PASSENGER IN-CAR COMMUNICATION ENHANCEMENT

Klaus Linhard and Juergen Freudenberger

DaimlerChrysler AG, Research & Technology, Dialogue Systems
Wilhelm-Runge-Str. 11, 89081 Ulm, Germany
phone: +49 731 505 2170, fax: +49 1731 505 4105, email: klaus.linhard@daimlerchrysler.com

ABSTRACT

The expression passenger in-car communication refers to the conversation (speech communication) of the car passengers while driving. In particular, the communication between different seat rows is difficult due to road noise and the acoustic situation within a car. A system which enhances the speech intelligibility is desirable for safety and comfort reasons. We discuss the basic building blocks and design issues of such an in-car communication system, particularly with regard to the acoustic conditions in a car.

1. INTRODUCTION

Although luxury vehicles provide a noise reduced environment in the car the communication between the front seats and the rear seats is difficult. In contrast to a normal conversation, in cars we have road noise and the positions of the passenger seats are fixed which impairs speech understanding. Usually passengers do not feel comfortable to conduct long conversations. Frequently, the car driver is tempted to turn the head in order to improve the communication. Thus for safety and comfort reasons, a system which supports natural communication between passengers is desirable. In this paper we discuss an in-car communication system which improves speech communication in a vehicle. Such a system basically works as an intercom between the different passenger seats. Furthermore, it can serve as the acoustic front-end for other applications like hands-free telephony, voice controlled devices, broadcast services, and dialog systems. Similar concepts are considered for example in [1], [2], and [3]. The system under consideration consists of dedicated microphones and loudspeakers (one unit for each passenger), and a digital signal processing part. We discuss the basic design issues for such a system. For example, an in-car communication system operates in a closed electro-acoustic loop similar to public address systems which causes stability problems. In order to provide a full-duplex communication echo cancellation is required. Moreover, due to the short delays for wave propagation, the tolerable processing delays are also very small. These conditions are in fact similar to the situation with digital hearing aids [4]. In section 2, we discuss the influence of the microphone/loudspeaker positioning on the achievable gains in terms of signal to noise ratios. In the subsequent section, we describe the basic signal processing components required for in-car communication. That is, we consider the necessary efforts in order to stabilize the system as well as for echo cancellation. The discussions are supported by acoustic measurement results obtained in a Mercedes S-Class. Some speech samples will be demonstrated at the conference.

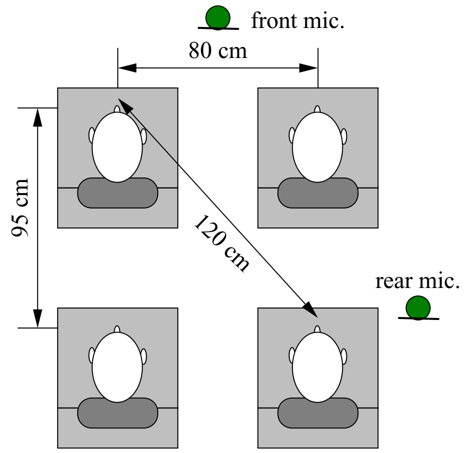

Fig. 1: Geometric situation for four passengers inside a Mercedes S-Class.

2. MICROPHONE/LOUDSPEAKER POSITIONING

In this section, we present measurement results for speech signal propagation within a car. In particular, we consider the transfer characteristics of speech signal propagation between talker and listener depending on the speaker position as well as on the positioning of the microphones and loudspeakers. The corresponding results are helpful for estimating the potential gains of in-car communication systems.

2.1 Transfer functions depending on the microphone position

Fig. 1 is an illustration of the geometric situation for four passengers inside a car. We give the approximate distances to each other in cm and depict two possible positions for microphones in the front and in the rear of the car cabin.

