# Planar Superdirective Microphone Arrays for Speech Acquisition in the Car

*Rainer Martin, Alexey Petrovsky*[1]*, and Thomas Lotter*

Institute of Communication Systems and Data Processing
Aachen University of Technology, D-52056 Aachen, Germany
{martin | thomas}@ind.rwth-aachen.de, apetrows@bas-net.by

## Abstract

In this paper we investigate a small broadside planar (2D) superdirective microphone array for speech acquisition in the car and compare its performance to linear arrays. The objective of this investigation is to replace an expensive directional microphone by a small array of inexpensive omnidirectional sensors. Since the array was designed to be used in the car environment it has to satisfy restrictions with respect to size and to the number of microphones. For all array configurations we present theoretical gains, actual measured gains using low-cost microphones, and beam patterns. For a fixed number of microphones and fixed array dimensions we show that the planar design leads to slightly superior array gains.

## 1. Introduction

Speech picked up by a hands-free device in a car can be significantly degraded. The degradation is mostly due to ambient noise but also to reverberation. The ambient noise in the car is not only very annoying in hands-free telephone conversations but can also render voice controlled applications, such as navigation systems useless. It is therefore of great importance to improve the signal-to-noise ratio (SNR) of the acquired speech.

To achieve a high SNR, the speech input device must be very close to the speakers mouth or, if this is not possible, must be designed to have a high directivity. The high directivity can be achieved either by using a single directional microphone or by combining multiple microphones into a microphone array. Since the price of multi-channel A/D converters has significantly decreased over the last years, and omnidirectional microphones are also much cheaper (and also smaller) than directional microphones, the array approach becomes increasingly attractive. Furthermore, the array approach allows adaptive steering of the direction of maximal sensitivity and can be therefore automatically adjusted to the position of the speaker (which can be the driver or a passenger).

The objective of this study is to investigate planar (2D) superdirective array designs. The planar design has a number of potential benefits: More microphones can be placed closer to the speakers mouth than with a linear design and the spatial sensitivity can be controlled in two spatial dimensions. As it will be shown below the gain of a planar array design can be larger than the gain of a linear array given the same number of microphones. Although endfire arrays do exhibit higher gains for low frequencies than broadside arrays [1, 2, 3], they are not very practical for the given application. Broadside arrays appear to be more adequate for the car environment as they are easily integrated into the car interior and do not present a hazard in case of an accident.

Most studies on microphone arrays in the car used a linear array [4, 5, 6, 7] or a curved array [8, 9], possibly with subarrays, and an adaptive beamformer, e.g, the *Generalized Sidelobe Canceller* (GSC). However, since the noise field in the car is almost diffuse, the adaptive beamformer will not give significantly better results than a fixed beamformer. Also, the fixed beamformer will give the same performance independently of the signal-to-noise ratio. We therefore focus on beamformers with fixed directivity patterns. A planar design was also considered in [7]. However, in this study the maximum dimension (length) of the planar design was smaller than the competing linear design and therefore its performance was inferior to the linear array. In our investigation, the maximum dimension is the same for linear and planar designs.

The remainder of this paper is organized as follows. In Section 2 we review the design of superdirective arrays with special emphasis on planar arrays. In Section 3 we discuss our array design and the placement of the array in the car. After presenting the results of the theoretical gain computation in Section 4 we conclude with experimental results in Section 5.

## 2. Design of Planar Superdirective Arrays

Figure 1 depicts $N = 4$ microphones in a cartesian coordinate system $(X, Y, Z)$ and the corresponding filter bank structure of the filter-and-sum beamformer. The array design task amounts to finding the best microphone positions and to determining the coefficients of the filters. In our planar design, the microphones are all located in the (shaded) $(X, Y)$ plane. The origin is placed at the center of gravity of the microphone configuration. The (column) vectors $\vec{r}_n$, $n = 1 \ldots N$, where $N$ is the number of microphones, denote the position of the n-th microphone with respect to the origin of the coordinate system. The (sampled) microphone signals are denoted by $x_n(k)$. These signals are filtered with impulse and frequency responses $a_n(k)$ and $A_n(\Omega)$, respectively, where $\Omega = 2\pi f/f_s$ denotes a normalized frequency variable and $f_s$ the sampling frequency. $\widehat{s}(k)$ is the output signal of the array.

