Seat Belt-Microphone Systems and their Application to Speech Signal Enhancement

Mohamed Krini, Klaus Rodemer

paragon AG, Division Acoustics, Schwalbenweg 29, 33129 Delbrück, Germany

mohamed.krini/klaus.rodemer@paragon.ag

Introduction

In automobiles, microphones used for hands-free telephony or speech dialogue systems are often integrated in the rear view mirror or in the roof above the speaker. Alternatively microphones can also be positioned on a seat belt. By integrating the microphones into the seat belt of a passenger the distance from the talker's mouth to the microphones is being reduced. An enhanced signal quality in terms of high signal-to-noise ratio (SNR) can usually be reached. The seat belt-microphone system can be used for hands-free telephony or for in-car communication.

Seat Belt-Microphone System

The seat belt-microphones are manufactured by the company paragon AG [1]. In Fig. 1 an example of a belt-microphone system installed in a car is shown. It consists

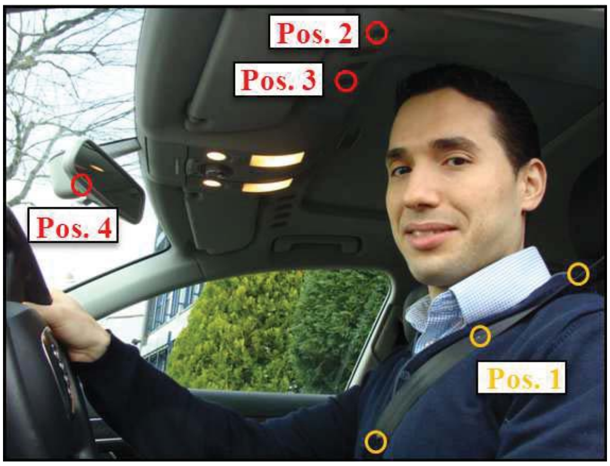

Figure 1: Belt-microphones (Pos. 1) and microphones positioned at the roof and at the mirror (Pos. $2-4$).

of three omnidirectional microphones spaced by 160 mm and are placed around the shoulder and chest of a sitting passenger. From the arrangement of the three microphones, that microphone is selected which is close to the speaker's mouth and exhibits the highest SNR. The backside of the belt has an even surface and all signal lines needed for signal transmission and voltage supply are integrated into the seat belt. A sophisticated mounting technique for the microphone capsules has been developed. The signal lines are made of a special alloy and a flexible structure of wires are weaved into the seat belts so that they appear invisible. It has been proven that the safety, usability and comfort of a seat belt with integrated microphones are still maintained. The microphones satisfy the VDA 1.5 specifications and receive signals between 100 Hz–6 kHz.

Microphone Selection

An overview about signal enhancement algorithms designed for belt-microphone systems was given in [2]. In this contribution we concentrate on selecting a best microphone from a belt system. The selection is based on the long-term SNR measure as well as on the time-delay between adjacent microphone signals. The number of microphones would be $M = 3$. To estimate the time-delay between signals received at adjacent microphones, the input signals $y_m(n)$ with $m \in [0, M-1]$ are firstly divided into overlapping blocks and subsequently windowed. After applying the DFT to the windowed blocks, the cross power spectral density (PSD) is estimated:

$$S_{y_i y_{i+1}}(\Omega_\mu, n) \;=\; Y_i\big(e^{j\Omega_\mu}, n\big)\, Y_{i+1}^{*}\big(e^{j\Omega_\mu}, n\big)\,, \qquad (1)$$

where $Y_i\big(e^{j\Omega_\mu}, n\big)$ denotes the short-term spectrum of the i-th microphone, with $i \in [0,\, M-2]$. The indices $\mu \in [0, N-1]$, N and n denote the subbands, the DFT order and the frame index respectively. The crosscorrelation function (CCF), $s_{y_i y_{i+1}}(k, n)$, is computed by applying the IDFT to the weighted cross PSD $S_{y_i y_{i+1}}(\Omega_\mu, n)$:

$$s_{y_i y_{i+1}}(k,n) = \frac{1}{N} \sum_{\mu=0}^{N-1} S_{y_i y_{i+1}}(\Omega_\mu, n)\, W\big(e^{j\Omega_\mu}, n\big)\, e^{j\frac{2\pi k}{N}\mu}. \qquad (2)$$

