SCORE: A LOW COMPLEXITY, ROBUST ALGORITHM FOR THE DETECTION OF CORRUPT SENSORS AND SELF-CALIBRATION OF MICROPHONE ARRAYS

Nilesh Madhu and Rainer Martin

{firstname}.{lastname}@rub.de
Institute of Communication Acoustics (IKA), Ruhr-Universität Bochum
44780 Bochum, Germany.

ABSTRACT

This article details a simple method developed to detect sensor degradation in microphone arrays and to equalize the microphone gains. The described method is online in nature and based upon the long-term power spectra of the microphone signals of each array. For a multi-array system, the calculations are done separately for each array, the close proximity of the microphones in an array being necessary for the application of the detection and calibration methods. The approach is tested on three sample degradation scenarios: complete sensor failure, non-linearity in the input, and parasitic harmonic injection. The results indicate that the developed approach is capable of detecting such failures. The calibration factors computed using the proposed approach are similar to that obtained by established methods.

***Index Terms*—** Microphone arrays, self-calibration, gain correction

1. INTRODUCTION

When using arrays for the localization and enhancement of acoustic sources, the proper functioning of each channel of the array is implicitly assumed. However, in applications where the arrays are deployed in adverse environments, and for long periods of time, this assumption needs to be verified periodically. Furthermore, the underlying matched-microphone characteristics assumption of most multi-channel signal processing algorithms is not guaranteed due to the manufacturing tolerances of the microphones, and that of the associated amplifier and sampling components. Such mismatches lead to a degradation in algorithm performance.

The problem of calibration has been well studied in literature. One approach is to calibrate the microphone arrays using known calibration sources, either in anechoic environments or *in situ* [1]. The calibration filters are then held constant during the operational phase. Periodic recalibration may be done to correct degradations due to changes in environment or age related changes. If done properly, such approaches are able to correct both gain and phase mismatches and considerably improve the performance of the localization/enhancement algorithms.

Alternatively, in [2], the calibration filters are computed *online* in a generalized sidelobe canceller (GSC) structure, using the enhanced signal after the fixed beamformer stage for better SNR conditions. The direction of the desired signal is assumed known and the calibration is done in the presence of the target signal. Similar structures are proposed in [3]. These approaches are better suited to dynamic environments – where a periodic recalibration using known signals is impractical

In [4] it is argued that the problem arising from mismatch is more due to the gain imbalance. Thus, to improve performance, it is sufficient to compute only a *scalar*, the gain correction factor, instead of a *filter*. This is also the focus of [5]. The difference between the two approaches lies mainly in the adaptation of the gain correction factor. Whereas [4] adapts the gain factors in the presence of a *single* dominant source with a strong direct path, [5] models the gains over a long-term average of the signal powers with the assumption that, over a long observation period, the average power received by all the microphones must be the same in the absence of mismatch.

The approaches summarized above, however, assume that all microphone channels are functional. While detection of sensor failure can be done by manually disassembling and re-calibrating the arrays, this procedure is not only time-consuming, but also impractical. Online algorithms for sensor degradation detection are thus necessary, and such an algorithm is the focus of this contribution.

The rest of the article is organized as follows. First, for clarity, we shall introduce the notations we adopt throughout the text. Next, we describe the system model and enumerate our key assumptions and simplifications. Based upon this model, Section 4 presents a detailed discussion of the proposed online approach: the spectral correlation (SCORE) approach.