The direction of incidence of the speech signal is given by the unit vector $\vec{u}_0$ and defined by the angles $\theta_0$ (elevation) and $\phi_0$ (azimuth). Using the far field assumption, the delays of the source signal at the microphone locations with respect to a reference signal $x_0(k)$ picked up at the origin of the coordinate system can be computed on the basis of a plane wave model. The near field assumption employs a spherical wave model and also takes the attenuation of the waves from the source to the individual microphones into account.

Generally, the directional characteristics of an array can be described by means of the beam pattern $\Psi(\Omega, \theta, \phi)$, which is

[1] This work was carried out while on leave from Computer Science Department, Belarusian State University of Informatics and Radioelectronics, 6, P. Brovky St., Minsk 220027, Republic of Belarus.

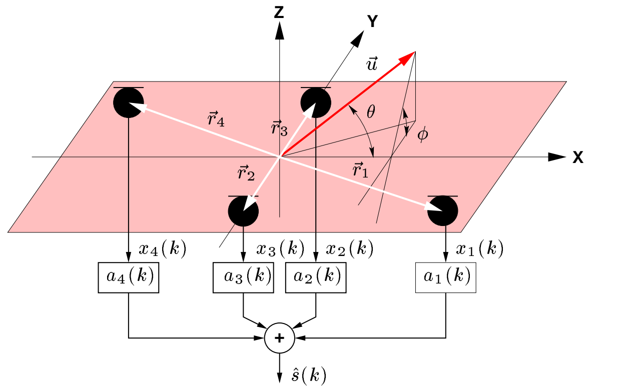


Figure 1: Design of a planar superdirective array with four microphones.

defined as the normalized power spectral density of the output signal $\hat{s}(k)$ as a function of frequency $\Omega$, and the direction of incidence, i.e., elevation $\theta$ and azimuth $\phi$. Using the far field assumption the beam pattern $\Psi(\Omega, \theta, \phi)$ can be written as

$$\Psi(\Omega, \theta, \phi) = \frac{\Phi_{\hat{s}\hat{s}}(\Omega, \theta, \phi)}{\Phi_{x_0 x_0}(\Omega)} = \left| \sum_{n=1}^{N} A_n(\Omega) \exp\left(j\beta \cdot \vec{r}_n^T \vec{u}\right) \right|^2 \tag{1}$$

where $\beta = 2\pi f/c = \Omega f_s/c$ is the wave number and $c$ is the speed of sound. With the notation shown in Fig. 1, the scalar product $\vec{r}_n^T \vec{u}$ in (1) is given by

$$\vec{r}_n^T \vec{u} = r_{xn} \cos\theta + r_{yn} \cos\phi \sin\theta \tag{2}$$

where $r_{xn}$ and $r_{yn}$ are the $X$ and $Y$ coordinates of vector $\vec{r}_n$, respectively. For $r_{yn} = 0$, $n = 1 \ldots N$, the array is a linear array. The beam pattern is then independent of $\phi$.

The directivity is measured by the gain $G(\Omega)$, which can be defined as the ratio of the beam pattern for the direction of principal incidence $\vec{u}_0$, $\Psi(\Omega, \theta_0, \phi_0)$, and the directivity averaged over all directions $\vec{u}$,

$$G(\Omega) = \frac{\Psi(\Omega, \theta_0, \phi_0)}{\sum_{n=1}^{N} \sum_{m=1}^{N} A_n(\Omega) A_m^*(\Omega) h_{mn}(\Omega)}\,, \tag{3}$$

where $h_{mn}(\Omega)$ is a normalized cross power spectral density of microphone signals $m$ and $n$. For the spherically isotropic (diffuse) noise field $h_{nm}(f)$ is given by

$$h_{mn}(\Omega) = \begin{cases} \frac{\sin(\Omega f_s d_{nm}/c)}{\Omega f_s d_{nm}/c} & \text{for } n \neq m \\ 1 & \text{for } n = m \end{cases} \tag{4}$$

where $d_{nm} = ||\vec{r}_m - \vec{r}_n||$ denotes the distance between the $n$-th and the $m$-th microphone.

