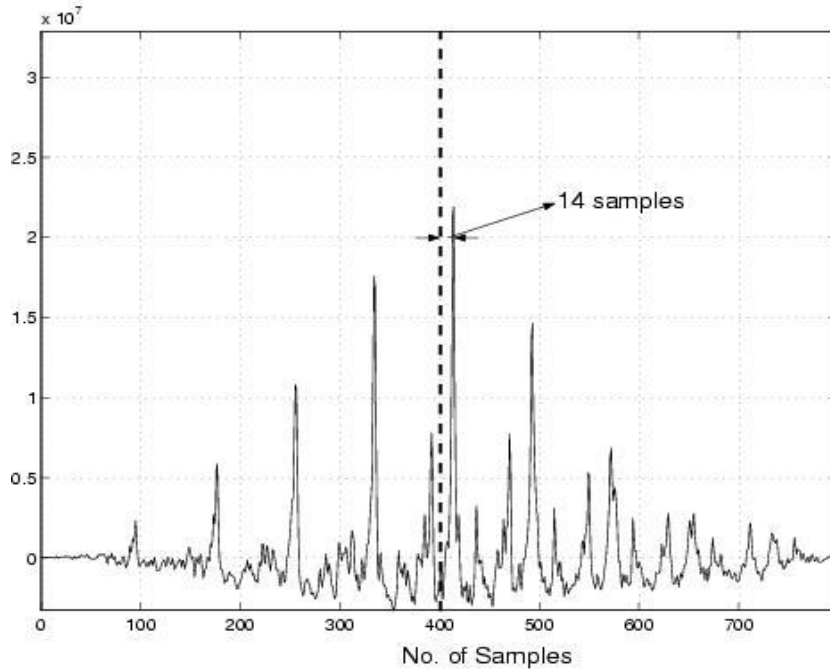
# Enhancement of Speech in Multichannel Case

**Objective:** To enhance speech components using signals from multiple microphones. Estimate delay between speech signals collected over a pair of microphones

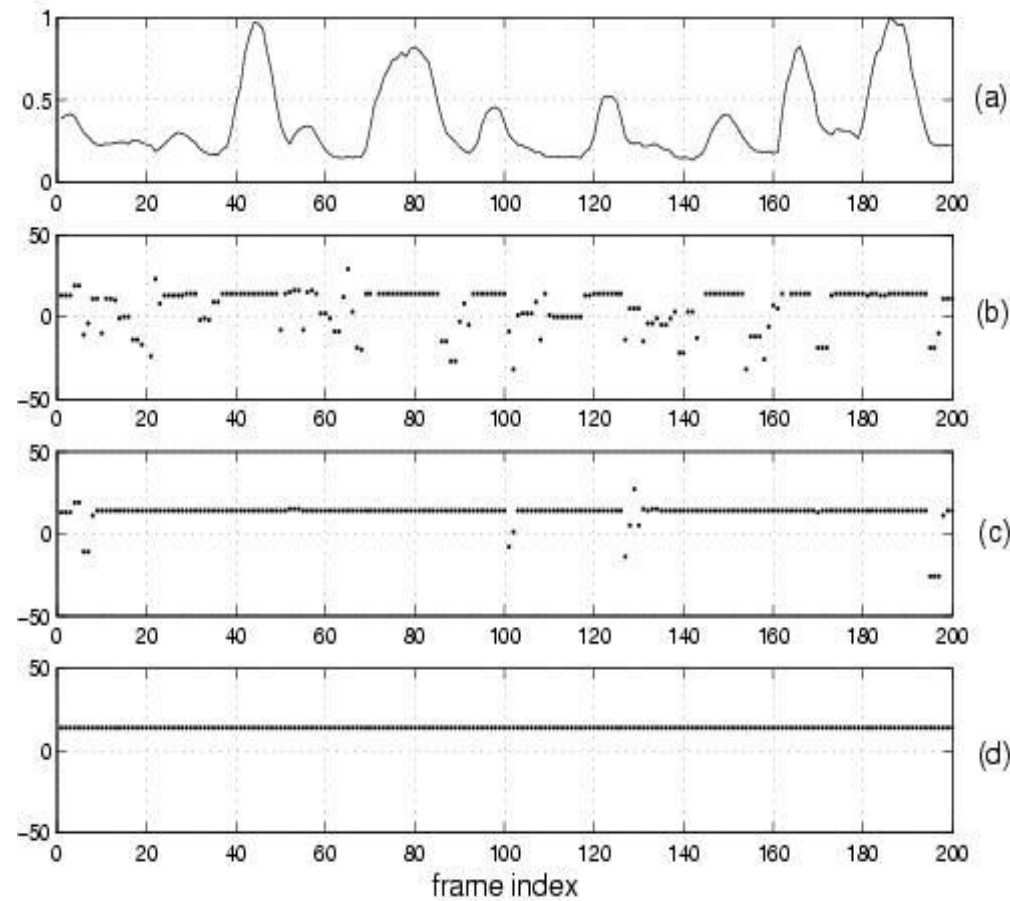
**Existing Methods:** Estimate characteristics of degradation and reduce the same from degraded speech