The time-delay can be determined by a search for the maximum of the CCF over the entire range of indices:

$$\widetilde{\tau}_{i,i+1}(n) \;=\; \mathrm{argmax}\Big\{ s_{y_i y_{i+1}}(k,n) \Big\}. \qquad (3)$$

The weighting function $W\big(e^{j\Omega_\mu}, n\big)$ has been used to attenuate highly disturbed speech components often concentrated at low frequencies in automotive environments. Furthermore, the reliability of the delay estimation is determined by the value of the normalized CCF:

$$p_{i,i+1}(n) \;=\; \frac{s_{y_i y_{i+1}}\big(\widetilde{\tau}_{i,i+1}(n), n\big)}{\sqrt{s_{y_i y_i}(0,n)\; s_{y_{i+1} y_{i+1}}(0,n)}}\,, \qquad (4)$$

where $s_{y_i y_i}(0,n)$ denotes the power of the input signal y_i. An estimation takes place only for large values of Eq. 4. For enhanced delay estimation, an IIR smoothing of first order in temporal direction is performed:

$$\overline{\tau_{i,i+1}}(n) \;=\; (1-\beta)\,\tau_{i,i+1}(n) \,+\, \beta\,\overline{\tau_{i,i+1}}(n-1)\,, \qquad (5)$$

where β falls in the interval $0.7 < \beta < 1$. The delay estimate $\widetilde{\tau}_{i,i+1}(n)$ is limited between $\tau_{\min} \simeq -460\ \mu s$ and $\tau_{\max} \simeq 460\ \mu s$ as follows:

$$\tau_{i,i+1}(n) \;=\; \min\Big(\tau_{\max}, \max\big(\widetilde{\tau}_{i,i+1}(n), \tau_{\min}\big)\Big)\,. \qquad (6)$$

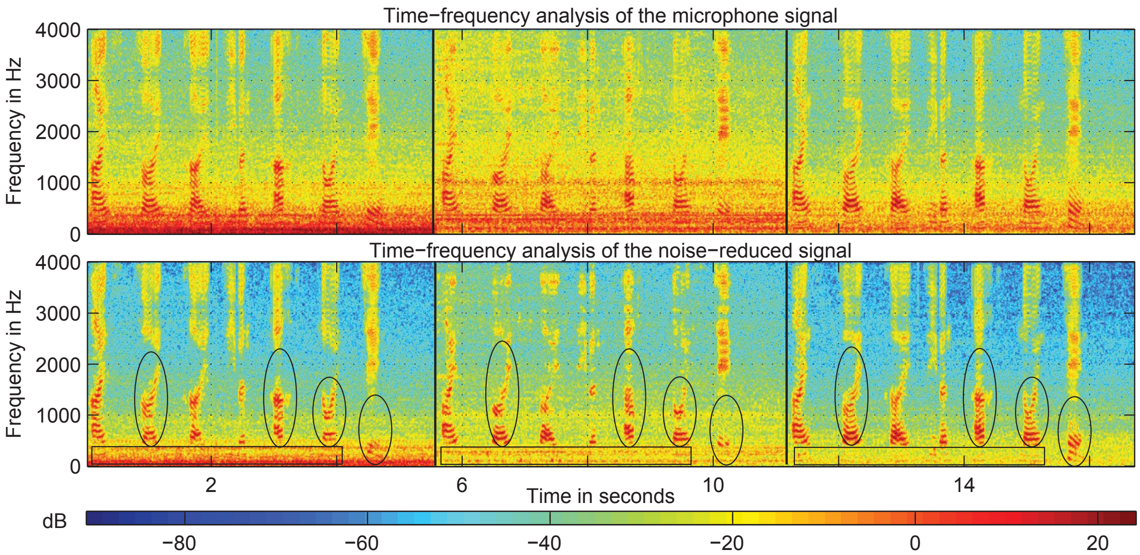

Figure 2: Spectrogams of the microphone signals (upper part) and of the noise reduced signals (lower part). The signals on the left, center and right correspond to microphone positions 2, 4 and 1 respectively.