When the noise in the microphone signals is spatially uncorrelated the array gain is given by the *white noise gain*

$$G_W(\Omega) = \frac{\Psi(\Omega, \theta_0, \phi_0)}{\sum_{n=1}^{N} \sum_{m=1}^{N} |A_n(\Omega)|^2}\,. \tag{5}$$

The inverse of $G_W(\Omega)$ is called susceptibility $K(\Omega) = 1/G_W(\Omega)$ and is used as a measure of the sensitivity of the array with respect to uncorrelated errors in the beamformer implementation. Such errors arise from self-noise as well as magnitude and phase errors of the microphones. The susceptibility is a crucial design parameter especially when low-cost microphones are used.

The design of a superdirective array aims at the maximization of the gain, while the susceptibility must not exceed a preset upper limit. As a result, the transfer functions $A_n(\Omega)$ is computed from [1]

$$A_n(\Omega) = \frac{\widetilde{A}_n(\Omega)}{\sum_{M=1}^{N} \widetilde{A}_m(\Omega) \exp\left(j\beta \cdot \vec{r}_m^T \vec{u}_0\right)} \tag{6}$$

where $\widetilde{A}_n(\Omega)$ is obtained from solving the system of linear equations

$$\sum_{m=1}^{N} h_{nm}(\Omega) \widetilde{A}_m(\Omega) + \mu \widetilde{A}_n(\Omega) = \exp\left(-j\beta \cdot \vec{r}_n^T \vec{u}_0\right) \tag{7}$$

for $1 \leq n \leq N$. In (7), $\mu$ denotes a Lagrangian multiplier, which allows to control the superdirectivity as well as the susceptability. For superdirective beamforming, the Lagrangian multiplier must be chosen from the interval $0 < \mu \ll 1$. The directivity and susceptability increase as $\mu$ tends towards zero. Since the Lagrangian multiplier cannot directly be computed from a preset susceptibility an iterative procedure was used to obtain the optimal multiplier for each frequency [10, 3]. The impulse responses $a_n(k)$ are obtained by solving the equations (6), (7) for equispaced frequencies, $0 \leq \Omega_i \leq \pi$, by taking the inverse DFT of length $M$ of frequency responses $A_n(\Omega_i)$, and by multiplying the resulting time-domain sequences with e.g., a Hamming window. For the array considered in this contribution, which are designed for a sampling frequency of $f_s = 8$ kHz, impulse responses $a_n(k)$ of length 256 have been used.

## 3. Array Placement and Geometry

After exploring several options we decided to place the array at the rear-view mirror. To allow comfortable viewing, the mirror will be adjusted to the position and the height of the driver and, hence, it will be also oriented towards the mouth of the driver. On the other hand, the integration of the array into the mirror restricts the maximum size of the array. Furthermore, since the array approach competes with a single directional microphone, the array elements have to be inexpensive omnidirectional (pressure) microphones.

Figure 2 depicts the microphone array which was used in our experiments. We found that using more than five microphones out of the available eight microphones did not give a significant advantage. Therefore, only a subarray with five microphones was be employed, the best linear and the best planar configuration as well as the corresponding delay-and-sum beamformers are specified in Table 1. When the array in Fig. 1 is oriented such that the positive X-axis points towards the driver side the driver in our test car is located at $\theta_0 = 70^o$ and $\phi_0 = 102^o$. The distance from the mirror to the drivers mouth is approximately 40 cm.

