Real-time steerable directional sound sources

Arthur Berkhoff[1,2], Raymond van der Rots[2]

[1] *TNO Technical Sciences, Acoustics and Sonar, The Hague, Netherlands, Email: arthur.berkhoff@tno.nl*

[2] *University of Twente, Faculty EEMCS, Netherlands, Email: a.p.berkhoff@utwente.nl*

Warning signal generators in cars are being developed that have the objective to warn vulnerable road users. Several sensor systems exist which are able to reveal the position of the vulnerable road users, which information can be used by the warning signal generator. Based on this information, the signal generator can be designed to generate the specifed warning signal at the location of the vulnerable road user while the acoustic response at other locations is minimized in order to reduce noise pollution. The directional sound beam was realized with an array of controlled acoustic sources. Changes of the relative positions of the vehicle and the vulnerable road user require continuous adjustments of the sound beam. Different methods to generate the sound beam are described and experimental results are shown. Rapid real-time adjustment of the beaming direction is possible even if the beamformer is recomputed everytime a new steering direction is required.

Introduction

Studies [1, 2] have shown that (hybrid) electrical vehicles pose an increased risk to pedestrians and bicyclists when compared to traditional vehicles with combustion engines. This increased risk is due to the low noise production at slower speeds, where tire noise is not yet dominant. These speeds are typically below 30 km/h. To improve safety, a sound source can be used to improve detection of electrical vehicles if required. In this paper, it is suggested to make use of a directional sound beam in order to reduce the noise pollution. This directional beam is designed to be able to produce a specified acoustic response in a given target direction whilst minimizing the response in the other directions. The algorithms that are able to create such a beam through producing filters for use with transducer arrays are commonly referred to as beamformers.

Much work on beamforming has been done in the electromagnetic domain, using antenna arrays to receive specific frequencies from specific directions. In [3], many of these earlier approaches are summarised and referenced. Later on, this concept was expanded for the use of microphone arrays of which the book [4] provides many approaches and serves as an important background for beamforming in general. Further work includes automatic steering of the array as in [5]. Due to the reciprocity theorem, the algorithms that are designed for these receiving arrays can be applied to transmitting arrays of point-sources as well. For source arrays the amount of radiated power is often critical and therefore these sources can not be considered to be small. Furthermore, in some configurations there is a strong coupling between the sources. Moreover, the radiation geometry can be complex and there can be differences between the sources as well. Therefore, the methods that are considered are all based on measured transfer functions.

The most simple method of beamforming is commonly known as delay-and-sum [4, 6]. This method is designed to compensate for the different phase differences between the sources and the bright spot in order to ensure that they are constructive. Additional compensation for different, possibly frequency dependent source sensitivities to each target sensor can be included. An approach which is specifically designed for transmitting type arrays is the contrast control approach [7, 8]. A variant known as acoustic energy difference was later proposed in [9]. These methods optimize the contrast between the bright and dark zones. The least-squares algorithm described in [4, 6] aims to find a sound pressure field that matches the desired field. In [10], a frequency-invariant approach is introduced for an omnidirectional receiving array.

A method for loudspeaker arrays which constrains the response at certain sensors, minimizes the response at other sensors, and which can be used with arbitrary transfer functions is described in Ref. [11]. This algorithm will be called the sound power minimization method. Yet another approach to creating a directional source was introduced in [12] under the name of the time-reversal approach. This algorithm reverses the impulse response from the loudspeaker to the focus point, effectively using the reflections that are present to focus the sound. As also noted in [13], the performance greatly depends on the surrounding which makes this algorithm impractical for dynamic environments. A framework for comparison of regularisation methods is presented by Elliott [14].

Methods

Let us define N loudspeakers at positions $r = \begin{bmatrix} r_1 & r_2 & \dots & r_N \end{bmatrix}$ and the vector $G(x)$ containing the transfer functions from the loudspeaker positions r to the receiver point x:

$$G(x) = \begin{bmatrix} G(r_1|x) & G(r_2|x) & \dots & G(r_N|x) \end{bmatrix}. \quad (1)$$

The source strength vector $q = \begin{bmatrix} q_1 & q_2 & \dots & q_N \end{bmatrix}^\top$ is defined to describe the input to each source. The pressure in a point x is therefore given by

$$p(x) = G(x)q. \quad (2)$$

The acoustic potential energy e_V in a region V is related to the mean magnitude squared value of the pressure integrated over the volume of V:

$$e_V = \frac{1}{V} \iiint_V p(x)^* p(x)\,\mathrm{d}x, \quad (3)$$

in which $*$ denotes the complex conjugate.