2. NOTATIONS

The discrete-time domain signal at any sensor m is denoted as $x_m(n)$, where n denotes the discrete-time index. The corresponding K-point discrete Fourier transform[1] of this signal segment is denoted as:

$$\mathrm{DFT}\left\{x_m(n)\right\} = X_m\left(k\right) \quad k = \{0, \ldots, K\}\,. \tag{1}$$

Further, stacking the time domain signals of all the channels at time n as:

$$\mathbf{x}(n) = \left(x_1(n)\,,\, x_2(n)\,,\, \ldots\,,\, x_M(n)\right)^T\,, \tag{2}$$

we obtain an M-dimensional vector for that time instant. We can write a similar expression for each frequency *bin* as:

$$\mathbf{X}\left(k\right) = \left(X_1\left(k\right)\,,\, X_2\left(k\right)\,,\, \ldots\,,\, X_M\left(k\right)\right)^T\,, \tag{3}$$

where $X_m\left(k\right)$ is the DFT coefficient for channel m and bin k.

[1] In practice, we perform frame-wise operations upon the input signals. Therefore, a more accurate expression would be:

$$\mathrm{DFT}\left\{x_m(bO+n)\right\} = X_m(k,b)\,,$$

where $O \in \mathbb{N}$ indicates the frame shift (in samples) between the frames and b, the frame index. The frame index shall be dropped for convenience and re-introduced where required.

3. SYSTEM MODEL

We consider the general case of an array of M channels and Q sources distributed in the environment of the array. Q is any unknown number, inessential to the approach. The general signal model we shall assume subsequently is:

$$\mathbf{X}(k) = \mathbf{A} \odot \left(\sum_{q=1}^{Q} \mathbf{S}_q(k) \right) + \mathbf{V}(k), \tag{4}$$

where $\odot$ represents the element-wise (Hadamard) product, $\mathbf{S}_q(k) = (S_{q1}(k), S_{q2}(k), \ldots, S_{qM}(k))^T \in \mathbb{C}^{M \times 1}$ represents the part of the input that is dependent upon the signals in the environment, and received at all the channels (the propagation effects are implicitly included in the definition); $\mathbf{V}(k) = (V_1(k), \ldots, V_M(k))^T$ represents the part of the input that contains self-induced noise and $\mathbf{A} = (A_1, \ldots, A_M)^T$ indicates the calibration factor where, following [4], they are assumed to be purely real and *independent* of frequency. The microphone signals are assumed to be zero-mean and the Q sources statistically independent of one another and of the self-induced noise.

Our approach is based on assumptions similar to that in [5]: we principally assume that each microphone receives the same power on an average. This is a justifiable assumption when the microphones of an array are placed close together. Therefore, for microphones that are properly calibrated, the values for the A_m would be around 1, whereas, when sensor degradations occur, the values diverge.

As a note, under some acoustic conditions, e.g., when the array is in a standing wave field, the assumption of equal average power may be grossly violated, leading to strong power variations across the microphones of the array. These variations do not reflect the mismatch being modelled. Consequently, adapting the gain functions and/or detecting degradation etc. leads to biased results in such conditions. These fields are usually generated by strongly directive sources, and in such cases it might be advisable to allow the algorithm to be active only during periods where these sources are *absent*. The detection of such directive sources may be accomplished, for example, as in [6, 7, 8].

4. THE SCORE APPROACH

For a healthy array, a sufficient condition for equal average power at each sensor is that the average spectral power of each sensor should be similar. For a microphone m, the power spectrum is computed by an averaging of the instantaneous power, $\mathcal{E}_m(k,b) = |X_m(k,b)|^2$, of the DFT coefficients over multiple frames B (temporal averaging). We define this value as:

$$\overline{\mathcal{E}}_m(k) = \frac{1}{B} \sum_b \mathcal{E}_m(k,b) \tag{5}$$

To detect sensor degradations, we treat the power spectrum of each channel as a realization of a random process and compute the *correlations* between the spectral powers of the channels. Thus, for each channel pair (m, m'), we obtain the corresponding correlation coefficient [9] $\Gamma_{mm'}$ as follows:

$$\widehat{\mu}_m = \frac{1}{K} \sum_{k=1}^{K} \overline{\mathcal{E}}_m(k), \tag{6}$$

