SPATIAL AUDIO ACTIVITY DETECTION FOR HEARING AIDS

Sriram Srinivasan and Kees Janse

Philips Research High Tech Campus 36, Eindhoven, The Netherlands {sriram.srinivasan,kees.janse}@philips.com

ABSTRACT

We present a multi-microphone signal activity detection scheme for hearing aids to differentiate between the periods of activity of desired and interfering sources. The method is designed to provide robust performance in the presence of simultaneously active desired and interfering sources. We exploit knowledge from the hearing aid domain, and the directional processing present in modern hearing aids, to present a framework to design appropriate thresholds for the detection. Experiments confirm robust performance under practical reverberant conditions.

Index Terms— Hearing aids, speech enhancement, array signal processing, speech activity detection, directional processing

1. INTRODUCTION

A robust signal activity detection scheme is critical to control the update of adaptive beamformers and adaptive noise cancellers. When the interfering signal is stationary background noise, a voice activity detector that tracks deviations from long-term spectral floors may be used to control the update of a beamformer [1]. The task is more complex in the context of hearing aids when the interfering signal can also be non-stationary, e.g., speech from another talker, where an approach based on tracking spectral floors fails. Instead, multiple microphones may be employed to exploit the spatial diversity arising from the fact that the desired and interfering signals are usually located at different points in space. However, the small microphone spacing (0.6-1.5 cm) in hearing aids, and the fact that both the desired and interfering signals are located at a distance from the microphones, introduce new challenges.

In the hearing aid scenario, we seek a detection scheme whose output is ideally one of three possible states: the desired signal is dominant and thus a beamformer steered to the desired source may be adapted, the interfering signal is dominant and thus a noise canceller may be adapted, and finally, neither signal is dominant (either both sources are inactive or both are equally strong) and no adaptation is performed. Thus, we are not interested in merely detecting signal activity but wish to detect dominant signal activity.

Conventional detection schemes rely on thresholds against which the estimated signal-to-interference ratio (SIR) is compared to determine whether the desired or interfering signal is dominant [2]. Using inter-microphone correlation between signals observed on a broad-side two element microphone array, a method to detect desired speech in the presence of spatially separated noise is described in [3], but considers a minimum spacing of 7 cm between the microphones. In this paper, we present a detection scheme that accounts for the small microphone spacing in current hearing aids. The algorithm benefits from the differential processing that almost all modern hearing aids perform to provide directionality [4]. We exploit knowledge from the hearing aid domain, specifically the assumption that the desired signal is located in a small neighborhood in front of the end-fire microphone array present on a hearing aid, and that the interfering sources are located in the rear-half plane. The desired and interfering signals are assumed to originate from point sources, and we also include the effect of ambient noise. We provide a framework where the detection thresholds can be theoretically derived based on application specific design parameters.

2. DETECTIONS USING DIRECTIONAL PROCESSING

We first present the motivation behind our detection algorithm by considering a single source in an anechoic environment. We then include the effect of ambient diffuse noise, and finally generalize the discussion to include simultaneously active desired and interfering sources.

2.1. Single source in an anechoic room

Consider a hearing aid with two omni-directional microphones spaced d meters apart in an end-fire configuration (oriented toward 0 degrees). We define the origin O to be the center of the end-fire array and assume that the point sources under consideration are at ear-level. Most modern hearing aids include a directional processing unit that combines the two omni signals to produce a forward facing cardioid response and a backward facing cardioid response [4]. Consider a point source $s(n, \theta)$ located at an angle θ , where θ is the angle between the vertical y-axis and a ray from the origin to the source. The received microphone signal $x_1(n)$ can be written in the noiseless anechoic case as $x_1(n) = s(n, \theta)$. We can then write the output of the forward cardioid as

$$f(n) \approx \frac{1}{2} (1 + \cos \theta) s(n, \theta) \equiv c(\theta) s(n, \theta), \tag{1}$$

where $c(\theta) = \frac{1}{2}(1 + \cos \theta)$ is the response of a cardioid facing 0 degrees to a signal incident on the array at an angle θ , normalized to have unit response to a signal arriving from 0 degrees.

A block diagram of our proposed detection scheme is depicted in Fig. 1. We introduce an adaptive weight w between one of the omnidirectional microphones, and the forward cardioid. The weight w is adapted continuously using e.g., the normalized least mean square (NLMS) algorithm [5, ch. 9], regardless of whether desired or interfering source is active, and minimizes the expected energy of the error signal defined by

$$e(n) = f(n) - wx_1(n).$$
 (2)

The herein disclosed information is secret until 2008-03-30; eyes only for the appointed reviewers of this conference.

