Active-Road-Noise- Cancellation (ARNC) – Speech-Enhancement through Noise-Reduction in Hands-free Systems

Juergen Zollner[1], Dr. Vasudev Kandade Rajan[1], Dr. Nikos Zafeiropoulos[1], Dr. Markus Christoph[1]

[1] *Harman/Becker Automotive Systems GmbH, 94315 Straubing, Germany*

Introduction

Structure-borne road-noise is the noise inside the vehicle caused by the transmission of vibrations and its structural resonances from the tires through the chassis to the inside of the cabin [1]. Here, the Active-Road-Noise-Cancellation (ARNC) is able to reduce the in-vehicle perceived road-noise in the acoustic domain for the passengers using existing cabin loudspeakers as secondary sources. Below, the benefit of an already available ARNC infrastructure for an omnipresent hands-free system is presented. Stationary and quasi-stationary vehicle noise is well handled in state-of-the art hands-free systems [2], but very low frequency and fast and variable noise components like impulsive road-noise scenarios are challenging. In such situations, the ARNC can work in tandem with a hands-free system to reduce such unpleasant road-noise. The ARNC-system here is applied to improve a local speech situation, where it is canceling road-noise caused by driving on a cobblestone road.

Acoustic and Electrical Domain

A typical ARNC-System [3] for road-noise cancellation in the acoustic domain usually has omnidirectional error microphones close to passenger's heads in order to have passenger's ears in a zone of silence, see figure 1.

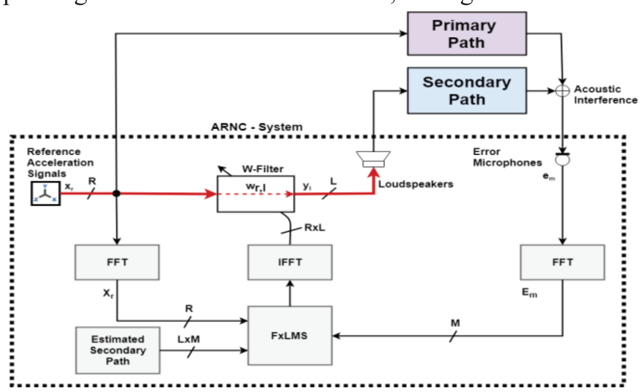

Figure 1: Block diagram of a multichannel ARNC-System in the acoustical domain with primary and secondary transfer paths utilizing an FxLMS [4] algorithm with vehicle components including accelerometer sensors, error-microphones and loudspeakers used as actuators.

Here, the ARNC-System use digital accelerometer sensors to acquire the electrical signals from the structure borne road-noise vibrations and use them as reference signals x_r for a feed-forward adaptive algorithm. In case of acoustical cancelation, the error microphones close to the passenger's ears capture the vehicle noise and feed their signals e_m to an adaptive algorithm to identify the underlying system and the loudspeakers generate the control signals y_l to cancel the road-noise. Unfortunately, the microphones used in hands-free systems are mostly placed in the middle of the cabin, close to or nearby the rear mirror, and have a built-in directivity. Therefore, it may be contra productive to create an additional acoustical zone of silence at the hands-free microphone locations, since this would cause more effort and complexity for the existing ARNC-System in the acoustic domain and potentially cause a compromise on cancelation performance for the front passenger positions. To avoid such interdependencies and because either a machine or a far-end person would listen to the hands-free microphone content, the ARNC for the hands-free system is applied in the electrical domain, as shown in figure 2.

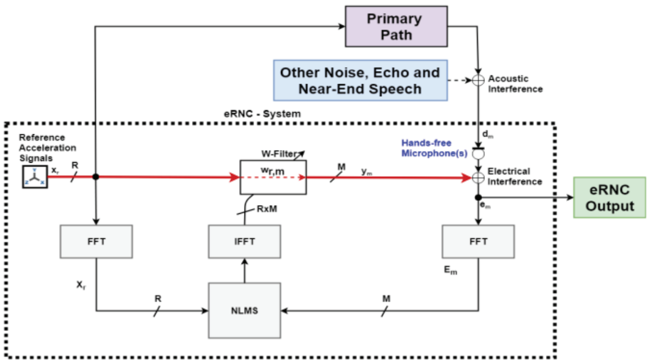

Figure 2: Block diagram of a multichannel ARNC-System in electrical domain with primary transfer paths utilizing an NLMS algorithm with vehicle components including accelerometer sensors and hands-free microphones.