$$\Gamma_{mm'} = \frac{\sum_k \left(\overline{\mathcal{E}}_m(k) - \widehat{\mu}_m \right) \left(\overline{\mathcal{E}}_{m'}(k) - \widehat{\mu}_{m'} \right)}{\sqrt{\sum_k \left(\overline{\mathcal{E}}_m(k) - \widehat{\mu}_m \right)^2 \sum_k \left(\overline{\mathcal{E}}_{m'}(k) - \widehat{\mu}_{m'} \right)^2}}, \tag{7}$$

where K is the length of the (discrete) Fourier transform. Stacking the $\Gamma_{mm'}$ according to the indices, we obtain the corresponding correlation matrix $\mathbf{\Gamma}$ for the array under consideration. Fig. 1 shows this correlation matrix $\mathbf{\Gamma}$ for a sample 8 channel array.

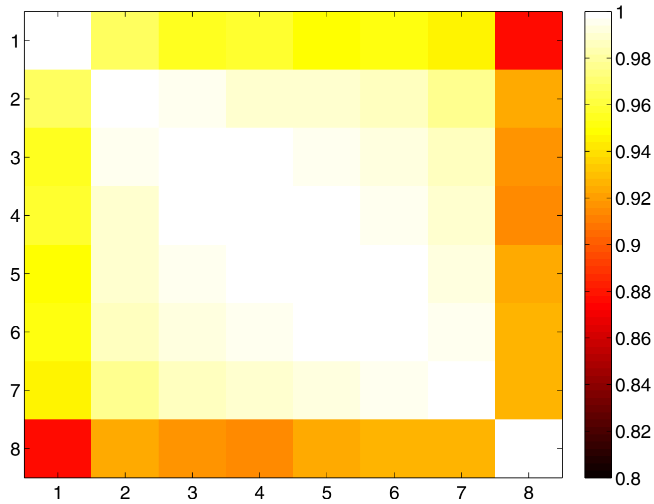

Fig. 1. Power spectrum correlation for a healthy array. Note the high values of the correlation coefficients.

The advantage of using the correlation coefficients as defined in (7) becomes evident when we consider a noise floor that is spectrally flat (which is a realistic assumption for the sensor noise). For this case, taking the statistical expectation of the instantaneous power, we obtain from (4) and the definition of $\mathcal{E}_m(k,b)$:

$$\begin{aligned} \mathrm{E}\{\mathcal{E}_m(k,b)\} &= \mathrm{E}\left\{|X_m(k,b)|^2\right\} \\ &= |A_m|^2 \sum_{q=1}^{Q} \Psi_{S_{qm}S_{qm}}(k) + \Psi_{V_m V_m}(k) \\ &= |A_m|^2 \Psi_{SS}(k) + \Psi_{V_m V_m}, \end{aligned} \tag{8}$$

where $\Psi_{SS}(k) = \sum_{q=1}^{Q} \Psi_{S_{qm}S_{qm}}(k)$ and $\Psi_{S_{qm}S_{qm}}(k)$ is the power spectral density of the qth source at sensor m, for bin k. Note that the equal average power assumption at each channel implies that $\Psi_{SS}(k)$ is the same for all the channels – indicated by dropping the channel index m in (8). Note, also, that the frequency index has been dropped for the noise because it is assumed to have a flat spectrum. Consequently,

$$\begin{aligned} \mu_m &= \frac{1}{K} \sum_{k=1}^{K} \mathrm{E}\{\mathcal{E}_m(k)\} \\ &= |A_m|^2 \frac{1}{K} \sum_{k=1}^{K} \Psi_{SS}(k) + \Psi_{V_m V_m}. \end{aligned} \tag{9}$$

We now see that computing the correlation as in (7) yields a value that is not sensitive to white noise and dependent upon the incident signals only, thus avoiding any bias in the correlation function.