The best solution to capture the speech source would be with a close talk microphone for each passenger (not shown in Fig. 1). Such a close talk microphone is not suitable for practical applications. However, we will use close talk signals as a reference in order to discuss potential gains for in-car communication. That is, in the following measurements we refer to the signal spectrum of a close talk microphone such that the corresponding spectrum is white with a signal level of 0 dB. In Fig. 2 we show measurement results for two speaking situations: driver is speaking and right hand side passenger in the rear is speaking. Because levels at left and right ears are different we give the maximum, respectively. These measures were conducted with frequency linear omnidirectional microphones and the frequency averaging is 1/3 octave. We observe from Fig. 2 that the received levels at the

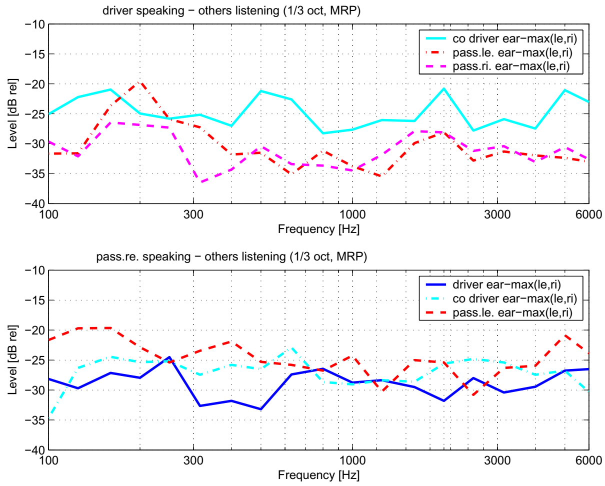

Fig. 2: Transfer functions for two different speaking positions with respect to close talk microphone (mouth reference point).

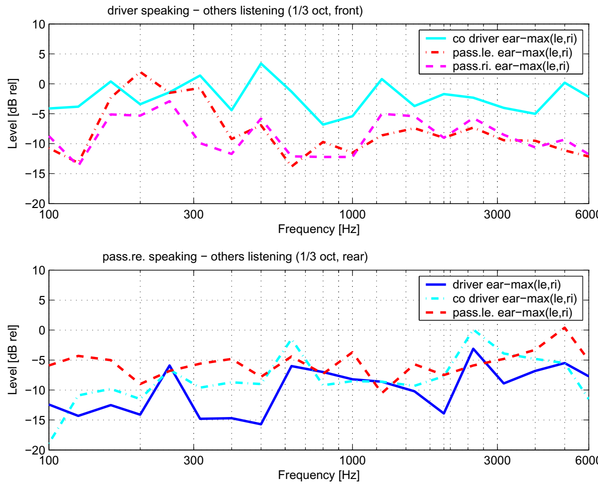

Fig. 3: Transfer functions for two different speaking positions with respect to front or rear microphone.

ear positions of the non-speaking passengers are about 20 to 30 dB smaller than the level of the close talk microphone. Furthermore, the lowest signal level reduction is obtained for the passenger closest to the speaker. Assuming that on all ear and microphone positions we observe the same level of road noise we can compare the different received speech levels. For example, if a speech level at the close-talk microphone position is 3 dB higher than at an ear position, using the microphone (and for example in-ear monitoring) would result in a 3 dB better quality (3 dB gain in signal to noise ratio). Hence, from Fig. 2, we would conclude a basic potential for an enhancement in the order of 20 dB.

However, a situation with close talk microphones is not suited for in-car applications. Hence, we also consider a more realistic scenario with positions suited for car integrated microphones. Instead of referring to the close talk microphone we refer to the front and rear microphone position (cf. Fig. 1), respectively. From the corresponding transfer functions depicted in Fig. 3 we obtain a more realistic estimate for the potential gain of an in-car communication system, i.e., an achievable improvement in signal to noise ratio in the order of 10 dB. Moreover, the front microphone gives no advantage for the communication support between the driver and co-driver. The microphone speech level (front mic) is almost the same as the ear level.

Apparently, these measurements neglect potential gains of directional microphones, microphone arrays, or noise cancellation. However, the effectiveness of noise cancellation algorithms is limited by the very restrictive demands on the tolerable processing delay.