In the discrete space domain, e_V can be approximated by sampling at sufficiently small intervals. This then leads to the following matrix-vector representation:

$$\begin{aligned} e_V &\approx \frac{1}{M} \sum_{m=1}^{M} p(x_m)^* p(x_m) \\ &= q^H R_V q. \end{aligned} \tag{4}$$

with

$$R_V = \frac{1}{M} \sum_{m=1}^{M} G(x_m)^H G(x_m). \tag{5}$$

Here, the matrix R_V can be interpreted as the spatially averaged correlation matrix of transfer functions defined on the sampled points of the sampled acoustic region V.

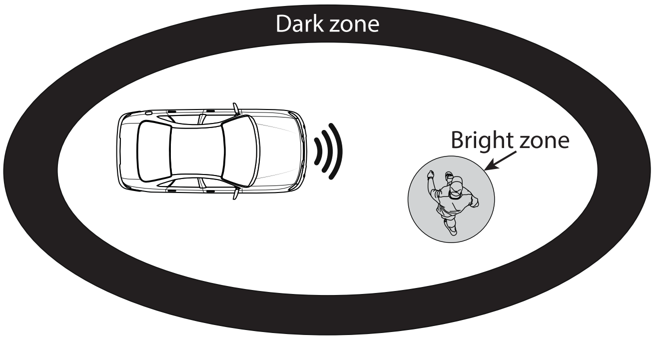

Figure 1: Illustration of an acoustic bright zone and an acoustic dark zone.

As illustrated in Fig. 1, two distinct types of regions are defined: a bright and a dark region. Furthermore, the total region is defined as the combination of both bright and dark zones. These regions will be abbreviated by subscripting b (bright), d (dark) or t (total). Then let e_b, e_d and e_t denote the energy of these regions and let R_b, R_d and R_t denote their correlation matrices of transfer functions (as defined in Eq. (5)).

The resulting algorithms for calculating the source strength vectors are as follows:

1. Delay and sum

$$q = \frac{G^*}{G^H G} \tag{6}$$

2. Contrast control

$$J_{\text{cc}} = \frac{e_b}{e_t} = \frac{q^H R_b q}{q^H \left(R_t + \beta I\right) q}. \tag{7}$$

Here, q is the eigenvector belonging to the maximum eigenvalue of $\left(R_t + \beta I\right)^{-1} R_b$.

3. Acoustic energy difference

$$J_{\text{aed}} = \frac{e_b - \alpha e_d}{e_q} = \frac{q^H \left(R_b - \alpha R_d\right) q}{q^H q}. \tag{8}$$

In this case, q is the eigenvector belonging to the maximum eigenvalue of $R_b - \alpha R_d$. Note that no regularization factor β is required since no matrix inversion is performed, but an eigenvalue computation instead.

4. Least squares

$$q = \hat{G}^{\dagger} p, \tag{9}$$

in which $\dagger$ denotes the pseudo-inverse.

5. Sound power minimisation

$$J_{\text{spm}} = q^H R_t q + \beta q^H q + f(\lambda, c), \tag{10}$$

in which λ is a Lagrange multiplier and c is a constraint.

For a quantitative and objective performance comparison, different measures were used. These are the directivity, also commonly referred to as acoustic contrast or signal to noise ratio, efficiency, also commonly referred to as sensitivity and white-noise gain, consistency and beam-width.

The directivity is defined as:

$$\begin{aligned} Q(\omega) &= 10 \log \frac{M |p_x(\omega)|^2}{\sum_{i=1}^{M} |p_i(\omega)|^2} \\ &= 10 \log \frac{e_b(\omega)}{e_t(\omega)} \quad \text{as } \dim R_b = 1 \times 1 \; \forall \omega, \end{aligned} \tag{11}$$

where p_x is the pressure at the target point and $p_1 \ldots p_M$ are the pressures in the entire region of interest. The directivity describes the ratio between the sound pressure at the target point against the sound pressure in the area of interest and is therefore a measure of how well the sound is directed towards the target point. Some other criteria to evaluate the performance were used as well. The efficiency is defined as:

$$\begin{aligned} \eta(\omega) = \frac{|p_x(\omega)|^2}{q(\omega)^H q(\omega)} &= \frac{e_b(\omega)}{q(\omega)^H q(\omega)} \\ \text{as } \dim R_b &= 1 \times 1 \; \forall \omega, \end{aligned} \tag{12}$$

which weighs the pressure in the target area against the amount of control energy that is required.