The best microphone channel is chosen as follows:

$$m_{\mathrm{sel}}(n) = \tag{7}$$
$$\begin{cases} 0, & \text{if } \left(\tilde{\tau}^{(\min)} < \overline{\tau_{0,1}}(n) < \tilde{\tau}^{(\max)}\right) \wedge \left(p_{0,1}(n) > \tilde{p}\right) \\ & \quad \wedge \left(\overline{\xi_0}(n) = \overline{\xi_{\max}}(n)\right), \\ m+1, & \text{if } \left(\tau^{(\min)} < \overline{\tau_{m,m+1}}(n) < \tau^{(\max)}\right) \\ & \quad \wedge \left(p_{m,m+1}(n) > \tilde{p}\right) \wedge \left(\overline{\xi_m}(n) = \overline{\xi_{\max}}(n)\right), \\ m_{\mathrm{sel}}(n-1), & \text{else}, \end{cases}$$

where $\overline{\xi_{\max}}(n) = \max\{\overline{\xi_m}(n)\}$ indicates the highest long-term broadband SNR. The lower and higher bounds of $\overline{\tau_{m,m+1}}(n)$ can be set by $\tilde{\tau}^{(\min)} = \tau^{(\min)} = 0\ \mu s$ and $\tau^{(\max)} = -\tilde{\tau}^{(\max)} = 400\ \mu s$. The reliability threshold is set by $\tilde{p} = 0.6$. The SNR is only updated during local speech activity and while no activity on the reference channel and from the neighboring talkers is detected.

Simulation Examples

In Fig. 3 the belt-microphone system is compared with three hands-free microphones placed at different positions (see Fig. 1) in terms of the average SNR for driving speeds between 120 and 160 km/h. The distances from different microphone positions to the mouth are: $20-27$ cm (Pos. 1), 28 cm (Pos. 2/3), 58 cm (Pos. 4). All microphones are calibrated to have the same power at stand-still. The proposed method from the previous section has been used to select the best microphone. This comparison shows that at higher frequencies the behavior of all microphones is more or less similar. Whereas at low and medium frequencies, the belt-microphone outperforms conventional hands-free microphones. An improvement of up to $6-10$ dB in SNR can be achieved. The upper part of Fig 2 shows microphone signals recorded at a speed of 130 km/h for three microphone positions (left: Pos. 2, center: Pos. 4, right: Pos. 1). The lower part of Fig. 2 demonstrates the analysis after a conventional noise suppression using [3]. On comparing these results, one can see improvement in terms of noise reduction and speech recovery with a belt-microphone system compared to other microphones positioned at the roof or at the rear-view mirror, especially at low frequencies. Subjective tests have also indicated that the overall speech quality is considerably improved when using belt-microphones.

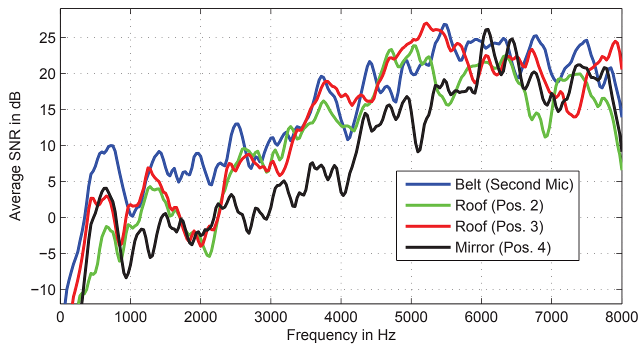

Figure 3: SNR measured at different microphone positions.

Conclusion

This contribution introduced the seat belt-microphone system and a method for selecting its best microphone. Using belt-microphone system for hands-free telephony, a considerable improvement in terms of speech quality and noise reduction can be achieved compared to other hands-free microphones positioned at different locations in the car. Overall, belt-microphones can be seen as an interesting alternative to conventional hands-free microphones. They are especially well-suited for cabriolets where a favorable position (like the roof) is not available. Besides hands-free telephony, belt-microphone systems can also be applied advantageously for in-car communication.

References

[1] paragon AG, URL: `http://www.paragon.ag/`

[2] V.K. Rajan, S. Rohde, G. Schmidt, J. Withopf, "Signal Processing for Microphone Arrays on Seat Belts," Workshop on DSP for In-Vehicle Systems, Seoul, Korea, 2013.

[3] Y. Ephraim and D. Malah, "Speech enhancement using a MMSE short-time spectral amplitude estimator," IEEE Trans. on ASSP, vol. 32, no. 6, pp. 1109–1121, 1984.