**Limitation:** Estimating degradation characteristics is a difficult task

# Time-Delay Estimation: Basis



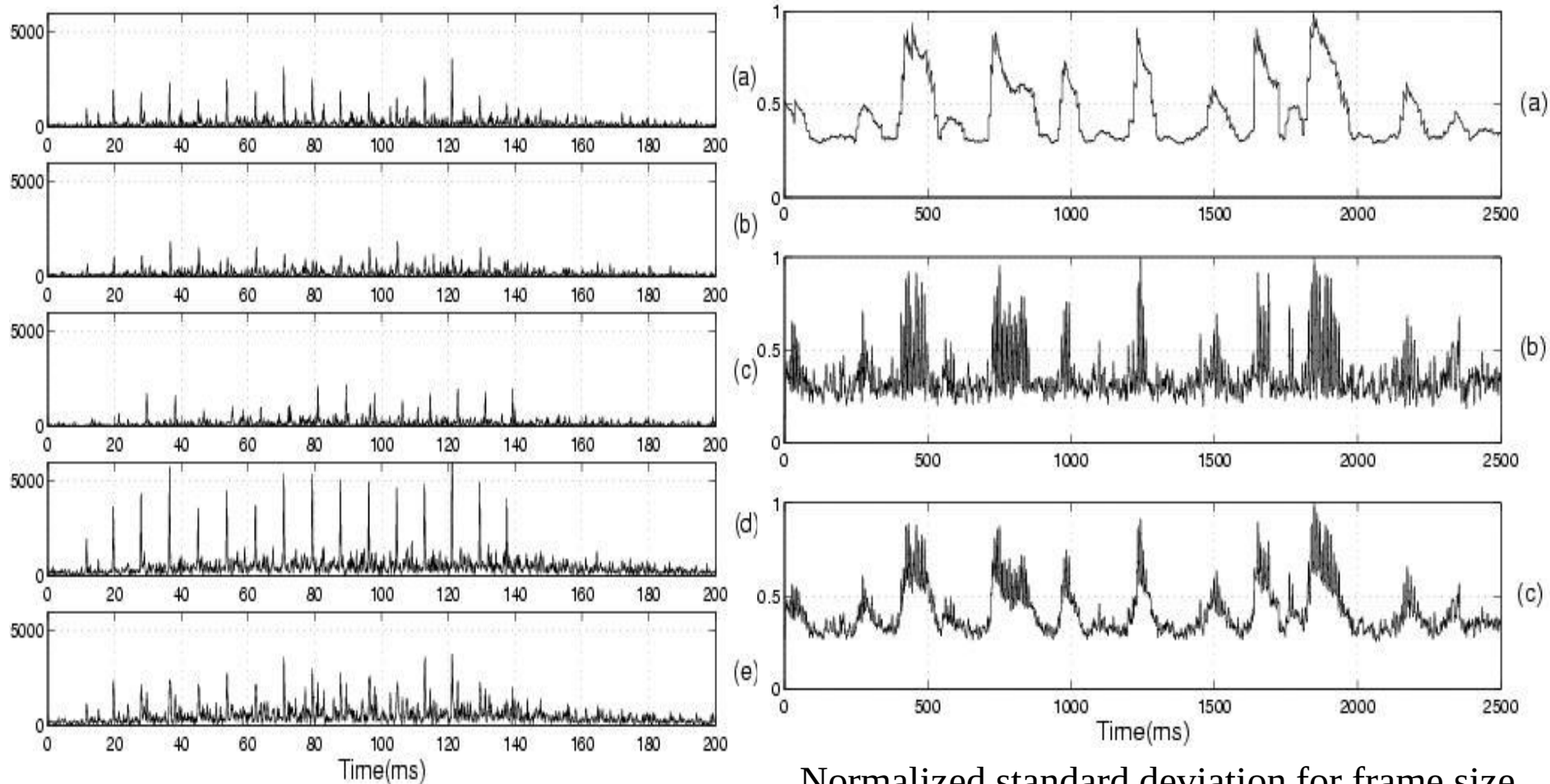
Cross-correlation of the Hilbert envelopes from the two microphones