The optimal solution for w is given by



Fig. 1. Signal activity detections using directional processing. f(n) and b(n) are the forward and backward facing cardioid signals respectively.

$$w = \frac{\mathrm{E}\{f(n)x_1(n)\}}{\mathrm{E}\{x_1^2(n)\}}.$$
(3)

When only a desired source located in the region $\Psi_s \equiv [2\pi - \psi_s, 0] \cup [0, \psi_s]$ (a small neighborhood around 0 deg., see Fig. 2) is active, from (1) and (3), we have

$$c_s \equiv \frac{1 + \cos\psi_s}{2} \le w \le 1. \tag{4}$$

When only an interferer located in $\Psi_i \equiv \left[\frac{\pi}{2} + \psi_i, \frac{3\pi}{2} - \psi_i\right]$ is active (in the rear half plane), we have

$$0 \le w \le c_i \equiv \frac{1 + \cos(\frac{\pi}{2} + \psi_i)}{2}.$$
 (5)

E.g., we may have $\psi_s = \pi/9$ (20 deg.), which is a typical value in hearing aid applications, and $\psi_i = 0$ to restrict the adaptation of the noise canceller only for interferences located in the rear half plane $\left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$. It is easy to see that (5) holds even in the presence of multiple interferences located within Ψ_i . When only a single source (either



Fig. 2. Origin O is at the center of the end-fire array facing 0 degrees. Desired source located in $\Psi_s \equiv [2\pi - \psi_s, 0] \cup [0, \psi_s]$. Interferer located in $\Psi_i \equiv [\frac{\pi}{2} + \psi_i, \frac{3\pi}{2} - \psi_i]$.

desired or interferer) is active in an anechoic environment, equations (4) and (5) suggest a simple detection scheme to differentiate between a source in the front half plane and an interferer in the rear half plane. A desired source is detected whenever $w \ge c_s$ and an interferer is detected whenever $w \le c_i$. Note that we decouple the two detections. An interferer is not automatically detected when the desired source is not active.

We note that a similar one-tap adaptive weight, but between the forward and backward cardioids, has been used in the context of adaptively placing a null in the rear-half plane [6,7]. A weight β is adapted to minimize the energy of the error signal $f(n) - \beta b(n)$, where b(n) is the response of a backward facing cardioid. The value of β is restricted to $0 \leq \beta \leq 1$ to avoid cancellation of desired sources ($\beta \geq 1$ in the front half plane) and such a scheme may also be used as an activity detector. However, in basing the detections on the value of either w or β , a practical detection scheme needs to first address the following:

- While (4) and (5) (or the equivalent conditions on β) are satisfied in the free field, the effect of ambient diffuse noise needs to be accounted for.
- 2. The case when both desired and interference signals are simultaneously active needs to be addressed. c_s and c_i can no longer be used as detection thresholds and new thresholds T_s and T_i need to be appropriately defined.

2.2. Effect of ambient diffuse noise

We assume that the ambient diffuse noise, can be modelled as a spherically isotropic noise field. In the presence of diffuse noise and a single source $s(n, \theta_s)$ located at θ_s , the microphone signal can be written as

$$x_1(n) = s(n, \theta_s) + v_1(n),$$
 (6)

and we model the diffuse noise $v_1(n)$ as

$$v_1(n) = \int_0^\pi \int_0^{2\pi} v(n,\theta,\phi) \sin\phi \, d\theta \, d\phi, \tag{7}$$

where $v(n, \theta, \phi)$ corresponds to a signal arriving at the array from an azimuth θ and elevation ϕ . We assume $E\{v(n, \theta, \phi)v(n, \tilde{\theta}, \tilde{\phi})\}$ is nonzero only when both $\theta = \tilde{\theta}$ and $\phi = \tilde{\phi}$, and that $E\{s(n, \theta_s)v(n, \theta, \phi)\}=0$, and $E\{v^2(n, \theta, \phi)\}=E\{v^2(n, \tilde{\theta}, \tilde{\phi})\}=\sigma_v^2$, $\forall \theta, \tilde{\theta}, \phi, \tilde{\phi}$, i.e., we model the diffuse noise as the sum of the equal energy signals emitted by an infinite number of independent sources located around the microphone. The response of the forward cardioid in this situation can be written as

$$f(n) = c(\theta_s)s(n,\theta_s) + \int_0^\pi \int_0^{2\pi} v(n,\theta,\phi) \frac{(1+\cos\theta\sin\phi)\sin\phi}{2} \, d\theta \, d\phi.$$
(8)