Here as well the ARNC-Control structure is very similar to the acoustic domain cancelation, both use the same reference signals x_r. Since the control signal y_m and hands-free microphone signal d_m interfere only in the electrical domain, no secondary path needs to be considered and the adaptive controller use a specialized least mean squares (LMS) algorithm instead a dedicated filtered reference least mean squares (FxLMS) algorithm in the acoustic domain. The implementation of an ARNC-System in the electric domain allows two major advantages. First, the ARNC-System can still fully rely on the available multi-reference sensors signals for good coherence but optimize control complexity on each desired individual hands-free microphone. Second, since no plant system needs to be considered, there are no issues and side-effects by deviations and limitations of secondary path. But there is also a possible disadvantage. The ARNC in the acoustic domain may cause increased road-noise levels next to the hands-free microphones, since the road-noise and the anti-noise from the loudspeaker may constructively interfere at hands-free microphone positions.

ARNC in the Electrical Domain

Although as mentioned the ARNC-System for the electrical domain can be simplified in dimension and complexity compared to the complexity of an ARNC-System in the acoustic domain, still the adaptive control is a complex MIMO-System, with m = 1..M hands-free microphone signals and r = 1..R reference signals. Below in figure 3, a block diagram of a multiple-channel system for active control of road noise in the electrical domain is provided.

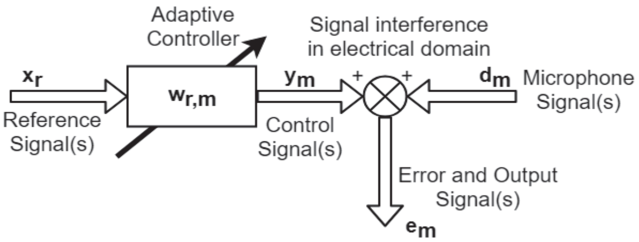

Figure 3: Top level control block diagram of an ARNC-System with signal interference in the electrical domain.

Here, the w-Filter matrix $w_{r,m}$ with a length of L taps needs to be calculated and applied to the reference signals x_r to generate the control signals y_m within the forward path in the time domain to minimize the system latency to meet the hands-free low-latency requirements. The update is performed in the spectral domain to lower the calculation effort which utilize the spectral reference signals X_r and spectral error-signals E_m for the normalized least mean squares (NLMS) algorithm to have an iterative W-Filter update. So, in context of above block diagram, the error-signal e_m in the time domain for each hands-free microphone signal can be calculated as:

$$e_m(n) = d_m(n) + \sum_{r=1}^{R}\sum_{i=0}^{L-1} w_{r,m}(i) \cdot x_r(n-i) \tag{1}$$

The equation in vector form can be written as:

$$e_m(n) = d_m(n) + \boldsymbol{w}^T \boldsymbol{x}(n) \tag{2}$$

With:

$$\boldsymbol{w} = [w(0), w(1), \dots, w(L-1)]^T \tag{3}$$

$$\boldsymbol{x}(n) = [x(n), x(n-1), \dots, x(n-L+1)]^T \tag{4}$$

For the sake of better readability we drop the microphone subscript index m. The cost function for each microphone J is defined by the mean square of the error-microphone signal:

$$J = E[e^2(n)] \tag{5}$$

$$J = E[e^T(n)e(n)] \tag{6}$$

By substituting $e(n)$ as defined in equation (2), the cost function can be rewritten as:

$$J = E[d^T(n)d(n)] + 2 \cdot \boldsymbol{w}^T E[\boldsymbol{x}^T(n)d(n)] + E[\boldsymbol{x}^T(n)\boldsymbol{x}(n)]\boldsymbol{w} \tag{7}$$

This quadratic form has a minimum value, once **w** is assumed to be the optimum value $\mathbf{w_0}$ in case of stationary time signals are available:

$$\boldsymbol{w_o} = -\{E[\boldsymbol{x}^T(n)\boldsymbol{x}(n)]^{-1}\}\{E[\boldsymbol{x}^T(n)d(n)]\} \tag{8}$$

The optimum $\mathbf{w_0}$ can be used to calculate the minimum cost function $\mathbf{J_0}$:

$$\boldsymbol{J_o} = E[d^T(n)d(n)] - \{E[\boldsymbol{x}^T(n)d(n)]\}^T \{E[\boldsymbol{x}^T(n)\boldsymbol{x}(n)]\}^{-1} \{E[\boldsymbol{x}^T(n)d(n)]\} \tag{9}$$