(a) Normalized Hilbert envelope energy.  
Time-delays estimated from analysis frames  
of size (b) 50 ms, (c) 200 ms and (d) 500 ms,  
each with a shift of 10 ms



# Enhancement of Speech in Multichannel Case



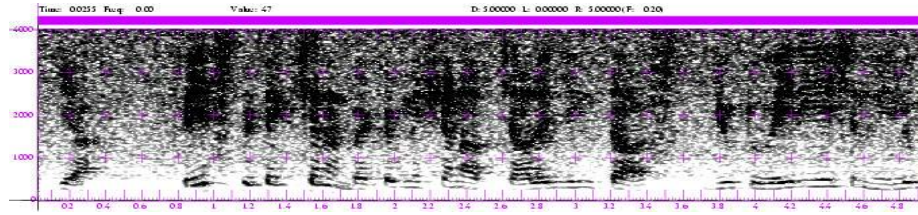
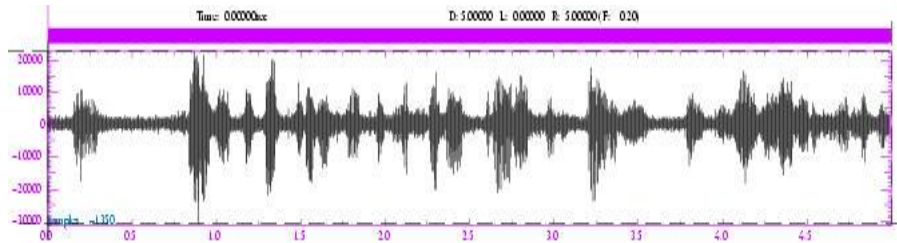
Hilbert envelope of the LP residual of a segment of (a) mic-1, (b) mic-2, (c) mic-3, (d) coherently-added and (e) incoherently-added

Normalized standard deviation for frame size of (a) 50 ms, (b) 3 ms, each with a shift of one sample and (c) Weight function obtained by adding (a) and (b)

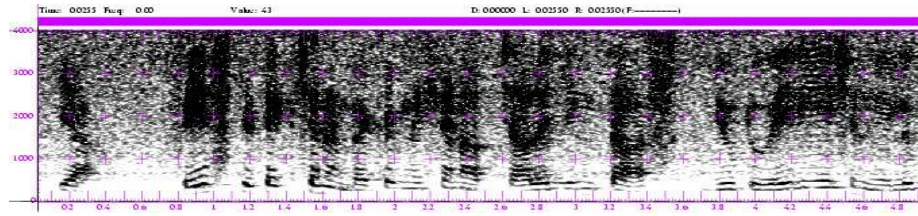
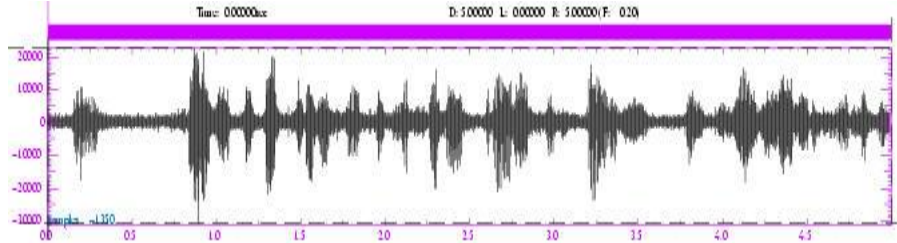