From (3) and (8), w can be simplified as

$$w = \frac{c_{\theta_s}\sigma_s^2 + 0.5\sigma_{v_1}^2}{\sigma_s^2 + \sigma_{v_1}^2},$$
(9)

where $\sigma_s^2 = E\{s^2(n, \theta_s)\}$ and $\sigma_{v_1}^2 = E\{v_1^2(n)\}.$

2.3. Simultaneously active desired and interfering sources

When a desired source located at θ_s and an interferer located at θ_i are both active, we have in the presence of diffuse noise

$$x_1(n) = s(n, \theta_s) + i(n, \theta_i) + v_1(n).$$
(10)

Assuming the desired signal, the interfering signal, and the diffuse noise to be pairwise uncorrelated, from (3) and the discussion in section 2.2, the optimal w is given by

$$w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) = \frac{c(\theta_s) + c(\theta_i)/\gamma + 0.5/\tilde{\gamma}}{1 + 1/\gamma + 1/\tilde{\gamma}},$$
(11)

where $\gamma = \sigma_s^2/\sigma_i^2$, $\tilde{\gamma} = \sigma_s^2/\sigma_{v_1}^2$, and the notation explicitly shows the dependence of the w on γ , $\tilde{\gamma}$, θ_s , and θ_i . We next present our framework to set appropriate detection thresholds on $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i)$ for simultaneously active desired and interfering sources in the presence of diffuse noise.

3. DETERMINING DETECTION THRESHOLDS

First, we discuss the threshold for detecting a dominant desired source. This is followed by a discussion on detecting the dominant interferer. Note that we decouple the two detections, i.e., a dominant interferer is not automatically detected when there is no dominant desired source (e.g., both could be inactive or both equally dominant).

3.1. Detecting a dominant desired source

We seek a threshold T_s such that the desired source is detected as dominant when $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) \geq T_s$. We adopt a deterministic approach and look at worst-case behaviors. To ensure that the desired source is not detected as dominant below a certain SIR γ_s^{low} , we require the following:

Condition 1 For all $\theta_s \in \Psi_s$, for all $\theta_i \in \Psi_i$, and for a given $\tilde{\gamma}$, $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) < T_s$ whenever $\gamma < \gamma_s^{\text{low}}$.

Since $c(\theta_s) \leq c(0) = 1$ and $c(\theta_i) \leq c(\frac{\pi}{2} + \psi_i) = c_i$, we have $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) \leq w(\gamma, \tilde{\gamma}, 0, \frac{\pi}{2} + \psi_i)$, and so we set

$$T_{\rm s} = w(\gamma_s^{\rm low}, \tilde{\gamma}, 0, \frac{\pi}{2} + \psi_i) = \frac{1 + c_i / \gamma_s^{\rm low} + 0.5 / \tilde{\gamma}}{1 + 1 / \gamma_s^{\rm low} + 1 / \tilde{\gamma}}.$$
 (12)

The threshold $T_{\rm s}$ completely determines the SIR $\gamma_s^{\rm high}$ such that:

Condition 2 For all $\theta_s \in \Psi_s$, for all $\theta_i \in \Psi_i$, and for a given $\tilde{\gamma}$, $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) \geq T_s$ whenever $\gamma \geq \gamma_s^{\text{high}}$.

Since $c(\theta_s) \ge c(\psi_s) = c_s$ and $c(\theta_i) \ge c(\pi) = 0$, we have $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) \ge w(\gamma, \tilde{\gamma}, \psi_s, \pi)$, and so we obtain γ_s^{high} by setting

$$T_{\rm s} = w(\gamma_s^{\rm high}, \tilde{\gamma}, \psi_s, \pi) = \frac{c_s + 0.5/\tilde{\gamma}}{1 + 1/\gamma_s^{\rm high} + 1/\tilde{\gamma}}, \qquad (13)$$

so that

$$\gamma_s^{\text{high}} = \frac{T_{\text{s}}\tilde{\gamma}}{\tilde{\gamma}(c_s - T_{\text{s}}) - (T_{\text{s}} - 0.5)}.$$
(14)

To summarize the discussion on threshold calculation so far, for a given ψ_s , ψ_i , and $\tilde{\gamma}$, we select a $\gamma_s^{\rm low}$. Then we compute the threshold T_s according to (12), which guarantees that a desired source will not detected as dominant for SIRs below $\gamma_s^{\rm low}$ and will be detected as dominant for SIRs above $\gamma_s^{\rm high}$.

