Influence of the Waveguide Propagation on the Antenna Performance in a Car Cabin

Leonid Krasny and Ali Khayrallah

Ericsson Research
P.O.Box 13969, RTP, N.C. 27709, USA
Email: euslekr@am1.ericsson.se
ali.khayrallah@ericsson.com

Abstract

This paper presents a novel array processing algorithm for noise reduction in a hands free car environment. The algorithm incorporates the spatial properties of the sound field in a car cabin and a constraint on allowable speech signal distortion. Our results indicate that the proposed algorithm gives substantial performance improvement of 15-20 dB in comparison with the conventional array processing which is based on a coherent model of the signal field.

1 Introduction

In a car environment, the speech signal is often corrupted by background noise, which degrades the performance of the coding and recognition algorithms. It is essential to reduce the noise level without distorting the original speech signal. Two conventional approaches have been developed to solve this problem. One is a single microphone noise reduction technique, which utilizes differences in the spectral characteristics of the speech signal and the background noise [1]-[2]. It is hampered by the fact that in many situations the speech and the noise tend to have similar spectral distributions. Under these conditions, the single-microphone noise reduction technique will not yield substantial improvement in speech intelligibility. On the other hand, the signal and the noise in a car cabin are acoustical fields, which have different spatial characteristics. This is the basis of the second approach, which exploits the spatial separation of the speech signal and the car noise.

It is known that spatial signal processing requires microphone that combine the outputs of several microphones. Several array processing algorithms have been proposed [3]-[6]. The general structure of these algorithms can be expressed using the same basic equation which has the following form in a frequency domain:

$$U_{out}(\omega) = \sum_{i=1}^{M} U(\omega, \mathbf{r_i})\, H^*(\omega, \mathbf{r_i}), \qquad (1)$$

where $U_{out}(\omega)$ and $U(\omega, \mathbf{r_i})$ are respectively the Fourier transform of the array output and the field $u(t, \mathbf{r_i})$ observed at the i-th microphone with the spatial coordinates $\mathbf{r_i}$, $H(\omega, \mathbf{r_i})$ is the frequency response of the filter at the i-th microphone-array element, and M is the total number of the microphones at the array.

The determination of the functions $H(\omega, \mathbf{r_i})$ in Eq.(1) is the major area of concern in array processing, and a considerable amount of effort has been spent in this area [3]-[6]. However, there exist some difficulties, which become apparent when the conventional array processing technique is applied to noise reduction in a car environment. The problem is that the conventional approach is based on an assumption that the signal field in a car cabin has a coherent spatial structure. Meanwhile, experimental data shows that this assumption is unrealistic for a car environment where the sound source and receiving antenna are in a small enclosure, and wall irregularities scatter the signal waves propagating inside the car cabin. Therefore, the effects of waveguide sound propagation should be taken into account for synthesis of the practical noise reduction system.

In this paper the structure and performance of array processing algorithms is investigated based on a non-coherent signal model. The paper is organized as follows. In Section 2 the signal and noise field models are presented. In Section 3 an optimization criterion which utilizes these models is introduced, and a new optimal array processing algorithm based on the criterion is proposed. In Section 4 the performance of the optimal algorithm is analyzed and compared with the array processing algorithm which is based on a coherent model of the signal field. Section 5 concludes the paper.

2 Signal and Noise Models

We assume that the field in a car is a superposition of the signal field $s(t, \mathbf{r})$ (speech) and background noise $N_b(t, \mathbf{r})$ (road noise, wind noise, engine noise). Let us assume also that the receiving microphone array with known but arbitrary geometry is placed into the car. When a mixture of the signal $s(t, \mathbf{r})$ and background noise $N_b(t, \mathbf{r})$ are incident on the array, the field $u(t, \mathbf{r_i})$ received by the i-th array element has the form:

$$u(t, \mathbf{r_i}) = s(t, \mathbf{r_i}) + N_b(t, \mathbf{r_i}) + n(t, \mathbf{r_i}), \quad (2)$$

where $n(t, \mathbf{r_i})$ is a spatial-time uncorrelated internal noise of the array.