# Enhancement of Speech in Multichannel Case

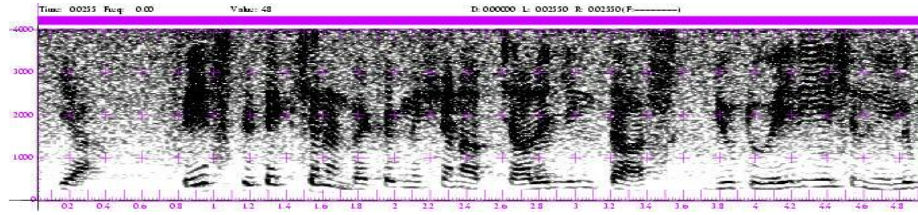
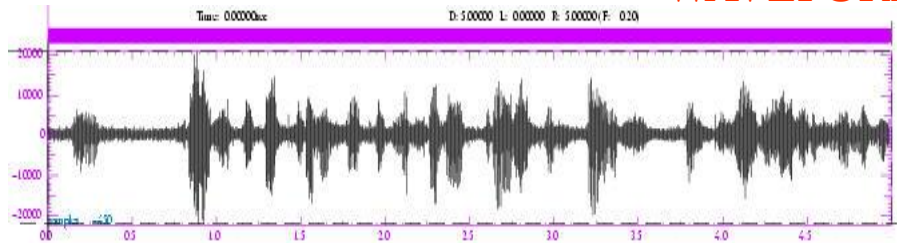
MICROPHONE-1



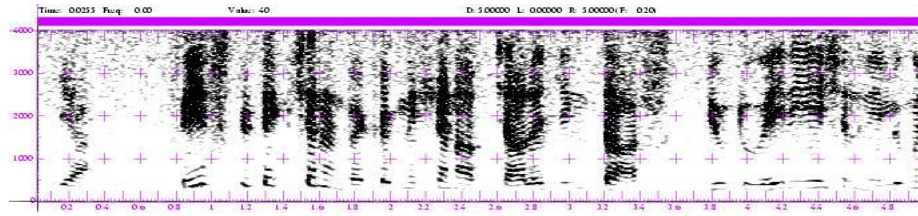
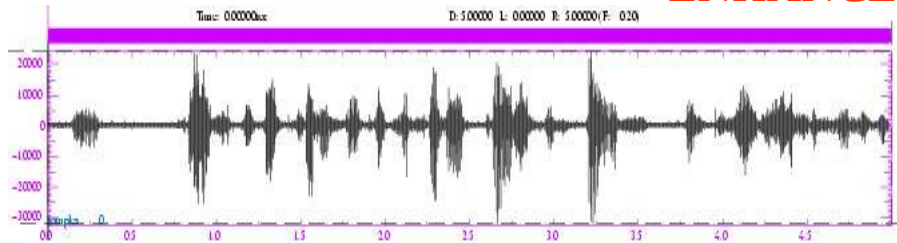
MICROPHONE-2



WAVEFORM ADDITION



ENHANCED SPEECH



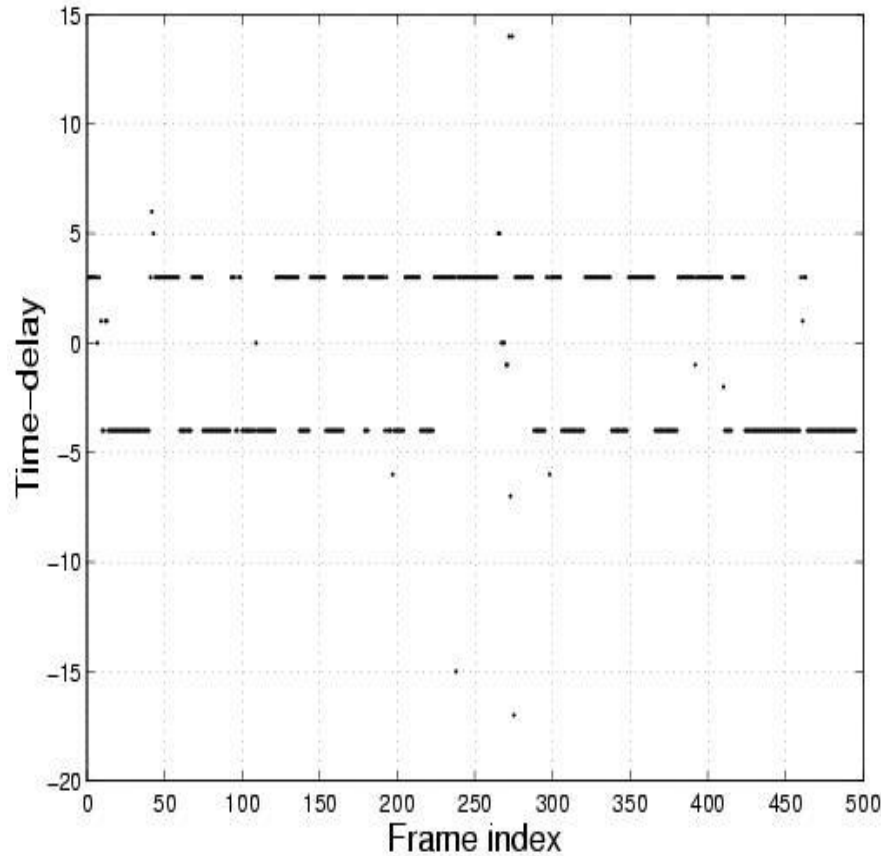
# Enhancement of Speech in Multispeaker Environment