While we are free to select the design parameter $\gamma_s^{\rm low}$, the range of valid values for $\gamma_s^{\rm low}$ depends on ψ_s , ψ_i , and $\tilde{\gamma}$, as we show next. In the following discussion, we assume that ψ_s , ψ_i , and $\tilde{\gamma}$ are fixed. Clearly, the lower bound for $\gamma_s^{\rm low}$ equals zero, corresponding to an SIR of $-\infty$ dB. In this case, $T_{\rm s} = c_i$. To obtain the upper bound of $\gamma_s^{\rm low}$, we first observe from (13) that an upper bound on $T_{\rm s}$ is given by

$$T_{\rm s} \le \frac{\tilde{\gamma}c_s + 0.5}{\tilde{\gamma} + 1},\tag{15}$$

which follows by setting $\gamma_s^{\rm high} = \infty$. Substituting (12) for $T_{\rm s}$ in (15), we obtain the upper bound on $\gamma_s^{\rm low}$ as

$$\gamma_s^{\text{low}} \le \frac{\tilde{\gamma}}{\tilde{\gamma}+1} \frac{c_s - c_i}{1 - c_s} + \frac{1}{\tilde{\gamma}+1} \frac{0.5 - c_i}{1 - c_s}.$$
 (16)

Thus the design parameter γ_s^{low} should be selected such that (16) is satisfied. We see from (16) that for a positive upper bound on γ_s^{low} ,

we require that the ranges Ψ_s and Ψ_i of the possible locations of the desired source and interferer, and thus c_s and c_i , are such that

$$c_s > c_i \frac{\tilde{\gamma} + 1}{\tilde{\gamma}} - \frac{0.5}{\tilde{\gamma}}.$$
(17)

The design parameters ψ_s , ψ_i , $\tilde{\gamma}$, and γ_s^{low} can be set during the individual fitting session between the user and an audiologist, or automatically be set to different values based on the output of the environment classification algorithm that is present in modern hearing aids [4]. We now consider an example.

Example 1 Assume the desired source to be located within $\Psi_s \equiv [0, 20] \cup [340, 360]$, and the interferer to be located within $\Psi_i \equiv [120, 240]$, i.e., $c_s = 0.9698$ and $c_i = 0.329$. We fix $\tilde{\gamma} = 10$, corresponding to a signal-tonoise ratio of 10 dB. For the above parameters, the range of valid values for γ_s^{low} is obtained from (16) as $0 \leq \gamma_s^{\text{low}} \leq 19.8$, i.e., on a dB scale, γ_s^{low} should be between $-\infty$ and ≈ 13 dB. In this example, we select $\gamma_s^{\text{low}} \approx 3.16$, i.e., we do not wish to detect the desired source as dominant for SIRs below 5 dB. From (12), we get $T_s = 0.8149$. For the above choice of γ_s^{low} , the corresponding γ_s^{high} is obtained by inserting the computed value of T_s in (14) and is found to be $\gamma_s^{\text{high}} \approx 6.6$, or ≈ 8.2 dB. Thus, for SIRs above 8.2 dB, the algorithm always detects a desired source is detected as dominant depends on the exact location $\psi_s^* \in \Psi_s$ of the desired source, and $\psi_i^* \in \Psi_i$ of the interferer.

In the above discussion, to compute the detection threshold $T_{\rm s}$ for a given ψ_s , ψ_i , and $\tilde{\gamma}$, we began by specifying a certain $\gamma_s^{\rm low}$ in condition 1. An alternative approach is to start by specifying a certain $\gamma_s^{\rm high}$ in condition 2, and compute the threshold $T_{\rm s}$ using (13). Analogous to the upper bound on $\gamma_s^{\rm low}$, a lower bound on $\gamma_s^{\rm high}$ is given by

$$\gamma_s^{\text{high}} \ge \frac{c_i}{c_s - c_i + (0.5 - c_i)/\tilde{\gamma}}.$$
 (18)

It can be shown that $\gamma_s^{\text{high}} \ge 0$ if (17) is satisfied.