4.1. Determining the corrupted channels from $\mathbf{\Gamma}$

As shown in Fig. 1, the correlation between the power spectra is high when the channels function properly – illustrated below for an

$M = 8$ sensor array:

$$\mathbf{\Gamma} = \begin{pmatrix} 1.000 & 0.970 & 0.963 & 0.964 & 0.964 & 0.962 & 0.953 & 0.943 \\ 0.970 & 1.000 & 0.992 & 0.986 & 0.982 & 0.969 & 0.942 & 0.914 \\ 0.963 & 0.992 & 1.000 & 0.997 & 0.994 & 0.983 & 0.960 & 0.929 \\ 0.964 & 0.986 & 0.997 & 1.000 & 0.999 & 0.992 & 0.972 & 0.943 \\ 0.964 & 0.982 & 0.994 & 0.999 & 1.000 & 0.996 & 0.979 & 0.950 \\ 0.962 & 0.969 & 0.983 & 0.992 & 0.996 & 1.000 & 0.989 & 0.964 \\ 0.953 & 0.942 & 0.960 & 0.972 & 0.979 & 0.989 & 1.000 & 0.983 \\ 0.943 & 0.914 & 0.929 & 0.943 & 0.950 & 0.964 & 0.983 & 1.000 \end{pmatrix}$$

Summing the $\Gamma_{mm'}$ values along the columns then yields values close to M:

$$\sum_{m'=1}^{8} \Gamma_{mm'} = \big(7.722,\ 7.758,\ 7.820,\ 7.8556,\ 7.865,\ 7.858,\ 7.782,\ 7.628\big)^T .$$

However, when one or more channels are degraded, this correlation goes down, as illustrated below for the case where sensor 1 is degraded.

$$|\mathbf{\Gamma}| = \begin{pmatrix} 1.000 & 0.026 & 0.029 & 0.033 & 0.032 & 0.033 & 0.026 & 0.012 \\ 0.026 & 1.000 & 0.992 & 0.984 & 0.981 & 0.975 & 0.962 & 0.946 \\ 0.029 & 0.992 & 1.000 & 0.997 & 0.995 & 0.988 & 0.973 & 0.957 \\ 0.033 & 0.984 & 0.997 & 1.000 & 0.999 & 0.994 & 0.981 & 0.963 \\ 0.032 & 0.981 & 0.995 & 0.999 & 1.000 & 0.997 & 0.985 & 0.966 \\ 0.033 & 0.975 & 0.988 & 0.994 & 0.997 & 1.000 & 0.992 & 0.972 \\ 0.026 & 0.962 & 0.973 & 0.981 & 0.985 & 0.992 & 1.000 & 0.983 \\ 0.012 & 0.946 & 0.957 & 0.963 & 0.966 & 0.972 & 0.983 & 1.000 \end{pmatrix} .$$

The column sum yields, in this case,

$$\sum_{m'=1}^{8} \Gamma_{mm'} = \big(0.805,\ 6.817,\ 6.873,\ 6.886,\ 6.893,\ 6.888,\ 6.852,\ 6.777\big)^T ;$$

i.e, values significantly lower than M, with the first element having the lowest score. Using this observation, we arrive at our procedure, described in Fig. 2. Note that the approach is *iterative*: we discard the defective sensors in the descending order of their magnitude of degradation. This is necessary because it allows for the setting of a uniform degradation threshold. It is not possible to obtain such a uniform threshold value if we decide to pick the degraded sensors in *one* step, on the basis of the column sum.

Further, the limitiation of $M \geq 3$ in step 7 is required in order to have a majority vote. Obviously, given only two sensors with low correlation it is difficult to decide which sensor is defective.