2.2 Loudspeaker positioning

Besides the microphones another important aspect is the positioning of loudspeakers for sound reinforcement. The reduction of costs would directly lead us to use the already installed speakers for the car radio. The disadvantage of this solution is that these speakers are often far way from the passengers head, resulting in reduced system performance. In order to examine the acoustic situation for the reinforced signal, we placed a small loudspeaker box on the dashboard be-

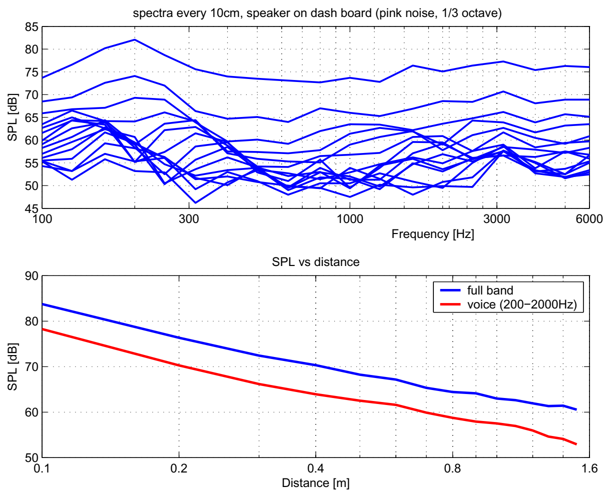

Fig. 4: Spectrum and signal level depending on the distance to the loudspeaker.

tween the head-rest for the rear passengers. We measured the spectrum 10 cm in front of this speaker and increased the distance in increments of 10 cm up to 150 cm toward the front of the car. The 150 cm point is close at the interior rear mirror. Calculating the total power reveals that each doubling of the distance yields a power decrease of 6 dB. This is also true if we only refer to the frequency range 200 Hz to 2 kHz, were most of the speech power is present. The corresponding curves are given in Fig. 4.

These results indicate that inside a car we are approximately in a free sound field in contrast to a diffuse sound field (dominated from reflection, sound level constant over distance). This 6 dB rule of thumb supports the argument to place the loudspeaker for the speech reinforcement close to the passengers head. For measurements we used loudspeakers mounted directly above the passengers. If, for example, the loudspeaker is in 20 cm distance to the ear, a low signal power is sufficient to provide a moderate listening level and the acoustic dissipation into the car cabin is also reduced. If

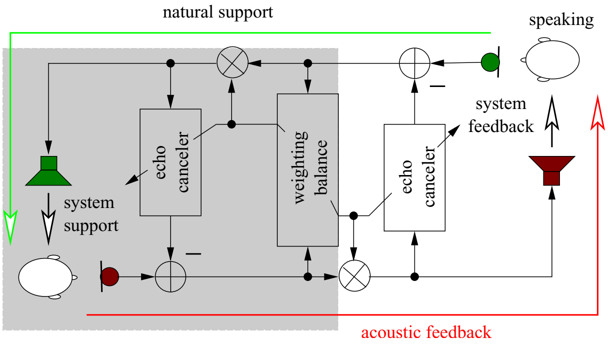

Fig. 5: Basic system structure with only two microphones and loudspeakers.

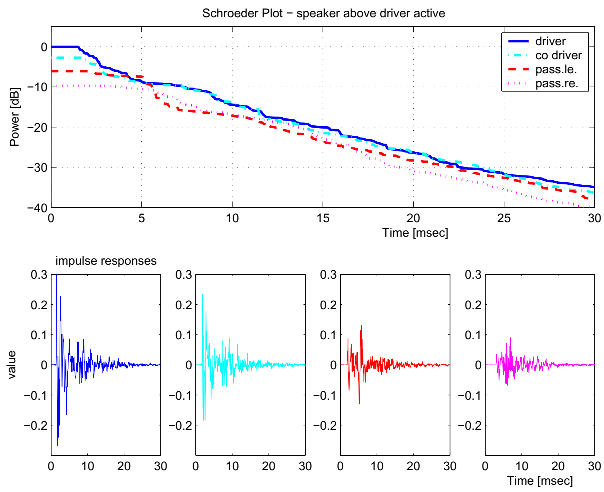

Fig. 6: Schroeder plots and impulse responses for different propagation paths (impulse responses from left to right: driver, co-driver, rear left, and rear right).

we assume the active speech source, or more precisely the microphone for this source, to be 80 cm away the receiving speech at the sending microphone (cross talk) is 12 dB lower than the level at the ear. This 12 dB gives a good basis for stabilizing the feedback problem (cf. Section 3).