If the speech generator (talker) is a simple point source with the spatial coordinates $\mathbf{r_0}$, the signal spectrum $S(\omega; \mathbf{r_i})$ at the i-th microphone has the form

$$S(\omega; \mathbf{r_i}) = S(\omega)\; G(\omega; \mathbf{r_i}, \mathbf{r_0}), \quad (3)$$

where $S(\omega)$ is a spectrum of the speech, and $G(\omega; \mathbf{r_i}, \mathbf{r_0})$ is the Green function which describes the propagation channel between the talker and the i-th microphone.

Let us assume that the car enclosure has a rectangular shape with sides H_x (x-axis), H_y (y-axis), and H_z (z-axis). Than the Green function has the form [7]:

$$G(\omega, \mathbf{r}, \mathbf{r_0}) = \sum_{n,m,p} A_{nmp}\; \Psi_{nmp}(\mathbf{r})\; \Psi_{nmp}(\mathbf{r_0}), \quad (4)$$

where

$$A_{nmp} = \{k^2 - \gamma_{xn}^2 - \gamma_{ym}^2 - \gamma_{zp}^2\}^{-1}, \quad (5)$$

$$\Psi_{nmp}(\mathbf{r}) = \frac{\varphi_{xn}(x)\; \varphi_{ym}(y)\; \varphi_{zp}(z)}{\|\varphi_{xn}\|\; \|\varphi_{ym}\|\; \|\varphi_{zp}\|}, \quad (6)$$

the functions $\varphi_{ln}(l)$ ($l = x, y$, or z) are the normal modes [7], the coefficients γ_{ln} are the characteristic values of the normal modes $\varphi_{ln}(l)$, $k = \omega/c$ is the wave number, c is the speed of sound.

Eq.(4) shows that the signal field in a car cabin is the sum of the waves corresponding to the different normal modes, each with amplitude proportional to the values of the coefficient A_{nmp}. If the walls are regular, the standing wave shapes are not changed by the yielding of the walls. Actually, of course, the fact that the walls have irregularities affect the shape of the standing waves. In other words, the original standing wave is modified by the scattered wave produced by the reaction of the wall to the standing wave. Each standing wave becomes a more or less a random mixture of a large number of the normal modes, corresponding to wave motion in many directions. As a result, the signal spatial correlation function has the form:

$$K_S(\omega; \mathbf{r_i}, \mathbf{r_k}) = g_s(\omega)\; E\{\; G(\omega; \mathbf{r_i}, \mathbf{r_0})\; G^*(\omega; \mathbf{r_k}, \mathbf{r_0})\;\}, \quad (7)$$

where $g_s(\omega)$ is the power spectral density (PSD) of the speech signal, and $E\{\cdot\}$ represents the statistical average over the normal modes.

Substituting Eq.(4) into Eq.(7) and averaging the normal modes we obtain the following equation for the signal spatial correlation function:

$$K_S(\omega; \mathbf{r_i}, \mathbf{r_k}) =$$
$$= g_s(\omega) \sum_{n,m,p} A_{nmp}^2\; |\Psi_{nmp}(\mathbf{r_0})|^2\; \Psi_{nmp}(\mathbf{r_i})\; \Psi_{nmp}^*(\mathbf{r_k}). \quad (8)$$

In a car environment the background noise is generated by the car engine, road, and wind hitting the windows. The acoustical model for the background noise field can be represented as a superposition of an infinite number of uncorrelated random point sources located over some surface S_0. Using this representation, it follows that the spatial correlation function of the background noise is given by

$$K_N(\omega; \mathbf{r_i}, \mathbf{r_k}) =$$
$$= \; g_N(\omega) \int_{S_0} E\{G(\omega, \mathbf{r_i}, \mathbf{r})\; G^*(\omega, \mathbf{r_k}, \mathbf{r})\}\; d\mathbf{r}, \quad (9)$$

where $g_N(\omega)$ is the power spectral density (PSD) of the noise sources.

