Automatic Evaluation of In-Car Communication Systems

Anne Theiß[1], Gerhard Schmidt[1]

[1] *Digital Signal Processing and System Theory, Christian-Albrechts-Universität zu Kiel, E-mail: {ath, gus}@tf.uni-kiel.de*

Abstract

In order to improve the impaired communication among the passengers of a driving vehicle, so-called in-car communication (ICC) systems recently entered the automotive market. Such systems record, process, and play back the desired speech signal of the talking passenger in order to increase the desired signal strength inside the vehicle. This contribution proposes an automatic evaluation strategy which is capable of testing an arbitrary ICC system by means of a multitude of single measures. Those single measures allow for a detailed analysis and understanding of the quality of ICC systems and will be further used to derive overall quality results. Those results are presented in a graphic scheme which eases the access to the found results.

Introduction

The communication between the passengers of a driving vehicle can be impaired, especially at medium or high velocities. In order to improve this situation so-called in-car communication (ICC) systems can be utilized. An ICC system records the speech signals of the talking passengers and plays back the enhanced speech signals via those loudspeakers positioned close to the listening passengers. A basic ICC system with only one loudspeaker and microphone is depicted in Fig. 1. Certainly, this system does not correspond to a complete ICC system as used in actual vehicles, but it conveys the basic functionality of such a system. Common ICC systems utilize several signal processing stages, such as noise and feedback suppressions or automatic gain control, in order to create an enhanced speech signal. Detailed descriptions of the functionality of ICC system can be found in [1, 2, 3].

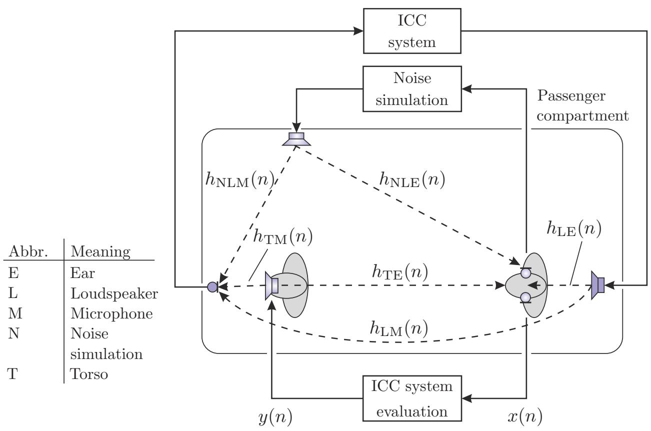

Figure 1: Basic structure of an ICC system and the evaluation environment including a necessary noise scenario inside a vehicle [4].

Whenever a new system is developed or when an existing one is extended, the question about the quality of the system is raised. Within this contribution, an automatic evaluation strategy is proposed, which is capable of testing arbitrary ICC systems. The strategy is divided into three different evaluation groups:

- The first deals with the classification of the environment inside the vehicle and the vehicle itself.
- In the second evaluation group, the quality of a ICC system is determined by analyzing the system behavior.
- The third group derives the quality of the communication situation inside the vehicle.

Within each group, a variety of different automated measures are derived, which allow a detailed root-cause analysis in case of non-optimal system. The single results are combined to an overall quality outcome by means of a sophisticated metric. In order to ease the access to the obtained results an graphic scheme is proposed.

Evaluation Strategy

In any case, to predict the quality by means of instrumental measures, a quality base needs to identified. The quality of an arbitrary speech processing system is given by the subjective judgment of human beings. Therefore, the quality base needs to be derived from the evaluation of a variety of different test persons experiencing the processed signals within the target environment. The judgment of each test person is depending on the experience, the expectations, and the physical characteristics of the test person as well as the environment. In order to reduce the variation of the subjectively determined quality a variety of different human beings need to be included to the test group. By means of the achieved quality results the mean opinion score (MOS) is derived and, further, utilized as a quality base. During the development of a new instrumental measure, predicting the quality of an system, the quality base is split up into two sets. One is needed to develop the instrumental measure, the other set is used for validation purposes. This approach, as it is described for example in [5, 6] for the speech quality of telephony, allows the creation of a reliable quality measure.