The consistency is defined as:

$$\sigma = \sqrt{\sum_{\omega=\omega_U}^{\omega_L} |p_x(\omega) - p_c|^2}, \tag{13}$$

which is a measure for the deviation from the target pressure. Note that this measure does not depend on frequency.

The beam width is defined as in equation (14) and represents the amount of area which is perceived as more than half as loud as in the focus point.

$$\begin{aligned} W(\omega) &= \frac{\sum_{i=0}^{M} \mu(i, \omega)}{M}, \\ \mu(i, \omega) &= \begin{cases} 1 & \text{when } 10 \log \frac{|p_i(\omega)|^2}{|p_c(\omega)|^2} \geq -10 \\ 0 & \text{otherwise} \end{cases}. \end{aligned} \tag{14}$$

Results

Experiments were performed to find the most suitable algorithm for use with this application. In these experiments, the acoustic performance was evaluated, as well as their stability and suitability for real-time implementation.

Three sets of experiments were defined with increasing order of realism: 1) simulations using ideal point sources in free-field conditions, 2) adding a fully reflective ground surface and 3) using measured transfer functions and a real-time implementation in the field.

Implementation

All methods as decribed above were implemented and evaluated in real-time. Experiments were performed to validate the simulation results. Measurements were performed in free-field, on asphalt using an 8-element uniform array designed for a maximum frequency of 3kHz, using a real-time implementation with Matlab/Simulink based on hardware as described in [15]. The sources were standard moving coil loudspeakers of nominal 5cm diameter, mounted in a closed box with a spacing of 5.8 cm between the centers of the sources. The height of the box above the asphalt was 60cm. The height of the microphones above the asphalt was 1.75 m. Measured transfer functions were used to determine the frequency domain control coefficients using a sampling frequency of 6 kHz. The transfer functions were obtained between the 8 sources and 15 microphones at a distance of 5 m from the center of the source array.

The Sound Power Minimization approach has the advantage that its implementation is relatively flexible. New beaming directions can be computed very efficiently based on stored transfer functions. In principle, for all methods the required beaming directions can be computed in advance and can be stored in memory if the number of possible beams is limited. In order to obtain smooth transitions between different beaming directions, a relatively dense interpolation grid may be required, leading to a large number of controller solutions. Furthermore, if several beams have to be produced at the same time without knowing in advance which combinations are required then the number of controller coefficients will increase even more. Direct computation of the controller coefficients once a request for a different beam is received is a solution providing more flexibility at a reasonable computational cost and less memory requirements. The control coefficients can be computed efficiently for each new beaming angle with the sound power minimization strategy. The size of the matrix inversion that needs to be recomputed depends on the number of the acoustic constraints, which is usually low. Other matrix inverses in this method are more complex, but do not have to be recomputed and therefore can be stored. Therefore, in the sound power minimization method this computation can be implemented efficiently, providing a flexible solution for generation of one or more beams in the directions specified by the sensor system.

Experiments

For all methods, it was found that there was agreement between the simulations and the real-time implementation. The Acoustic Energy Difference method and the Contrast Control method led to a reduction of the signal output if the coherence was significantly lower than unity, such as near the dips in the response caused by the ground reflection. The Sound Power Minimization approach sustained the beam at more frequencies than the Acoustic Energy Difference method and the Contrast Control method. The beam of the Sound Power Minimization method was found to be slightly more narrow than the beam of the Acoustic Energy Difference mthod. The Acoustic energy difference method had somewhat lower sidelobes than the Sound Power Minimization, but precise tuning was required through the parameter α. The beam of Delay and Sum is relatively wide, especially at low frequecies. For arrays with a relatively small size expressed in wavelengths such as considered in this application, the beams produced by Delay and Sum are found to be too wide. The Sound Power Minimization method and the Acoustic Energy difference method were found to give the best overall acoustic performance for this application. Figure 2 shows the resulting beamforming result for the Sound Power Minimization method. The dips at certain frequencies due to the ground reflection are clearly visible. In addition to the real-time implementation on the Matlab-Simulink system, the sound power minimization method was implemented on an embedded SHARC DSP based platform. On this embedded platform a new beam could be computed and produced within 30ms.