4.2. Determining the calibration factors

Once the degraded channels have been weeded out, the gain calibration can be done for the remaining elements of the array. For this, we first compute the average power at each sensor by summing the power spectrum:

$$\overline{\mathcal{E}}_{m, m \not\subseteq \text{DegradedChannels}} = \sum_{k=1}^{K} \overline{\mathcal{E}}_m(k) \tag{10}$$

Next, the sensor with the maximum power is taken as the reference:

$$\overline{\mathcal{E}}_{\text{ref}} = \max \overline{\mathcal{E}}_{m, m \not\subseteq \text{DegradedChannels}} \tag{11}$$

and the calibration factors may be computed for the remaining channels, with respect to this sensor, as:

$$g_{m, m \not\subseteq \text{DegradedChannels}} = \frac{\overline{\mathcal{E}}_{\text{ref}}}{\overline{\mathcal{E}}_m} . \tag{12}$$

In order to prevent any particular block/time interval from influencing the gain factors, we should use iterative updates for the gains as proposed in [4].

1. DegradedChannels = {}
2. $\forall$ channels m of M, do:
   Calculate the $\overline{\mathcal{E}}_m(k)$ (using (5)).
   End
3. Calculate $\mathbf{\Gamma}$ as in (6) & (7).
4. Let ϱ_0 := allowed degradation in the channels.
5. Calculate the column sum: $\boldsymbol{\gamma} = \sum_{m'=1}^{M} \Gamma_{mm'}$.
6. $\forall$ elements γ_m of $\boldsymbol{\gamma}$, do:
   $$\varrho_m = \frac{\gamma_m}{M}$$
   End
7. If ($\exists\ \varrho_m < \varrho_0$) & ($M \geq 3$), do:
   Find $m_{\text{bad}} = \underset{m}{\operatorname{argmin}}\ \varrho_m$
   DegradedChannels = DegradedChannels $\cup\ m_{\text{bad}}$
   Remove the row and column for channel m_{bad} from $\mathbf{\Gamma}$
   $M \leftarrow M - 1$
   Go to Step 5.
   End

Fig. 2. Algorithm to determine degraded channels using the SCORE approach.

4.3. Additional considerations for the SCORE approach

The above discussion has assumed a stationary power floor so far. However, in practical situations, we need to consider the eventuality that the average signal power could change with time, necessitating some kind of moving average operator similar to (5). There may also be signal segments which cannot be used (as discussed previously). This would lead to breaks in the averaging periods. Consequently, we split the averaging of (5) into two parts: first, for each block o of length T ms that contains mainly diffuse signal components (corresponds to $B \approx (T f_{\text{s}} - K)/O + 1$ frames[2]), the short-term temporal average is computed:

$$\overline{\mathcal{E}}_m^{\,o}(k) = \frac{1}{B} \sum_b \mathcal{E}_m(k, b) . \tag{13}$$

Next, the long term average is computed on the selected blocks in a recursive manner – in order to conserve memory – as:

$$\overline{\mathcal{E}}_m(k) \leftarrow \eta\, \overline{\mathcal{E}}_m^{\,o}(k) + (1 - \eta) \overline{\mathcal{E}}_m(k) , \tag{14}$$

where η is the smoothing factor for the update. Equation (13) provides a robust estimate of the signal power *within* the block o, whereas the second part (equation (14)) adapts the power spectrum to changes in the average signal power with time. As the detection of the anomalous sensors is based on long term averages, it suffices to compute the necessary parameters – the g_m and $\mathbf{\Gamma}$ – once every few *seconds*, reducing the computational load on the processing engine.

5. EVALUATION

The proposed approach was tested on a system designed for the automotive industry, where the focus was on the localization of harmonics. Due to space constraints, only the following results are presented:

a. Complete failure of multiple sensors (such sensors record only noise)

[2] with a frame shift of O samples, sampling rate f_{s}, and K-point DFT

b. One (or more) microphones are nonlinearly corrupted (clipping/rectification)

c. Multiple microphones pick up parasitic sinusoids (corresponding, e.g., to cross-talk with power supplies).

The system parameters used in the experiments were: $K = 1024$, $O = 256$, $f_\mathrm{s} = 32\,\mathrm{kHz}$, $T = 200\,\mathrm{ms}$, and $\varrho_0 = 0.7$. The signal segment considered was 5 s long.