3. BASIC SYSTEM STRUCTURE

Analyzing the structure of passenger in-car enhancement systems we may start with a common hands-free unit (see light grey box, Fig. 5). The hands free unit consists of an echo canceler and a weighting balance. The task of the echo canceler is to compensate the undesired microphone signal received from the loudspeaker near by (called echo). The weighting balance reduces residual echoes and thus improves the total system performance (see e.g. [5]). If the system is switched off the speaking person will hear his/her own voice as an echo.

3.1 Echo cancellation

The system shown in Fig. 5 has two modifications compared with a common hands free unit. Firstly, there is a second echo canceler on the right hand side. This is straight forward and is in fact similar to the situation when there is a hands free unit at each terminal side. Secondly, we used the control output from the weighting balance to modify the convergence step size of the echo cancellation. As a consequence only the echo canceler where the loudspeaker is active is allowed to adapt. This is useful to avoid misadjustment of the estimated impulse response. The issue of step-size control for the echo cancellation is of particular concern, because of the noisy environment inside the car [6]. Furthermore, the speech excitation signal is highly correlated which leads to slow convergence of the adaptation. A decorrelation filter (linear predictor) may be inserted in the excitation path of the filter adaptation so as to reduce the problem of slow convergence.

In order to estimate the necessary effort for a signal processing system we have measured the impulse responses of the involved transfer paths and derived the corresponding Schroeder plots (backward integrated energy of the impulse response). The Schroeder plots given in Fig. 6 show the decrease of the reverberation energy. Reverberation time usually is defined as the time for the energy to drop down by 60 dB. If we extrapolate this from our plots we get approximate 50 msec. Comparing this result with, e.g., the result for a recording studio with about 300 msec reverberation time we notice that the in-car installation has very small reverberation. Even more important, for modeling the transfer path we only need to consider impulse responses of about 15 msec. With this short impulse responses we can model the first 20 dB (see Schroeder plot in Fig. 6). This 15 msec correspond to a filter length of 150 samples with 10 kHz sampling frequency.

3.2 Feedback problem

The car environment is different from the usual telecommunication application, because there is an additional acoustic support path. If the system support from the loudspeaker is active this acoustic path also works as an acoustic feedback path back to the speaker. Due to echo cancellation the system feedback may be negligible, but the electro-acoustic feedback still exists and may cause stability problems if the turn around gain is higher than one for one or more frequencies. Hence, this electro-acoustic feedback limits the maximum gain that can be achieved without additional extensions to avoid feedback. The feedback problem itself is not new and is similar to the feedback problem in public addressed systems which frequently suffer from suddenly appearing feedback ringing or howling. The conventional method to solve this problem is manually setting of damping factors to potential feedback frequencies with the aid of a graphic equalizer. Automatic operation is possible with devices called feedback destroyer or feedback eliminator. Such devices can automatically detect feedback frequencies and automatically set very narrow notch filters with a resolution of the order of one Hz. Other applications with feedback problems are for example hearing aids. Due to the imperfect fitting in the ear or venting in the hearing device itself, there is an acoustic leakage from the receiver (loudspeaker) to the microphone [4].