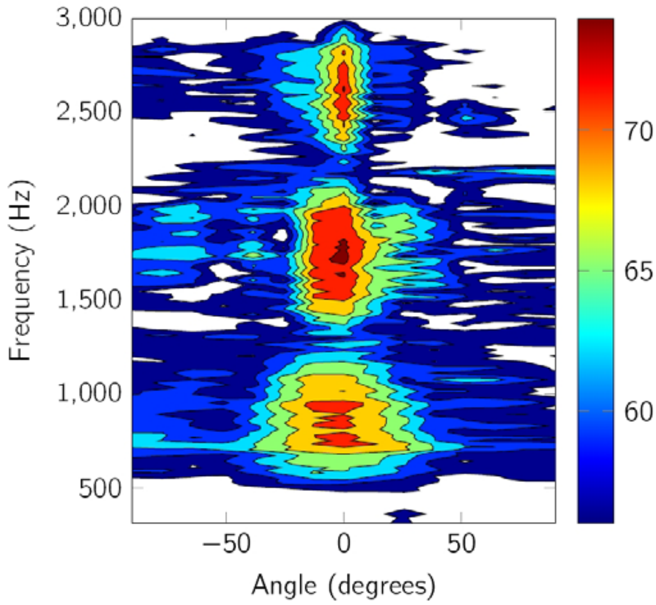

Figure 2: Real-time beamforming result in dB re 20 μPa at 5m distance for the Sound Power Minimization method.

Acknowledgement

Part of this work was supported by the European Commission DG RTD in the 7th Framework Programme, Theme 7 Transport - SST, SST.2011.RTD-1 GA No. 285095, project Electric Vehicle Alert for Detection and Emergency Response (eVADER).

References

[1] Ryan L Robart and Lawrence D Rosenblum. Are hybrid cars too quiet? *Journal of the Acoustical Society of America*, 125(4):2744, 2009.

[2] Refaat Hanna. Incidence of Pedestrian and Bicyclist Crashes by Hybrid Electric Passenger Vehicles. Tech-

nical Report September, NHTSA, 2009.

[3] BD Van Veen. Beamforming: A versatile approach to spatial filtering. *ASSP Magazine, IEEE*, 1988.

[4] H Van Trees. Optimum array processing. 2002.

[5] Francisco Pinto and Martin Vetterli. *Near-field adaptive beamforming and source localization in the spacetime frequency domain*. IEEE, 2010.

[6] Markus Guldenschuh. Transaural stereo in a beamforming approach. *Proceedings of the 12th Int. Conference on Digital Audio Effects (DAFx-09)*, (7):10–15, 2009.

[7] Joung-Woo Choi and Yang-Hann Kim. Generation of an acoustically bright zone with an illuminated region using multiple sources. *Journal of the Acoustical Society of America*, 111(4):1695–1700, 2002.

[8] Ji-Ho Chang, Chan-Hui Lee, Jin-Young Park, and Yang-Hann Kim. A realization of sound focused personal audio system using acoustic contrast control. *Journal of the Acoustical Society of America*, 125(4):2091–2097, 2009.

[9] Mincheol Shin, Sung Q Lee, Filippo M Fazi, Philip a Nelson, Daesung Kim, Semyung Wang, Kang Ho Park, and Jeongil Seo. Maximization of acoustic energy difference between two spaces. *The Journal of the Acoustical Society of America*, 128(1):121–31, July 2010.

[10] L.C. Parra. Least squares frequency-invariant beamforming. *Applications of Signal Processing to Audio and*, (7):102–105, 2005.

[11] A.P. Berkhoff. Lecture notes Signal Processing in Acoustics and Audio (191210960), University of Twente. 2011.

[12] Sylvain Yon, Mickael Tanter, and Mathias Fink. Sound focusing in rooms: the time-reversal approach. *Journal of the Acoustical Society of America*, 113(3):1533–1543, 2003.

[13] Markus Guldenschuh, Alois Sontacchi, and Franz Zotter. Principles and considerations to controllable focused sound source reproduction. *7th Eurocontrol INO*, 2008.

[14] Stephen J Elliott, Jordan Cheer, Jung-woo Choi, and Youngtae Kim. Robustness and Regularization of Personal Audio Systems. *IEEE Transactions on Audio, Speech and Language processing*, 20(7):2123–2133, 2012.

[15] J. M. Wesselink and A. P. Berkhoff. Fast affine projections and the regularized modified filtered-error algorithm in multichannel active noise control. *J Acoust Soc Am*, 124:949–960, 2008.