**Objective:** To enhance speech of desired speaker

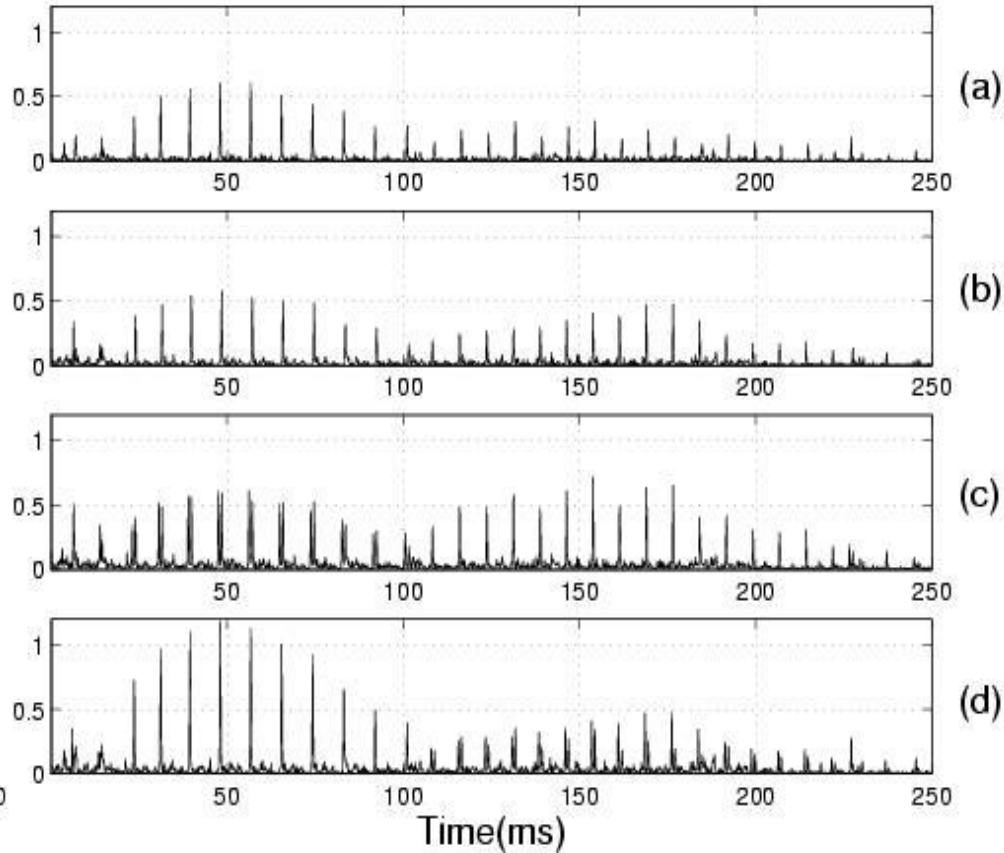
**Existing Methods:** Exclusively use knowledge of pitch to separate speech components of each speaker

**Limitation:** Estimating pitch in a multispeaker environment is a difficult task

# Enhancement of Speech in Multispeaker Environment



Time-delays computed using the Hilbert envelopes of the two microphones for every frame of 50 ms with a shift of 10 ms