Besides the quality base, which is a mandatory part in the development process of an instrumental measure, the evaluation environment needs to be specified and designed in such a way that different ICC systems can be evaluated. In Fig. 1 the passenger compartment of a vehicle equipped with an ICC system is depicted. In order to create a reproducible noise scenario an independent noise simulation system is necessary. In [7] such a mobile noise simulation within a vehicle is presented. In

Fig. 2) the arrangement of the noise simulation is exemplary depicted. This particular noise simulation is independent since it can be utilized in an arbitrary vehicle and, hence, an arbitrary ICC system can be excited with a predefined noise scenario. Within such a scenario, for example a drive with a velocity of 150 km/h, the ICC system evaluation module excites the ICC system with a certain test signal via an artificial talking passenger on the driver seat. In addition, the response of the overall system is observed by microphones located at the ears of the listening passenger and processed by the system evaluation. Such a setup allows to utilize not only the signals processed by the ICC system, but also the test signal, so-called reference-based measures can be applied in order to measure the quality.

Figure 2: Noise simulation in a laboratory environment.

Nevertheless, the terms *quality* respectively *speech quality* are complex concepts [8, 9], which can not be captured by only one instrumental measure. A sophisticated evaluation strategy consisting of a variety of different instrumental measures is necessary in order to cover all influences on the quality and to divide the main objective into several smaller tasks.

Evaluation Groups

The complex situation of evaluating the quality of an ICC system is split up in three evaluation groups:

- **Vehicle characteristics**

  The measures of this group capture the characteristics of the given environment inside the vehicle. Measures that determine the loudness of the background noise while driving at a high velocity or the distances between listening and talking passengers as well as between the passengers and the transducers (microphones and loudspeakers) of the ICC system, are part of this group. The individual results have only an indirect influence on the overall quality evaluation of the ICC system. The findings are utilized to reflect and include the complexity of the given environment into the evaluation. Thus, a system which provides adequate results, by means of other quality measures, in a difficult environment, is judged as a better system than a system which achieves the same results in a less complex environment.

- **ICC system behavior**

  This second evaluation group combines instrumental measures, which derive the behavior of the ICC system under test. The idea is to find out, if the system itself operates as expected. If the system does, for example, increases the noise level inside the compartment, the overall quality needs to be reduced. But this ICC system behavior is linked to the given environment, since, an increase of the background noise of 3 dB within a loud situation is less crucial than in a quiet environment. Therefore, the results of this evaluation group are weighted with the vehicle attributes as described, in more detail, in the next section.

- **Communication quality**

  The last evaluation group should analyse the situation inside a vehicle, which is equipped with an ICC system, in terms of communication quality. Common measures such as the improvement of the signal-to-noise ratio (SNR) at the ear of the listening passenger as well as the improvement of the speech transmission index (STI) could be examined within this group. But also linguistic measures as the number of turns or hesitations of the speaking passenger are possible communication quality measures [10]. The main idea is to consider not only the quality in terms of improvement for the listening passenger but also for the talking passenger.

The results of the evaluation group *ICC system behavior* and *communication quality* are further mapped into an absolute category rating scale consisting of four different categories. The four levels are chosen according to [11], were 1 equals to an exceptional good system and 4 to a poor system.

In a next step all derived results need to be combined into an overall quality result in order to determine whether the system under test corresponds to an excellent or poor ICC system.

Overall Quality Evaluation

As mentioned before, the results of the evaluation group *vehicle characteristics* are utilized to specify the given environment. Therefore, the distance between the talking and listening passenger or between the transducers and the passengers can be analyzed, as well as the loudness of the background noise occurring at high velocities. By means of the derived results the attributes *vehicle size* g_{S}, *background noise* g_{N} and *transducers of the ICC system* g_{T} can be defined. Of course further measures as well as further attributes are possible to be included in this section.

The determined attributes are further utilized to weight the results from the evaluation group *ICC system behavior*, see Fig. 3. The attributes are assigned to those evaluation measures that are influenced by the considered vehicle characteristic. The results from the evaluation group *ICC system behavior* are combined to the following result vector:

$$\boldsymbol{c}_{\mathrm{I}} = \left[c_1^{\mathrm{I}},\, c_2^{\mathrm{I}}, ..., \, c_N^{\mathrm{I}}\right]^{\mathrm{T}}, \tag{1}$$

where N corresponds to number of instrumental measures considered within this evaluation group. According to the composition of this vector the corresponding weighting vector needs to be defined by

$$\boldsymbol{w}_{\mathrm{I}} = \min\left\{\max\left\{0.1, \left[w_1(\boldsymbol{g}),\, w_2(\boldsymbol{g}), ..., \, w_N(\boldsymbol{g})\right]^{\mathrm{T}}\right\}, 1\right\}. \tag{2}$$

where

$$\boldsymbol{g} = \left[g_{\mathrm{N}},\, g_{\mathrm{T}},\, g_{\mathrm{S}}\right]^{\mathrm{T}}. \tag{3}$$

For example, if the insertion of an residual noise by the ICC system is evaluated, the result is weighted with the attribute *background noise* g_{N} since an increased background noise increases also the impact for the ICC system algorithms.