Let us assume that the window noise sources are located on the surfaces $x = \pm(H_x/2)$ and the front noise sources are located on the surface $y = H_y/2$. Under this assumption the spatial correlation function of the background noise has the form:

$$K_N(\omega; \mathbf{r_i}, \mathbf{r_k}) =$$
$$= g_N(\omega) \sum_{n,m,p} A_{nmp}^2\; \Psi_{nmp}(\mathbf{r_i})\; \Psi_{nmp}^*(\mathbf{r_k}) \times$$
$$\times\; [\; \varphi_{xn}^2(H_x/2)\; + \varphi_{ym}^2(H_y/2)\;]. \quad (10)$$

In the next section we use these models to synthesize the optimal array processing algorithm and analyze its performance.

3 Optimal Algorithm

Our goal is to develop a noise reduction space-time processing system for speech signals, the output of which is acceptable to the human ear. We realize that significant degradation in the desired signal is unacceptable, even if the noise level is greatly reduced. Therefore, the constraint $g_S^{out}(\omega) = g_s(\omega)$ is imposed when formulating the optimization criterion, where $g_S^{out}(\omega)$ is the signal spectral density after array processing.

We can also introduce some weighting function $B(\omega)$ which represents *a priori* desired distortion of the speech signal. Then the constraint for desired output signal has the form

$$g_S^{out}(\omega) = g_s(\omega)\ |B(\omega)|^2. \tag{11}$$

Based on this idea, we consider the optimization problem for the functions $H(\omega, \mathbf{r_i})$ as one of minimizing the noise spectral density after array processing subject to the constraint (11).

It follows from Eq.(1) that output signal spectral density

$$g_S^{out}(\omega) = \sum_{i=1}^{M}\sum_{k=1}^{M} K_S(\omega; \mathbf{r_i}, \mathbf{r_k}) H^*(\omega, \mathbf{r_i}) H(\omega, \mathbf{r_k}), \tag{12}$$

and the output noise spectral density

$$g_N^{out}(\omega) = \sum_{i=1}^{M}\sum_{k=1}^{M} K_N(\omega; \mathbf{r_i}, \mathbf{r_k}) H^*(\omega, \mathbf{r_i}) H(\omega, \mathbf{r_k}). \tag{13}$$

Therefore, the optimal weighting functions $H(\omega, \mathbf{r_i})$ in (1) are obtained as a solution of the variation problem

$$H_{opt}(\omega, \mathbf{r_i}) =$$

$$= arg\{\ \min \sum_{i=1}^{M}\sum_{k=1}^{M} K_N(\omega; \mathbf{r_i}, \mathbf{r_k}) H^*(\omega, \mathbf{r_i}) H(\omega, \mathbf{r_k})\ \} \tag{14}$$

subject to the constraint

$$\sum_{i=1}^{M}\sum_{k=1}^{M} K_S(\omega; \mathbf{r_i}, \mathbf{r_k}) H^*(\omega, \mathbf{r_i}) H(\omega, \mathbf{r_k}) \ =$$

$$=\ g_s(\omega)\ |B(\omega)|^2. \tag{15}$$

The solution of the optimization problem (14)-(15) gives the optimal weighting functions at antenna elements

$$H_{opt}(\omega, \mathbf{r_i}) \ = \ \frac{B(\omega)}{\sqrt{\nu_{max}(\omega)}}\ \Phi_{max}(\omega, \mathbf{r_i}), \tag{16}$$

where $\Phi_{max}(\omega, \mathbf{r_i})$ are the elements of the eigenvector $\mathbf{\Phi}_{max}(\omega) : \{\Phi_{max}(\omega, \mathbf{r_1}), .., \Phi_{max}(\omega, \mathbf{r}_M)\}$ which corresponds to the largest eigenvalue $\nu_{max}(\omega)$ of the matrix $\mathbf{K}(\omega) = \mathbf{K_N}^{-1}(\omega)\mathbf{K_S}(\omega)$, i.e.