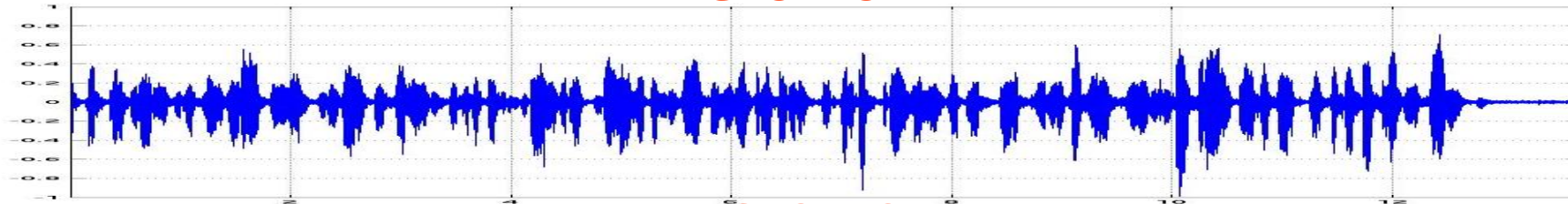


Hilbert envelope of the LP residual from (a) mic-1, (b) mic-2, (c) coherently-added using delay-1 and (d) coherently-added using delay-2

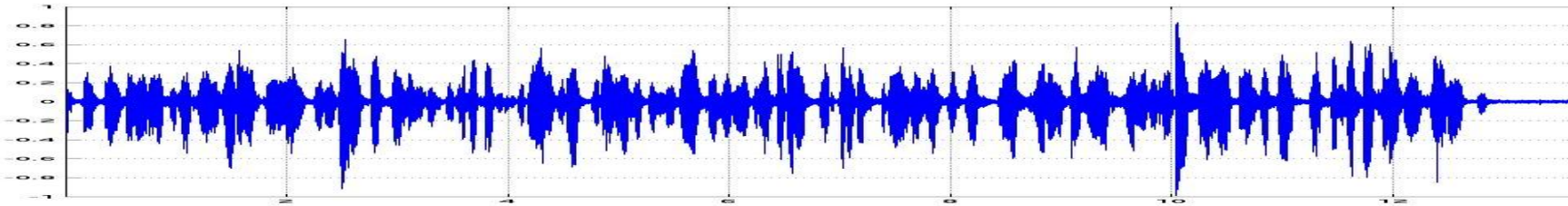


# Enhancement of Speech in Multispeaker Environment

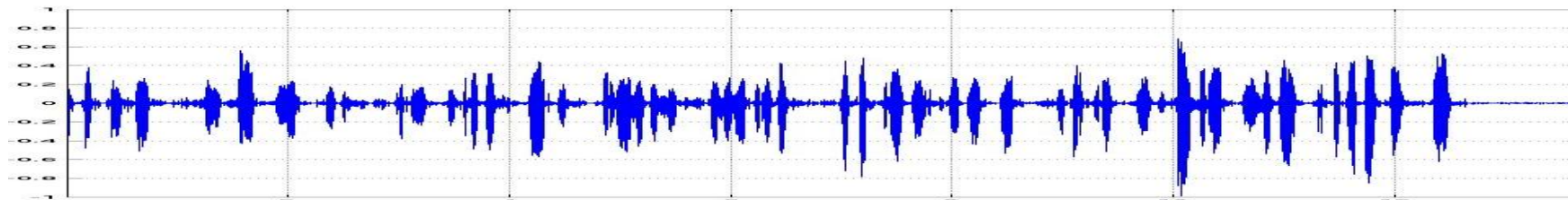
MICROPHONE-1



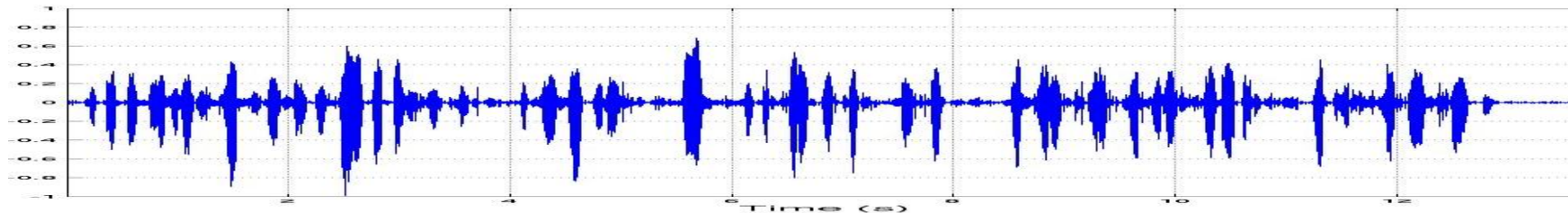
MICROPHONE-2



SPEAKER-1 ENHANCED



SPEAKER-2 ENHANCED



Thank You