The result of the evaluation group *ICC system behavior* is derived by:

$$Q_{\mathrm{I}} = \frac{c_{\mathrm{I}}^{\mathrm{T}} \cdot \boldsymbol{w}_{\mathrm{I}}}{\|\boldsymbol{w}_{\mathrm{I}}\|_1}, \tag{4}$$

with $\|...\|_1$ being the L_1 norm of a vector. In order to not exclude one result of vector $\boldsymbol{c}_{\mathrm{I}}$ the weighting vector $\boldsymbol{w}_{\mathrm{I}}$ is limited to 0.1. The combination of the results of the evaluation group *vehicle characteristics* and *ICC system behavior* is depicted in the upper leg of Fig. 3.

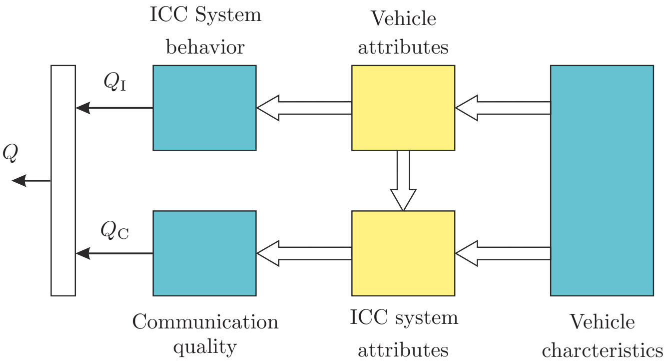

Figure 3: Derivation of the overall quality result of the instrumental evaluation according to [9].

In addition, the vehicle attributes and also further measures of the *vehicle characteristics* are combined and result in the *ICC system attributes*. In this category attributes like the necessity g_{R} of an ICC system is included. Those attributes contribute to the result of each single evaluation measure with in the evaluation group *communication quality*. In general the measure, like the SNR at the ear of the listening passenger, is evaluated during a scenario without the support of an ICC system $c_{\mathrm{E}}^{(\mathrm{off})}$ and with an activated ICC system $c_{\mathrm{E}}^{(\mathrm{on})}$. In order to achieve reasonable results, it is important to excite an almost equal noise scenario for both cases and to consider an appropriate calibration time of the ICC system. The result of one measure in this evaluation group can be defined by the following general relation:

$$c_{\mathrm{E}} = \min\left\{\max\left\{1, \hat{c}_{\mathrm{E}}^{(\mathrm{on})} + g_{\mathrm{R}} \cdot \left(\hat{c}_{\mathrm{E}}^{(\mathrm{on})} - \hat{c}_{\mathrm{E}}^{(\mathrm{off})}\right)\right\}, 4\right\}. \tag{5}$$

The results of all instrumental measures are combined in the result vector of the evaluation group *communication quality*

$$\boldsymbol{c}_{\mathrm{C}} = \left[c_1^{\mathrm{C}},\, c_2^{\mathrm{C}}, ..., \, c_{\mathrm{M}}^{\mathrm{C}}\right]^{\mathrm{T}}, \tag{6}$$

where M correspond to the number of instrumental measure within this evaluation group. The single elements of vector $\boldsymbol{c}_{\mathrm{C}}$ are derived by different quality evaluation measures and by means of Eq. (5). Since the weighting of the single measures by means of the system attribute necessity g_{R} is already performed, the mean value of all results within this group determines the result of the evaluation group. It follows, that the result of the evaluation group *communication quality* can be determined as follows:

$$Q_{\mathrm{C}} = \frac{\boldsymbol{c}_{\mathrm{C}}^{\mathrm{T}} \cdot \boldsymbol{w}_{\mathrm{C}}}{\|\boldsymbol{w}_{\mathrm{C}}\|_1}, \tag{7}$$

where the weighting vector is given by

$$\boldsymbol{w}_{\mathrm{C}} = [1,\, 1, ..., \, 1]^{\mathrm{T}}. \tag{8}$$

The weighting vector $\boldsymbol{w}_{\mathrm{C}}$ has the same length M as the result vector $\boldsymbol{c}_{\mathrm{C}}$.