For a real time implementation and for a more practical approach, a gradient decent algorithm is used to calculate the controller coefficients **w** so the adaptive controller converge over time using a steepest decent algorithm:

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \frac{\mu}{2} \cdot \frac{\partial \boldsymbol{J}}{\partial \boldsymbol{w}} = \boldsymbol{w}(n) - \mu \cdot E[\boldsymbol{x}^T(n)e(n)] \tag{10}$$

The effort to calculate the mean value can be dropped as shown in literature [5], 6], with a compromise in convergence speed as each iteration may not be aligned with its steepest gradient. Therefore, the update equation (10) can be rewritten as:

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \mu \cdot \boldsymbol{x}^T(n)e(n). \tag{11}$$

Here, the step-size μ is a small fixed value for robust and stable operation. Further, the least mean squares (LMS) algorithm can be made robust by reference signal normalization:

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) - \mu \cdot \frac{\boldsymbol{x}^T(n)e(n)}{\|x(n)\|^2}. \tag{12}$$

This algorithm allows a reference signal independent step-size selection. The NLMS is implemented in the spectral domain for reducing the computational complexity.

NLMS – Algorithm Evaluation

The electrical ARNC algorithm is evaluated using real vehicle data recordings. The recordings contain 12 reference channels acquired from accelerometer sensors mounted on the chassis of a Porsche Panamera, but only 6 channels were required for font-left position cancelation. The recordings also in include one microphone signal synchronized with accelerometer data, mounted on the cabin ceiling in the front-left driver position. All measurements were done while the car was driving over a cobblestone road at a constant speed of 30 [km/h] for about 40 seconds. With the above discussed algorithm which was tuned for the best performance in the electrical domain a robust and effective cancelation performance could be achieved, see below fig. 4.

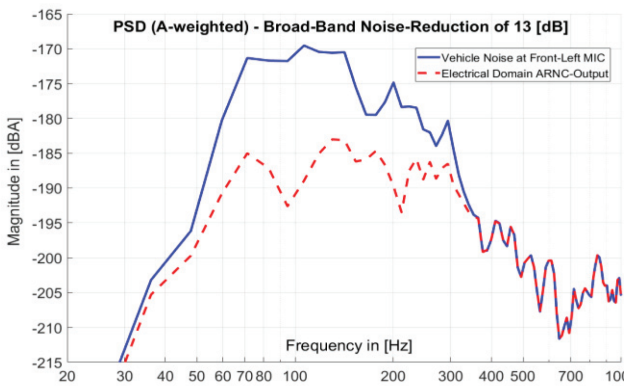

Figure 4: PSD diagram of ARNC-System in the electrical domain for a cobblestone road at 30km/h (solid line ARNC off, dashed line ARNC on) utilizing an NLMS algorithm.

The performance was not a surprise, as the same sensor and microphone setup is used within the already integrated ECU solution for a real-time ARNC-System in the acoustic domain [7]. Here, clearly the limitation of available digital sensors can be observed due to this constraint, the limited working band-width is causing insufficient coherence above 300 [Hz]. Therefore, the cobblestone road-noise example was chosen, as here the rumble and drum structure-born road-noise components are very dominant below 300 [Hz]. As already highlighted, the electrical cancelation of structure-born road-noise is not limited by a secondary path and its physical wavelength constraint in the acoustic domain. So, the only limiting factor is the coherence between the reference sensor data and the microphone signal [8].

Speech Enhancement and Noise Reduction

An additional evaluation was made regarding speech enhancement. Here, by simulation the prerecorded microphone data was mixed with clean speech and the output after the road-noise cancelation was analyzed. Both signal components are shown below in figure 5.

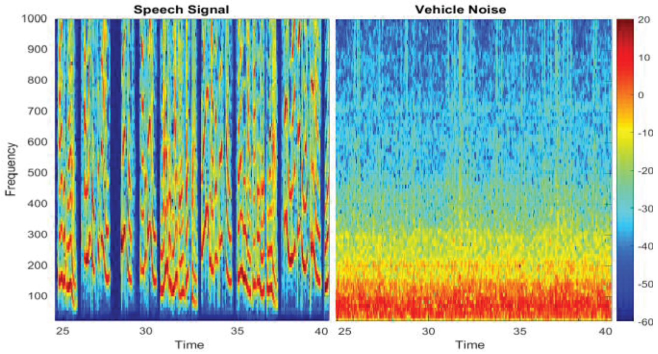

Figure 5: Spectrogram of last 10 [sec.] of clean speech signal (left) and prerecorded vehicle noise (right) for a cobblestone road at a speed of 30 [km/h] at the front-left driver ceiling position.