## 4. Directivity Pattern and Gain

We begin our evaluation of the proposed arrays by computing the array gain and the beam pattern using the delay-and-sum and

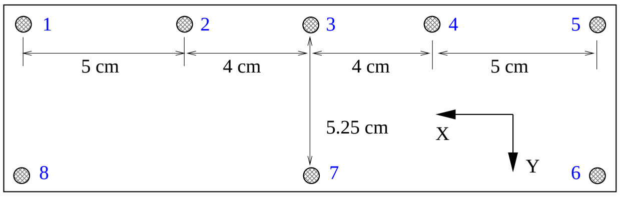


Figure 2: Array dimensions and microphone placement.

| name | beamformer type | uses microphones |
|---|---|---|
| G5L-DS | delay-and-sum | 1, 2, 3, 4, 5 |
| G5P-DS | delay-and-sum | 1, 2, 4, 5, 7 |
| G5L | superdirective | 1, 2, 3, 4, 5 |
| G5P | superdirective | 1, 2, 4, 5, 7 |

Table 1: Definition of arrays.

the superdirective approach with limited susceptibility. The susceptibility of the array design was adjusted such that it matches the susceptibility of the experimental setup. This leads to a rather low value $K(\Omega) = 2$. Figure 3 depicts the array gains for the arrays with five microphones. The superdirective designs G5L and G5P are compared to the delay-and-sum beamformers G5L-DS and G5P-DS. The look direction of the array was steered towards the driver at $\theta_0 = 70^o$ and $\phi_0 = 102^o$. From these results we conclude that compared to the linear delay-and-sum beamformer the superdirective approach improves the results by a maximum of 2 dB. We also note that the G5P design achieves a higher gain than the linear array G5L for all frequencies.

Figures 4, 5, and 6 show the beam pattern for the arrays G5P-DS, G5L, and G5P, respectively, as a function of frequency and elevation $\theta$. We note that both superdirective designs achieve approximately the same beam width with less variation than the delay-and-sum beamformer.

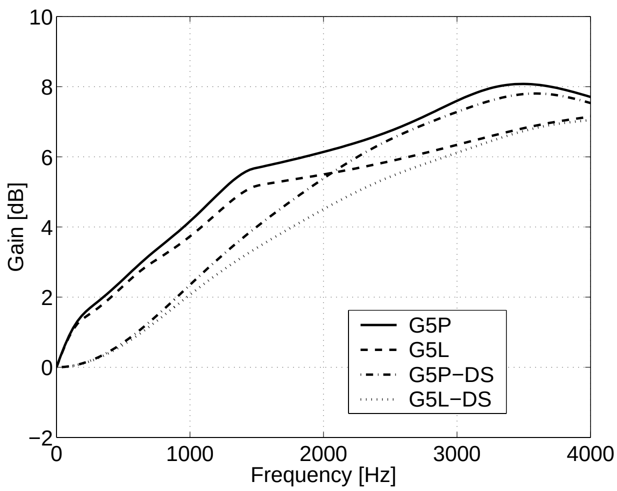


Figure 3: Gain of delay-and-sum arrays G5L-DS (dotted) and G5P-DS (dash-dotted) and of superdirective arrays G5L (dashed) and G5P (solid) in a diffuse sound field ($\theta_0 = 70^o$, $\phi_0 = 102^o$).

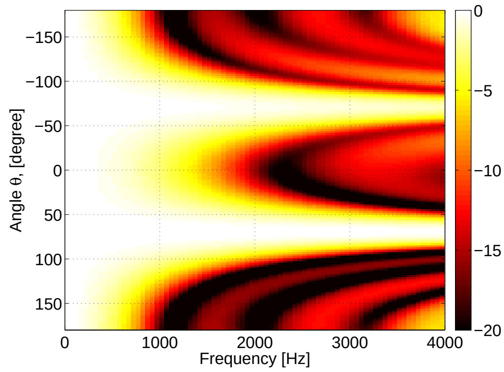


Figure 4: Beam pattern of microphone array G5P-DS.