$$\sum_{k=1}^{M} K(\omega; \mathbf{r_i}, \mathbf{r_k})\ \Phi_{max}(\omega, \mathbf{r_k}) = \nu_{max}(\omega) \cdot \Phi_{max}(\omega, \mathbf{r_i}). \tag{17}$$

Let us consider now a special case when the scattering is absent. In this case the elements of the matrix $\mathbf{K}(\omega)$ are

$$K(\omega; \mathbf{r_i}, \mathbf{r_k}) = H_0(\omega, \mathbf{r_i})\ G^*(\omega; \mathbf{r_k}, \mathbf{r_0}), \tag{18}$$

where the functions $H_0(\omega, \mathbf{r_i})$ satisfy the system of equations

$$\sum_{k=1}^{M} K_N(\omega; \mathbf{r_i}, \mathbf{r_k}) H_0(\omega, \mathbf{r_k}) = G(\omega; \mathbf{r_i}, \mathbf{r_0}). \tag{19}$$

The largest eigenvalue of the matrix (18) is

$$\nu_{max}(\omega) = \sum_{i=1}^{M} G(\omega; \mathbf{r_i}, \mathbf{r_0})\ H_0^*(\omega, \mathbf{r_i}). \tag{20}$$

Substituting Eq.(18) and Eq.(20) into Eq.(17) we obtain the eigenvector which corresponds to the maximum eigenvalue (20):

$$\Phi_{max}(\omega, \mathbf{r_i}) = \frac{H_0(\omega, \mathbf{r_i})}{\sqrt{\sum_{i=1}^{M} G(\omega; \mathbf{r_i}, \mathbf{r_0})\ H_0^*(\omega, \mathbf{r_i})}}.$$

Therefore, the optimal weighting functions for coherent model of the signal field has the form

$$H_{opt}(\omega, \mathbf{r_i}) = \frac{B(\omega)\ H_0(\omega, \mathbf{r_i})}{\sum_{i=1}^{M} G(\omega; \mathbf{r_i}, \mathbf{r_0})\ H_0^*(\omega, \mathbf{r_i})}, \tag{21}$$

which coincides with the result in [6].

4 Performance

In this section, we demonstrate the performance of the microphone array with the optimal array processing, which takes into account the scattering effects, and the conventional array processing

$$U_{out}(\omega) = \frac{B(\omega) \sum_{i=1}^{M} U(\omega, \mathbf{r_i})\ H_0^*(\omega, \mathbf{r_i})}{\sum_{i=1}^{M} G(\omega; \mathbf{r_i}, \mathbf{r_0})\ H_0^*(\omega, \mathbf{r_i})}, \tag{22}$$

which is based on a coherent model of the signal field.

The array performance can be characterized by the output signal-to-noise ratio $\mu(\omega)$, which we define as a ratio of the signal spectral density to the noise spectral density after array processing.

The signal spectral density after optimal array processing is described by Eq.(11) and the output noise spectral density is described by Eq.(13). Substituting Eq.(16) into Eq.(13) and taking into account orthogonal properties of the eigenvectors $\Phi_{max}(\omega, \mathbf{r_i})$ we obtain

$$g_N^{out}(\omega) = \frac{B^2(\omega)}{\nu_{max}(\omega)}.$$

Therefore, the signal-to-noise ratio (SNR) after the optimal array processing has the form:

$$\mu_{opt}(\omega) = g_s(\omega)\ \nu_{max}(\omega). \tag{23}$$

By virtue of Eq.(19) the SNR after the conventional array processing (22) has the form

$$\mu(\omega) = \frac{\sum_{i=1}^{M}\sum_{k=1}^{M}\ K_S(\omega; \mathbf{r_i}, \mathbf{r_k}) H_0^*(\omega, \mathbf{r_i}) H_0(\omega, \mathbf{r_k})}{\sum_{i=1}^{M} G(\omega; \mathbf{r_i}, \mathbf{r_0})\ H_0^*(\omega, \mathbf{r_i})} \tag{24}$$

It is clear that for coherent signal field the Eq.(24) has the form

$$\mu(\omega) \equiv \mu_{opt}(\omega) = g_s(\omega) \sum_{i=1}^{M} G(\omega; \mathbf{r_i}, \mathbf{r_0})\, H_0^*(\omega, \mathbf{r_i}). \tag{25}$$