For in-car communication, it is also possible to employ one of the following two approaches for adaptive feedback cancellation. A first approach exploits the fact that we are only interested in speech communication. An adaptive notch filter realizes a linear predictor with a prediction delay D (cf. Fig. 7). D should be higher than the correlation length of

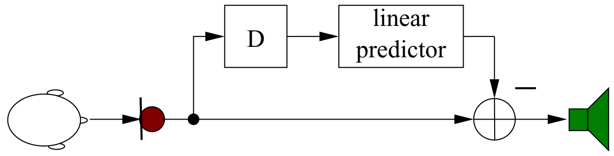

Fig. 7: Feedback cancellation based on linear predication.

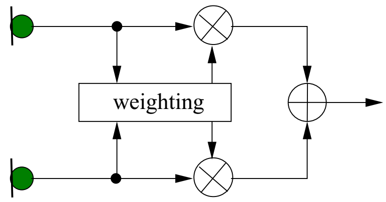

Fig. 8: Microphone weighting balance.

speech, e.g., if the sampling frequency is about 10 kHz a reasonable value for D is 10. Speech components will be hardly effected from the predictor, because after 10 samples there is only little correlation. Sinusoidal components from feedback are still correlated after 10 samples and therefore will be predicted and finally subtracted.

A second approach is to use an additional echo canceler and utilizes the fact that the impulse responses within the car are relatively short. The filter of this canceler should model the natural feedback path from the loudspeaker back to the microphone. The problem with this approach is the adaptation of the filter coefficients. Even if there would be no acoustic feedback path, such a feedback compensation filter would nevertheless try to compensate some components of the speech.

3.3 Expanding the system for four passengers

Up to now he have discussed the situation for two passengers, one in the front and for another position in the rear of the car. According to the acoustic measures inside the car, supporting the direction to/from the left or right neighbor may be dispensable. Thus we may add the two microphones for the front speakers and also the two microphones for the rear speakers. The loudspeakers in the front of the car and the rear may work in parallel, respectively. Instead of just adding the two microphone signals it may, however, be advantageous to weight the signals before adding in order to emphasize the louder microphone. This will reduce noise and reverberation. The basic principle of such a microphone weighting balance is illustrated in Fig. 8.

4. CONCLUSIONS

In this work we have presented measurement results for the acoustic situation in-side a car. Furthermore, we have discussed these results with regard to an in-car communication system. For the sake of conciseness we have restricted the measurements to omni-directional microphones. Hence, our results neglect potential gains of directional microphones or microphone arrays.

As we have demonstrated, the transfer characteristic of speech signal propagation depends on the speaker position as well as on the positioning of the microphones and loudspeakers. With microphone positions suitable for in-car installation an enhancement of the signal to noise ratio for the propagation path from front to rear (and vice versa) in the order of 10 dB seems achievable. While the system support for positions in the same seat row may be dispensable. Signal processing is required for echo cancellation and to stabilize the system.

REFERENCES

[1] B. M. Finn, “Integrated vehicle voice enhancement system and hands-free cellular telephone system,” *European Patent EP 0 932 142 A2*, Jul. 1999.

[2] K. Schaaf, J. Schultz, and K. Tontch, “Digital voice enhancement for improved in-car communication,” in *Proc. 3rd IFAC Workshop Advances in Automotive Control, Karlsruhe, Germany*, March 2001.

[3] T. Haulick and G. Schmidt, “Signal processing and performance evaluation for in-car communication systems,” in *Proc. ITU-T Workshop on Standardisation in Telecommunication for motor vehicles, ITU Headquarters*, Nov. 2003.

[4] J. M. Kates, *Signal Processing for hearing aids*, Kluwer Academic Publishers, 1998, Chapter 6 in M. Kahrs, K. Brandenburg: Applications of Digital Signal Processing to Audio and Acoustics.

[5] T. Burger and U. Schultheiss, “A robust acoustic echo canceller for a hands-free voice-controlled telecommunication terminal,” in *Proc. EUROSPEECH 1993, Berlin, Germany*, 1993, vol. 3, pp. 1809–1812.

[6] A. Mader, H. Puder, and G. Schmidt, “Step-size control for acoustic echo cancellation filters, an overview,” *Signal Processing*, vol. 80, no. 9, pp. 1697–1719, Sep. 2000.