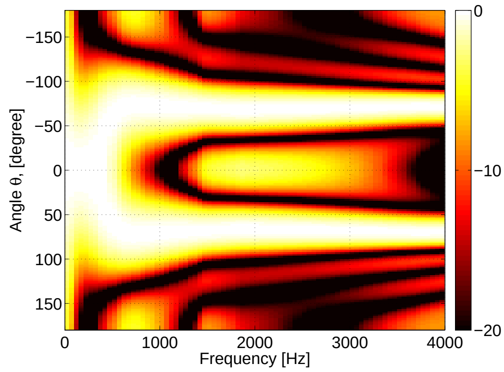


Figure 5: Beam pattern of microphone array G5L.

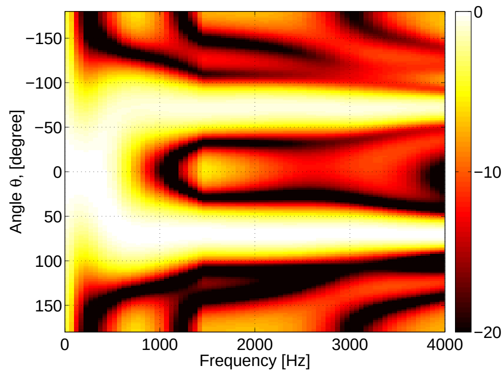


Figure 6: Beam pattern of microphone array G5P.

## 5. Experimental Results

For all our experiments, we used a Volkswagen Golf III vehicle. The array was mounted at the rear-view mirror. Recordings were made at standstill and while driving at 50 km/h with an 8 channel, 20 bit ADAT audio recorder. Prior to A/D conversion, the signals were highpass filtered to attenuate noise below 150 Hz. The digitized audio data was downsampled from 48 kHz to 8 kHz and was processed on a personal computer. We evaluated the actual gain of the arrays by computing the SNR improvement with respect to a single microphone of the array. For computing the SNR improvement we used real speech signals recorded during standstill and noise signals recorded while driving. The reference microphone for the gain computation was microphone #1 (see Fig. 2) which was the microphone with the largest SNR among the microphones of the array. Therefore, the results below are conservative gain estimates.

Figures 7, 8, and 9 plot the theoretical gain and the measured gain for the arrays G5P-DS, G5L, and G5P, respectively. The large fluctuations observed in the gain measurements are due to the geometry of the array environment, reverberation,

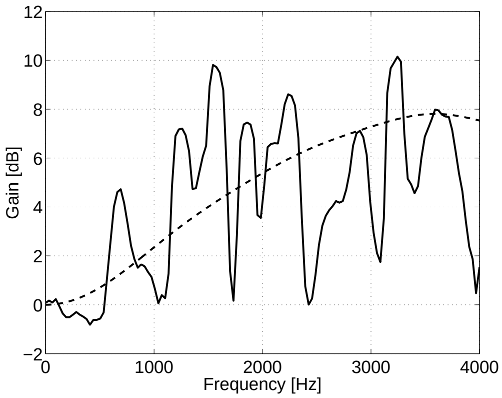


Figure 7: Theoretical (dashed) and measured (solid) gain of planar microphone array G5P-DS in a car.

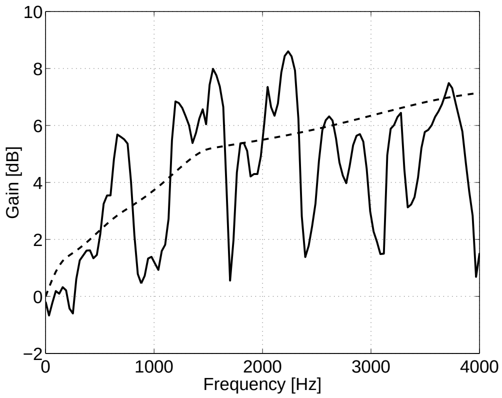


Figure 8: Theoretical (dashed) and measured (solid) gain of linear microphone array G5L in a car.