In a final step the overall quality measure is derived by

$$Q = \frac{Q_{\mathrm{I}} + Q_{\mathrm{C}}}{2}. \tag{9}$$

The combination and influence of the single attributes and results of the instrumental measures is also depicted in Fig. 3.

Result Diagram

In order to comprehend the resulting values and to ease the root cause analysis for an impaired system, all made results as well as attributes are concentrated within a graphical representation, see Fig. 4. In this particular case, an evaluation strategy with $N = 3$ instrumental measure in evaluation group *ICC system behavior* and with $M = 3$ measures utilized in the evaluation group *communication quality*. Also the already mentioned three vehicle attributes are derived by means of the results from the evaluation group *vehicle characteristics*.

Fig. 4 depicts this graphically the evaluation results of one vehicle respectively one ICC system. Similar to diagrams described in [12, 9], the colors indicate whether the result is identified as an exceptional good (green) or a poor (red) result. Also the filling of the color is associated with the achieved quality.

It can be seen that the results show an ICC system, which has reached a good overall quality result. The result of the evaluation group *communication quality* also is rated as a good, since, the semicircle is almost filled in green. Similar results are also achieved by two of the three single measures of this group. The third measure obtained a lower evaluation result in comparison. The group *ICC system behavior* was rated, in total, as good. This is caused by the three different evaluation results of the single measures, which ranged from excellent to fair. In conclusion, the shown evaluation results corresponds to a good system, which showed some weakness in the results of the evaluation group *ICC system behavior*.