3.2. Detecting a dominant interferer

Here, we seek a threshold T_i such that $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) \leq T_i$ when the interferer is dominant. Analogous to section 3.1, to ensure that the interferer is not detected as dominant above a certain SIR γ_i^{high} , we require:

Condition 3 For all $\theta_s \in \Psi_s$, for all $\theta_i \in \Psi_i$, and for a given $\tilde{\gamma}$, $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) > T_i$ whenever $\gamma > \gamma_i^{\text{high}}$.

Since $c(\theta_s) \ge c(\psi_s) = c_s$ and $c_{\theta_i} \ge c(\pi) = 0$, we have $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) \ge w(\gamma, \tilde{\gamma}, \psi_s, \pi)$, and so we set

$$T_{\rm i} = w(\gamma_i^{\rm high}, \tilde{\gamma}, \psi_s, \pi) = \frac{c_s + 0.5/\tilde{\gamma}}{1 + 1/\gamma_i^{\rm high} + 1/\tilde{\gamma}}.$$
 (19)

The threshold T_i completely determines the SIR γ_i^{low} such that:

Condition 4 For all $\theta_s \in \Psi_s$, for all $\theta_i \in \Psi_i$, and for a given $\tilde{\gamma}$, $w(\gamma, \tilde{\gamma}, \theta_s, \theta_i) \leq T_i$ whenever $\gamma \leq \gamma_i^{\text{low}}$.

where γ_i^{low} is obtained from the following relation

$$T_{i} = w(\gamma_{i}^{\text{low}}, \tilde{\gamma}, 0, \frac{\pi}{2} + \psi_{i}) = \frac{1 + c_{i}/\gamma_{i}^{\text{low}} + 0.5/\tilde{\gamma}}{1 + 1/\gamma_{i}^{\text{low}} + 1/\tilde{\gamma}}.$$
 (20)

Analogous to section 3.1, a lower bound can be obtained on the values that $\gamma_i^{\rm high}$ can assume and is given by

$$\gamma_i^{\text{high}} \ge \frac{c_i}{c_s - c_i + (0.5 - c_i)/\tilde{\gamma}}.$$
 (21)

For a meaningful detection scheme, we need $T_{\rm s} > T_{\rm i}$. From (13) and (19), the above inequality is satisfied if $\gamma_s^{\rm high} > \gamma_i^{\rm high}$.

4. EXPERIMENTS

A hearing aid, with two microphones spaced 0.008 m apart, was mounted on to the right ear of a KEMAR (Knowles Electronics Manikin for Acoustic Research) placed in a room with a measured reverberation time (T_{60}) of about 800 ms. Processing was performed on a PC on the signals obtained by directly tapping the microphones. In the first experiment, white Gaussian noise was played from a loudspeaker placed at ear level, and at two azimuths, 0 deg. (desired) and 135 deg. (interferer), one meter away from the KEMAR. The recorded signals were downsampled to 16 khz. A mixed signal was created with only the desired signal active for the first 10 seconds, only the interferer active for the next 10 seconds and both being active for the last 10 seconds at an SIR of 0dB. The weight w given by (3) was adaptively estimated using the NLMS algorithm [5, ch. 9]. Fig. 3 shows that in this simple example w behaves as expected with respect to the detection thresholds.



Fig. 3. Evolution of the adaptive weight w. For first 10s, source at 0 deg. is active, for next 10s interferer at 135 deg. is active, and for last 10s both are simultaneously active (SIR ≈ 0 dB). The values T_s and T_i correspond to $\Psi_s \equiv [0, 20] \cup [340, 360], \Psi_i \equiv [120, 240], \tilde{\gamma}=10 (10 \text{ dB}), \gamma_s^{\text{low}}= 3.1623 (5 \text{ dB})$ and $\gamma_i^{\text{high}} = 1 (0 \text{ dB})$.

As mentioned earlier, the proposed detection scheme can be used to control the update of an adaptive filter whenever the desired source is detected as dominant. In a second experiment, we evaluated the performance of the scheme in such a task. A residual signal was generated using the observed microphone signals as $r(n) = x_1(n) - h(n) * x_2(n)$, where * denotes convolution and the filter h(n) was updated (NLMS) to minimize the energy of r(n)whenever the desired signal was detected as dominant. Such a filter is useful, e.g., in generating a speech free noise reference. The SIR in the residual is a good performance metric. The lower the SIR in the residual, the better the performance (there is less speech leakage in the noise reference). To facilitate objective performance evaluations, the desired and interference signals were recorded separately and then added together. In a first run, the filter was updated using the mixed signals, and the filter coefficients were stored for each processed short-time segment. In a second run, the desired and interfering signals were separately processed using the stored filter coefficients to obtain two residual signals $r_s(n) = s_1(n) - h(n) * s_2(n)$ and $r_i(n) = i_1(n) - h(n) * i_2(n)$, where $s_k(n)$ and $i_k(n)$, k = 1, 2correspond to the desired and interfering signals at the k^{th} microphone respectively. The SIR in the residual was then computed as $10 \log_{10} \sum r_s^2(n) / \sum r_i^2(n).$