5.1. Complete sensor failure

In this evaluation, a degraded channel is represented by replacing the time domain signal by random, uncorrelated, white noise of unit variance. This would correspond to the case where a sensor is defective to the point where it does not respond to acoustic input. For the correlation matrix we have:

$$|\mathbf{\Gamma}| = \begin{pmatrix} 1.000 & 0.094 & 0.030 & 0.036 & 0.016 & 0.042 & 0.050 & 0.034 \\ 0.094 & 1.000 & 0.030 & 0.034 & 0.043 & 0.041 & 0.042 & 0.022 \\ 0.030 & 0.030 & 1.000 & 0.996 & 0.098 & 0.987 & 0.970 & 0.947 \\ 0.036 & 0.034 & 0.996 & 1.000 & 0.090 & 0.994 & 0.976 & 0.948 \\ 0.016 & 0.043 & 0.098 & 0.090 & 1.000 & 0.090 & 0.096 & 0.080 \\ 0.042 & 0.041 & 0.987 & 0.994 & 0.090 & 1.000 & 0.990 & 0.957 \\ 0.050 & 0.042 & 0.970 & 0.976 & 0.096 & 0.990 & 1.000 & 0.970 \\ 0.034 & 0.022 & 0.947 & 0.948 & 0.080 & 0.957 & 0.970 & 1.000 \end{pmatrix}.$$

Note, the low correlation values for channels 1,2 and 5, clearly delineating them as corrupted channels. Also, computing the average power we obtain:

$$\overline{\mathcal{E}}_m = \big(193.8 \times 10^3,\ 194.6 \times 10^3,\ 1.269,\ 1.359,\ 197 \times 10^3,\ 1.321,\ 1.181,\ 1.161\big),$$

where we see abnormal power levels for channels 1,2 and 5, although, this *alone* is not always indicative of the improper sensor functioning.

5.2. Non-linear degradation

The non-linearity considered here is full wave rectification: $\tilde{x}_m(n) = |x_m(n)|$, i.e., the absolute value of the input signal is taken. For this case, again channels 1,2 and 5 were considered to be degraded. The correlation matrix obtained is:

$$\mathbf{\Gamma} = \begin{pmatrix} 1.000 & 0.885 & 0.376 & 0.384 & 0.902 & 0.401 & 0.443 & 0.499 \\ 0.885 & 1.000 & 0.361 & 0.363 & 0.948 & 0.384 & 0.428 & 0.498 \\ 0.376 & 0.361 & 1.000 & 0.997 & 0.371 & 0.993 & 0.979 & 0.895 \\ 0.384 & 0.363 & 0.997 & 1.000 & 0.376 & 0.996 & 0.981 & 0.895 \\ 0.902 & 0.948 & 0.371 & 0.376 & 1.000 & 0.397 & 0.438 & 0.513 \\ 0.401 & 0.384 & 0.993 & 0.996 & 0.397 & 1.000 & 0.991 & 0.910 \\ 0.443 & 0.428 & 0.979 & 0.981 & 0.438 & 0.991 & 1.000 & 0.941 \\ 0.499 & 0.498 & 0.895 & 0.895 & 0.513 & 0.910 & 0.941 & 1.000 \end{pmatrix},$$

In this case, too, we see that the power spectra of degraded channels show a lower correlation with the power spectra of healthy channels, and the iterative procedure described in Fig. 2 is able to correctly pick out the degraded microphones. The high correlations between pairs (1,2), (1,5) and (2,5) are to be expected, as they suffer from the same degradation.