Since the road-noise cancelation is expected to take effect below 300 [Hz] the expectations were not an enhancement in articulation index [9] but for sound quality [10] to improve listening comfort for a person on far-end side without any degradation on the local speech signal. A low valued fixed step-size was found to be effective here, most likely due to the low correlation between speech and reference signal. The simulation output, shown in figure 6, could match those expectations.

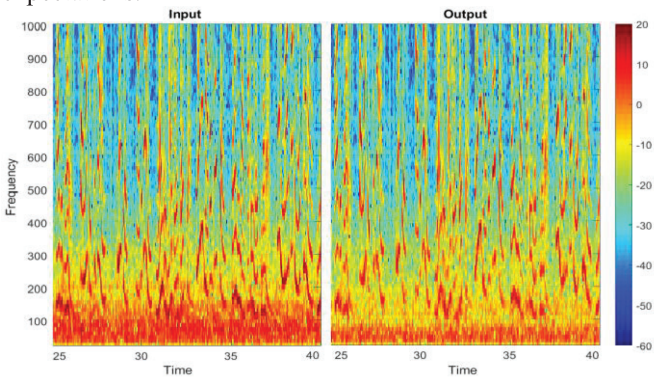

Figure 6: Spectrogram of last 10 [sec.] of mixed clean speech with vehicle noise (left) and simulated ARNC output (right) for a cobblestone road at a speed of 30 [km/h] at the front-left driver ceiling position.

Further evaluations on different driving situations confirmed that a fixed but moderately selected step-size allows a stable and robust but still sufficient ARNC performance. In case the step-size was to large, the performance might further improve but at the cost of perceivable speech degeneration.

Summary

In parallel to an ARNC-System in the acoustic domain an ARNC-System in the electrical domain is advisable, since the signals from the already existing accelerometer reference sensors are already available. In case the hands-free signal processing is not in the same ECU as the ARNC-Controller, a delay within the reference signals can be tolerated to some extent, since in the electrical domain a delay can be easily compensated, as long as the overall end-to-end latency between near-end and far-end talker is not significantly affected. For example, a delay of 5...10 [ms], caused by vehicle analog or digital bus transport from the ARNC-Controller ECU to hands-free processing ECU, can be handled by a tunable delay after the hands-free microphone signal before it is summed with the control signal. In order to guarantee the causality such compensation delay should allow some latency reserve. This would benefit the algorithm stability and robustness on unexpected primary path changes and may cost a reasonable number of control filter taps. Simulation results of real vehicle recordings have shown, that the structure-born road-noise can be adequately canceled and can significantly improve far-end listening comfort without causing artefacts on the desired speech signal.

Literature

[1] Zeller, P. (2009). Handbuch Fahrzeugakustik (1st Edition ed.). Vieweg+Teubner.

[2] Kandade Rajan, V. (2017). Speech Enhancement in Hands-free Systems for Automobile Environments (Doctoral dissertation), Shaker Verlag.

[3] Sutton, T., Elliott, S., McDonald, M., Saunders T. (1994). Active control of road noise inside vehicles. Noise Control Engineering Journal 42.

[4] Kuo, S. M., Morgan, D. R. (1996). Active Noise Control Systems: Algorithms and DSP Implementations. John Wiley & Sohns, Inc.

[5] Widrow, B., Stearns, S. D. (1985). Adaptive Signal Processing, Prentice Hall.

[6] Elliott, S. J., Stothers, I. M. and Nelson, P. A. (1987). A multiple-error LMS algorithm and its application to the active control of sound and vibration, IEEE Trans ASSP-35, pp. 1423–1434.

[7] Zollner, J., Zafeiropoulos, N., Christoph, M., Kandade Rajan, V. (2017). Active Road Noise Cancelation: Off-the-Shelf Digital Audio Amplifier Platform. AES International Conference on Automotive Audio.

[8] Zafeiropoulos, N., Moorhouse, A., Ballatore, M., Mackay, A., (2015). Active control of road noise: The relation between the reference sensor location and the effect on the controller's performance. Proceedings in 22nd International Congress of Sound of Vibration.

[9] Mueller, G. and Killion, M. (1990). An Easy Method for Calculating the Articulation Index, The Hearing Journal, 43(9), pp.14-17.

[10] Zafeiropoulos, N., Zollner, J., Kandade Rajan, V. (2018). Active Road Noise Cancellation for the Improvement of Sound Quality in the Vehicle. ATZ worldwide 120.