The design parameters ψ_s , ψ_i , and γ_s^{low} were set as in Example 1. Again, the desired source was located at 0 deg. and the interferer at 135 deg. We expect the beamformer not to adapt for SIRs lower than γ_s^{low} , which was 5 dB in our example. To verify this behavior, the recorded desired and interfering signals were mixed such that for the first 10s the SIR in the input signal was 20 dB after which the SIR dropped to 4 dB. As a baseline case, we consider a detection scheme that assumes the desired speech to be anywhere in the front half plane, which results in $T_s = 0.5$.

For the first 10s, since the SIR is high, both detection methods result in similar performance as seen in Figs. 4a and 4b, which plots the SIR (averaged over blocks of length 1024) in the residual as a function of time. As mentioned earlier, lower values indicate better performance (less leakage of the desired signal). After 10s, when the SIR falls to 4 dB, filter update freezes using the proposed detection scheme. The baseline method however continues to detect desired speech at 4 dB SIR, which causes the filter to diverge and allows leakage of the desired signal in the noise reference resulting in the poor performance (high SIR in residual) seen in Figs. 4a and 4b. Conventional energy based detection schemes typically attempt to avoid this divergence by making a positive detection only if $E\{f^2(n)\} > E\{b^2(n)\} + \Delta dB$, where Δ is chosen in an adhoc manner. The proposed detection scheme allows the selection of appropriate thresholds based on the design constraints of the application instead of in an adhoc manner.



Fig. 4. SIR in the noise reference. *Lower* values indicate *better* performance. Two cases are shown, $T_s = 0.8149$ as in Example 1 (solid), and the baseline case $T_s = 0.5$ (dotted). The SIR in the input signal is 20 dB for the first 10 seconds and then falls to 4 dB. Desired (0 deg.) and interfering (135 deg.) signals are (a). white noise (b). two different speech signals. In both (a) and (b), for $T_s = 0.5$ the filter continues to adapt after 10s at the low SIR of 4 dB, causing leakage of desired signal in the noise reference, and thus poor performance. $T_s = 0.8149$ avoids this problem.

5. CONCLUSIONS

A spatial signal activity detection scheme based on directional processing has been proposed for use in hearing aids. Threshold selection follows a theoretic framework based on design parameters that can be appropriately set for a given application. Experiments in practical reverberant environments confirm the validity of such a scheme when used to control the update of an adaptive filter.

6. ACKNOWLEDGEMENTS

The authors are grateful to Rene Derkx at Philips Research for valuable suggestions.

7. REFERENCES

- J. Ramirez, *et al.*, "Efficient voice activity detection algorithms using long-term speech information," *Speech Communication*, vol. 42, pp. 271–287, Apr. 2004.
- [2] M. W. Hoffman, et al., "GSC-based spatial voice activity detection for enhanced speech coding in the presence of competing speech," *IEEE Trans. Speech Audio Processing*, vol. 9, no. 2, pp. 175–179, Mar. 2001.
- [3] A. Koul and J. E. Greenberg, "Using intermicrophone correlation to detect speech in spatially separated noise," *EURASIP Journal on Applied Signal Processing*, pp. Article ID 93 920, 14 pages, 2006.
- [4] V. Hamacher, et al., "Signal processing in high-end hearing aids: State of the art, challenges, and future trends," EURASIP Journal on Applied Signal Processing, no. 18, pp. 2915–2929, 2005.
- [5] S. Haykin, Adaptive Filter Theory, 3rd ed. Prentice Hall, 1995.
- [6] G. W. Elko and A.-T. N. Pong, "A simple adaptive first-order differential microphone," in *Proc. IEEE Workshop Appln. Sig. Proc. Audio and Acoustics*, 1995.
- [7] F.-L. Luo, et al., "Adaptive null-forming scheme in digital hearing aids," IEEE Trans. Signal Processing, vol. 50, no. 7, pp. 1583–1590, July 2002.