5.3. Random parasitic sinusoids

For this purpose, a sinusoid of 100 Hz was added to the signals of microphone 1,2 and 5, at 0 dB SNR. We present, here, only the case of a single harmonic as it is more difficult to detect than the case of multiple harmonics. The parasitic sinusoid at each microphone has a random phase offset. The correlation matrix for this case is:

$$\mathbf{\Gamma} = \begin{pmatrix} 1.000 & 0.999 & 0.096 & 0.102 & 0.999 & 0.097 & 0.106 & 0.101 \\ 0.999 & 1.000 & 0.085 & 0.091 & 0.999 & 0.083 & 0.091 & 0.086 \\ 0.096 & 0.085 & 1.000 & 0.996 & 0.080 & 0.987 & 0.970 & 0.947 \\ 0.102 & 0.091 & 0.996 & 1.000 & 0.086 & 0.994 & 0.976 & 0.948 \\ 0.999 & 0.999 & 0.080 & 0.086 & 1.000 & 0.080 & 0.087 & 0.081 \\ 0.097 & 0.083 & 0.987 & 0.994 & 0.080 & 1.000 & 0.990 & 0.957 \\ 0.106 & 0.091 & 0.970 & 0.976 & 0.087 & 0.990 & 1.000 & 0.970 \\ 0.101 & 0.086 & 0.947 & 0.948 & 0.081 & 0.957 & 0.970 & 1.000 \end{pmatrix}$$

Note that the correlations between channel-pairs (1,2), (1,5), and (2,5) are high – which is not surprising since they suffer the same kind of degradation. However, their respective correlation with the other channels is low.

6. CONCLUSIONS

Current algorithms for online self-calibration of microphone arrays do not consider the eventuality of sensor degradation. Accordingly, in this contribution, we have introduced a simple, online approach – SCORE – for the *in situ* detection of corrupt sensors. The merits of the approach were tested under three possible sensor degradation scenarios: complete sensor failure (only noise at the degraded sensor); non-linear distortion of the sensor signals; and parasitic harmonics in the channels. The approach performs well for all these cases of sensor degradation. The proposed method is computationally inexpensive – the averages computed are evaluated only once every few seconds.

In addition, Section 4.2 indicates how the method can also be used for the gain calibration of the 'healthy' microphones of an array. Our approach yields gain compensation factors similar to that of [5] while, at the same time, providing an indication regarding the state of the sensor. In contrast to other state-of-the-art calibration algorithms [2, 3], this algorithm does not need phase equalization or time-delay compensation prior to calibration.

7. REFERENCES

[1] S. Nordholm, I. Claesson, and M. Dahl, "Adaptive microphone array employing calibration signals. an analytical evaluation," *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 3, May 1999.

[2] P. Oak and W. Kellermann, "A calibration algorithm for robust generalized sidelobe cancelling beamformers," in *Proceedings of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2005.

[3] M. Buck, T. Haulick, and H.-J. Pfleiderer, "Self-calibrating microphone arrays for speech signal acquisition: A systematic approach," *EURASIP Journal on Applied Signal Processing*, vol. 86, pp. 1230–1238, June 2006.

[4] I. Tashev, "Gain self-calibration procedure for microphone arrays," in *Proceedings of International Conference for Multimedia and Expo (ICME)*, June 2004.

[5] T. P. Hua, A. Sugiyama, and G. Faucon, "A new self-calibration technique for adaptive microphone arrays," in *Proceedings of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*, 2005.

[6] N. Madhu, D. Oswald, and R. Martin, "Speaker localization: New algorithm and practical aspects," in *Proceedings of the 19th Conference "Elektonische Sprachsignalverarbeitung (ESSV)"*, Frankfurt am Main, Germany, Sept. 2008.

[7] I. Potamitis, "Estimation of speech presence probability in the field of microphone array," *IEEE Signal Processing Letters*, vol. 11, no. 12, pp. 956–959, Dec. 2004.

[8] N. Madhu, R. Martin, H.-W. Rehn, and A. Fischer, "Brake squeal localization," in *Proceedings of the Annual Meeting of the German Acoustical Society (DAGA)*, Mar. 2006.

[9] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill Inc., 3 edition, 1991.