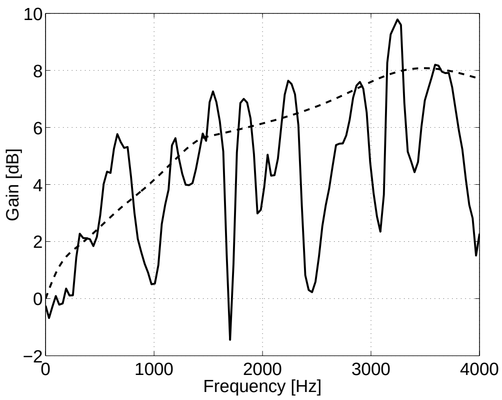


Figure 9: Theoretical (dashed) and measured (solid) gain of planar microphone array G5P.

and the measurement signals. The average gain is 4.8 dB for the array G5L and 5.1 dB for the planar design G5P. Interestingly enough, the planar delay-and-sum beamformer G5P-DS achieved the highest average gain (5.3 dB). The output signal of array G5P-DS, however, sounded much worse than the output of array G5P since low frequency noise is not adequately suppressed. Informal listening did not reveal significant differences between the output signals of the superdirective linear and the superdirective planar array design. The output signals of all arrays appear to be significantly less reverberant.

## 6. Conclusions

The gain computations indicate that the planar array slightly outperforms linear arrays in our speech acquisition task. Even with low-cost omnidirectional microphones and a rather low limit on the susceptibility ($K(\Omega) = 2$) the superdirective design achieves an audibly higher gain than the delay-and-sum beamformer at low frequencies. In a listening test no significant difference between the planar and the linear array designs could be observed. We also found that given the space restrictions and the limited number of microphones it was not possible to achieve the same performance as a well placed fixed directional microphone (e.g, a Primo EMU 4747). To further improve the array gain the array size must be increased.

## 7. References

[1] E. Gilbert and S. Morgan, "Optimum Design of Directive Antenna Arrays Subject to Random Variations," *Bell Syst. Tech. J.*, pp. 637–663, May 1955.

[2] H. Cox, R. Zeskind, and T. Kooij, "Practical Supergain," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 34, pp. 393–398, June 1986.

[3] M. Dörbecker, *Multi-channel Algorithms for the Enhancement of Noisy Speech for Hearing Aids (in German)* . PhD thesis, Aachen University of Technology (P. Vary, ed.), 1998.

[4] D. Van Compernolle, W. Ma, and M. Van Diest, "Speech Recognition in Noisy Environments with the Aid of Microphone Arrays," in *Proc. EUROSPEECH*, vol. 2, pp. 657–660, 1989.

[5] S. Oh, V. Viswanathan, and P. Papamichalis, "Hands-free Voice Communication in an Automobile with a Microphone Array," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, vol. 1, pp. 281–284, 1992.

[6] G. Elko, "Microphone Array Systems for Hands-Free Telecommunication," in *Proc. Intl. Workshop Acoustic Echo Control and Noise Reduction*, pp. 31–38, 1995.

[7] M. Dahl and I. Claesson, "Acoustic Noise and Echo Cancelling with Microphone Array," *IEEE Trans. Vehicular Technology*, vol. 48, no. 5, pp. 1518–1526, 1999.

[8] Y. Grenier, "A Microphone Array for Car Environments," in *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, vol. 1, pp. 305–308, 1992.

[9] S. Nordholm, I. Claesson, and B. Bengtsson, "Adaptive Array Noise Suppression of Handsfree Speaker Input in Cars," *IEEE Trans. Vehicular Technology*, vol. 42, pp. 514–518, Nov. 1993.

[10] M. Dörbecker, "Small Microphone Arrays with Optimized Directivity for Speech Enhancement," in *Proc. EUROSPEECH*, pp. 327–330, 1997.