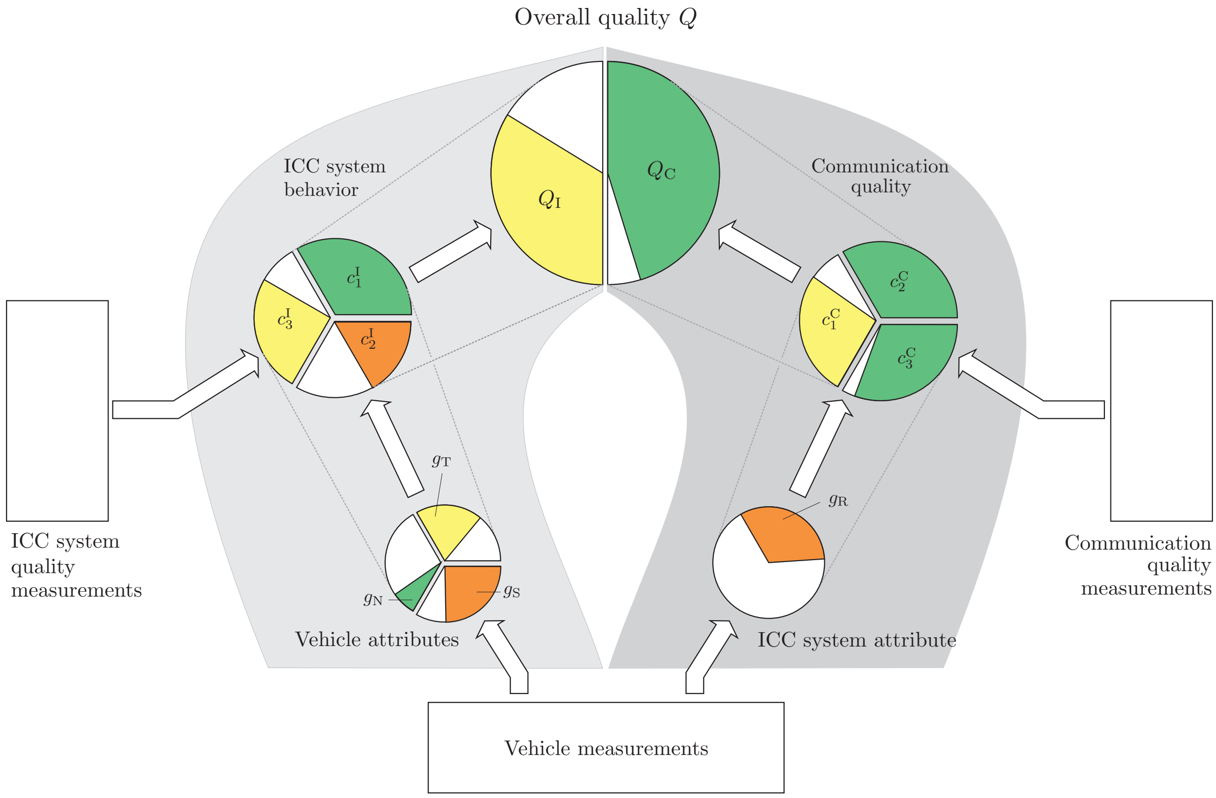

Figure 4: Graphical representation of the overall evaluation result of an arbitrary ICC system. In detail, some individual results of possible evaluation measures within each evaluation group, the attributes of the vehicle and the ICC system, as well as the corresponding measurement blocks are depicted [9].

Conclusion and Outlook

In conclusion, a comprehensive evaluation strategy has been introduced, which includes a variety of possible extensions either with new evaluation methods or with sophisticated weighting functions. Meaning that besides further measures within each of the three evaluation groups, also the weighting vectors $\boldsymbol{w}_{\mathrm{C}}$ and $\boldsymbol{w}_{\mathrm{I}}$ could be extended.

In any case, further subjective tests need to be performed in order to increase the quality base of such evaluation strategies. This would lead to an increased understanding of the quality of ICC systems and ease the development and validation of new evaluation measures.

References

[1] C. Lüke, G. Schmidt, A. Theiß, and J. Withopf, "In-car Communication", in Smart Mobile In-Vehicle Systems (G. Schmidt, H. Abut, K. Takeda, and J. H. Hansen, eds.), pp. 97–118, Heidelberg, Germany: Springer New York, 2014.

[2] K. Linhard and J. Freudenberger, "Passenger In-car Communication Enhancement", in Proc. EUSIPCO, Vienna, Austria, 2004.

[3] C. Lüke, H. Özer, G. Schmidt, A. Theiß, and J. Withopf, "Signal Processing for In-car Communication Systems", 5th Biennial Workshop on DSP for In-Vehicle Systems, Kiel, Germany, 2011.

[4] A. Theiß, G. Schmidt, J. Withopf, and C. Lüke, "Instrumental Evaluation of In-car Communication Systems," in Proc. 11. ITG-Fachtagung Sprachkommunikation, Erlangen, Germany, 2014.

[5] S. Möller, ed., Assessment and Prediction of Speech Quality in Telecommunications. Springer Science + Bussiness Media Drodrecht, 2000.

[6] K. Scholz, "Instrumentelle Qualitätsbeurteilung von Telefonbandsprache beruhend auf Qualitätsattributen", PhD thesis, Christian-Albrechts-Universität zu Kiel, Arbeiten über Digitale Signalverarbeitung Nr. 32, Shaker-Verlag, Aachen, 2008.

[7] C. Lüke, A. Theiß, G. Schmidt, O. Niebuhr, and T. John, "Creation of a Lombard Speech Database Using an Acoustic Ambiance Simulation with Loudspeakers", 6th Biennial DSP Workshop for In-Vehicle Systems, Seoul, Korea, 2013.

[8] H. Gierlich and F. Kettler, "Advanced Speech Quality Testing of Modern Telecommunication Equipment: An Overview", Signal Processing, Applied Speech and Audio Processing, vol. 86, no. 6, pp. 1327 – 1340, 2006.

[9] A. Theiß, "Instrumentelle Evaluierung von Innenraumkommunikationssystemen", PhD thesis, Christian-Albrechts-Universität zu Kiel. 2017.

[10] R. Landgraf, O. Niebuhr, G. Schmidt, T. John, C. Lüke, and A. Theiß, "Von der Straße ins Labor: Die Modifikation der Sprachproduktion bei lauten Fahrgeräuschen", Proc. "Fortschritte der Akustik", 41. Jahrestagung für Akustik (DAGA), Nürnberg, Germany, 2015.

[11] ITU-T Focus Group on Car Communication, Final Report, Version 1.0, November 2013.

[12] ITU-T Recommendation P.505, One-view Visualization of Speech Quality Measurement Results, Nov. 2005.