The results of calculations according to Eq.(23) and Eq.(24) are shown in Fig.1, which presents the normalized signal-to-noise ratio $SNR_{norm}(\omega) = \mu(\omega)/g_s(\omega)$ as a function of frequency. The calculations have been carried out for a linear array with 4 equally spaced microphones. The distance between microphones is 5 cm. The solid line corresponds to the SNR for optimal array processing which takes into account the scattering effects in a car cabin, and the dashed line shows the SNR for conventional array processing (22) which is based on a coherent signal field model.

It is evident from Fig.1 that the SNR for conventional array processing, which does not take into account scattering effects, is significantly poorer than for the optimal processing. At high frequencies the SNR for optimal processing is about 6-8 dB better than for conventional array processing. The highest amount of loss occurs at low frequencies where the array is shortest with respect to the acoustic wavelength and thus has the poorest directivity. In this case, the optimal processing gives an improvement of 15-20 dB with respect to conventional processing. It should be mentioned also that the optimal algorithm gives substantial performance improvement of 20 dB at low frequencies in comparison with the single microphone (dashed line with marks in Fig.1).

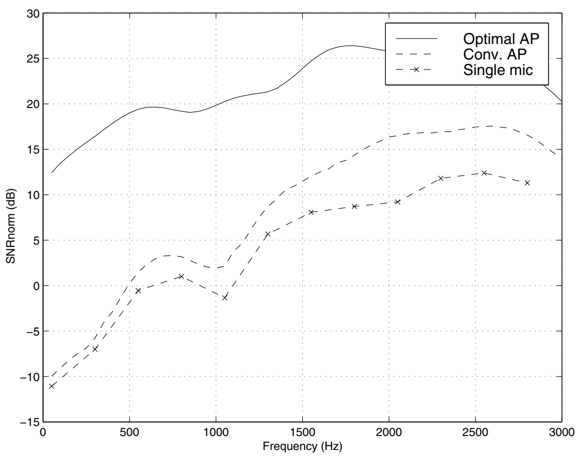

Figure 1: SNR after array processing

5 Conclusions

In this paper a novel array processing algorithm for the hands-free mobile telephone applications is proposed. This algorithm employs an optimization criterion with a constraint which allows to minimize output noise spectral density and takes into account the scattering effects of sound propagation in a car cabin.

Our results indicate that the optimal algorithm can work very well in car cabin and substantially (15-20 dB) attenuate a noise field without suppression and degradation of the speech signal. Meanwhile, the array processing algorithm which is based on a coherent model of the signal field leads to a significant signal distortion after array processing and has significantly poorer performance than the optimal algorithm.

References

[1] Boll S. F., "Suppression of acoustic noise in speech using spectral subtraction", *IEEE Trans. on Acoustics, Speech, and Signal Processing*. ASSP-27(2):113-120, April 1979.

[2] McAulay R. J. and Malpass M. L., "Speech enhancement using a soft-decision noise-suppression filter", *IEEE Trans. on Acoustics, Speech, and Signal Processing*. ASSP-28:137-145, 1980.

[3] Kaneda Y. and Ohga J., "Adaptive microphone-array system for noise reduction", *IEEE Trans. on Acoustics, Speech, and Signal Processing*. ASSP-34(6):1391-1400, December 1986.

[4] Cox H., Zeskind R. M., and Owen M. M., "Robust adaptive beamforming", *IEEE Trans. on Acoustics, Speech, and Signal Processing*. ASSP-35(10):1365-1376, October 1987.

[5] Kates J. M., "Superdirective arrays for hearing aides", *J.Acoust.Soc.Amer.*, Vol.94, 1930-1933, 1993.

[6] Krasny L. G., "Noise reduction in a car cabin using microphone antenna arrays", *17th International Congress on Acoustics*, Vol.3, pp.37-39, Rome 2000.

[7] Morse P. M., *Vibration and sound*, 1991.