Gerhard Schmidt / Tim Haulick
Recent Tends for Improving Automotive Speech Enhancement Systems

Geneva, 5-7 March 2008
“Speech Communication Channels” in a Vehicle

- Into the vehicle
- Within the vehicle
- Out of the vehicle
Speech Enhancement for the “Communication Channels”

**Send side processing:**
- beamforming,
- echo and noise suppression

**Receive side processing:**
- gain control

**In-car communication:**
- feedback suppression,
- automatic gain adjustment
Often the car is owned by the driver or the passenger:

The communication channels “within the vehicle” and “into the vehicle” should be in focus for improving the perceived system quality.

Often a car is used only by a few (3 ... 5) people:

Speaker dependent speech processing might be a good choice.

Often the speech signals of passengers who are using the car can be recorded in periods with high SNR:

Speaker dependant speech models might be extracted in periods of high SNR and used in periods of low SNR.
**Extended Speech Enhancement**

**Send side processing:**
- beamforming, echo
- and noise suppression,
- speech reconstruction

**Receive side processing:**
- gain control,
- bandwidth extension,
- adaptive equalization

**Speaker independent speech knowledge**

**Speaker (in-)dependent speech knowledge**

**Phone**

**Speech dialog system**
Speech Enhancement in the Receiving Path (1)

Speaker independent speech knowledge

Receive side processing

Control for receive side processing

To loudspeaker

Phone (downlink)

Analysis filter bank

Echo cancellation

Analysis filter bank

Microphone

Synthesis filter bank

Residual echo and noise suppression

Noise PSD

Echo PSD

Phone (uplink)
Speech Enhancement in the Receiving Path (2)

Phone (downlink) → Packet loss correction → Noise suppression → Bandwidth extension → Adaptive equalization → Adaptive limiter → To loud-speaker

Control for receive side processing

Speaker independent speech knowledge

Noise PSD

Echo PSD
Bandwidth Extension – Basic Principle

Input signal → **Low-frequency extension** → **Lowpass filter** → **Highpass filter** → **High-frequency extension** → Output signal

*Speaker independent speech knowledge*
**Bandwidth Extension – Example**

**Narrowband connection (current standard):**

Bandwidth-extension for narrowband speech signals (bandwidth 3.4 …3.8 kHz) – extension of low frequency components and extension of high frequency components up to 5.5 or 8 kHz.

**Wideband connection:**

Bandwidth-extension for wideband speech signals (bandwidth 7 kHz, e.g. AMR wideband codec G.722.2) – extension of high frequency components up to 11kHz.
**Gain adjustment:**

The echo and the background noise power are analyzed and a gain correction is computed in order to achieve a predefined SNR in the passenger compartment.

**Shape adjustment:**

In addition to the short-term powers also the short-term spectra of the noise and the echo are analyzed and a correction filter is designed in order to boost frequencies with low SNR while slightly attenuating those with good SNR. The design process is computed 10 to 20 times per second. An improvement of the speech intelligibility can be achieved while maintaining the loudness of the output signal.
Adaptive Equalization – Example

Measurement:
Signal was recorded in an accelerating vehicle, while entering a motorway.
Motivation:

At medium and high speed the SNR often drops at low frequencies below the 0 dB threshold. Thus, standard noise suppression schemes perform only an attenuation at these frequencies.

For further improvement of the speech quality a reconstruction approach is an alternative.

However, speech reconstruction starts now where conventional noise reduction fails …
Speech Enhancement in the Sending Path (2)

Phone (downlink) → Receive side processing → Analysis filter bank → Analysis filter bank

Speaker (in-) dependent speech knowledge → Echo cancellation

Speech reconstruction

Mixer → Residual echo and noise suppression

Synthesis filter bank → Phone (uplink)

To loud-speaker → Microphone
Speech Reconstruction – Audio Examples

Time-frequency analysis of noisy speech before and after processing

- Microphone signal
- Conventional noise suppression
- Mixed suppression and reconstruction

Frequency in Hz
0 1000 2000 3000 4000 5000

Time in seconds
0 1 2 3 4 5 6 7 8 9
Additional Information:

Harman/Becker Automotive Systems
Acoustic Signal Processing

Gerhard Schmidt    Tim Haulick
geschmidt@harmanbecker.com    thaulick@harmanbecker.com

Söflinger Str. 100
89077 Ulm, Germany