15

Applications of Adaptive Signal Processing Methods in High-End Hearing Aids

Volkmar Hamacher, Eghart Fischer, Ulrich Kornagel, Henning Puder

Siemens Audiological Engineering Group, Erlangen, Germany

15.1 Introduction

In the past ten years the technical capabilities of hearing aids have considerably increased. One important mile stone was the changeover from analog to digital technology enabled by the continuous progress in semi-conductor technology. This chapter focuses on the powerful digital signal processing of modern high-end hearing aids.

In principal, the development of hearing aids incorporates two aspects, namely the audiological and the technical point of view. The former focuses on items like the recruitment phenomenon, the speech intelligibility of hearing impaired persons or just on the question of hearing comfort. Concerning these topics different algorithms intending to improve the hearing ability are presented in this chapter. These are automatic gain controls, directional microphones and noise reduction algorithms. Besides the audiological point of view there are several purely technical problems which have to be solved. An important one is the acoustic feedback. Another instance is the proper automatic control of all hearing aid components by means of a classification unit.

Fig. 15.1 schematically shows the main signal processing units of a high-end hearing aid [23, 24]. We will follow the depicted signal flow and discuss the state-of-the-art and the challenges for the different components. A coarse overview is given below.

First, the acoustic signal is captured by up to three microphones. The microphone signals are processed into a single signal within the directional microphone unit which will be discussed in Sec. 15.2.

The obtained mono-signal is further processed separately for different frequency ranges. In general this requires an analysis filterbank and a corresponding signal synthesis. The main frequency band dependent processing steps are noise reduction as detailed in Sec. 15.3 and signal amplification combined with dynamic compression as discussed in Sec. 15.4.

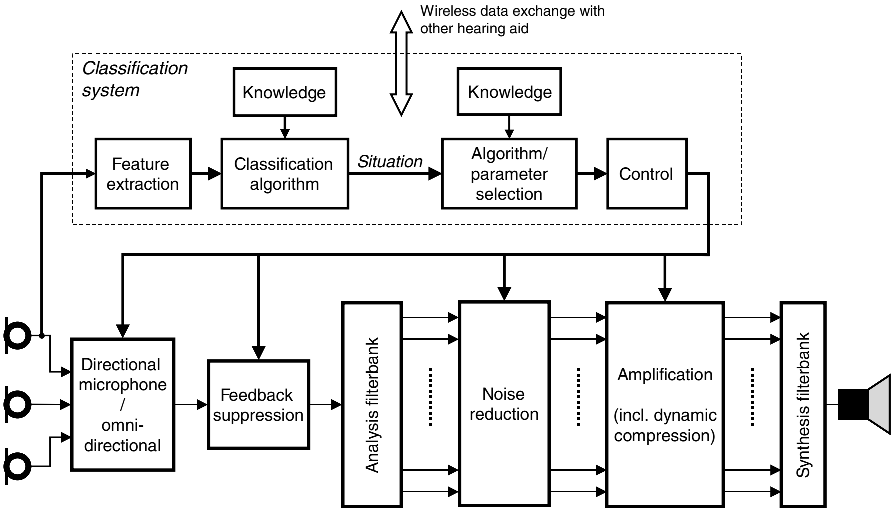

Fig. 15.1. Processing stages of a high-end hearing aid.

A technically challenging problem of hearing aids is the risk of acoustic feedback that is provoked by strong signal amplification in combination with microphones and receiver being close to each other. The two major alternatives to remedy feedback are the feedback suppression approach and the feedback compensation approach. Details regarding the feedback problem and possible solutions are discussed in Sec. 15.5. Note that feedback suppression can be applied at different stages of the signal flow dependent on the chosen strategy. One reasonable solution is shown in Fig. 15.1, where feedback suppression is applied right after the (directional) microphone unit.

Almost all mentioned hearing aid components can be tuned differently for optimal behavior in various listening situations. Providing different "programs" that can be selected by the hearing impaired is a simple means to account for this difficulty. However, the usability of the hearing aid can be significantly improved if control of the signal processing algorithms can be handled by the hearing aid itself. Thus, a classification and control unit, as shown in the upper part of Fig. 15.1 and described in Sec. 15.6, is required and offered by advanced hearing aids. In binaural use, the effectiveness of this unit can be significantly improved by means of wireless coupling of both hearing aids.

15.2 Directional Microphones

One of the main problems for the hearing impaired is the reduction of speech intelligibility in noisy environments, which is mainly caused by the loss of temporal and spectral resolution in the auditory processing of the impaired ear. The loss in signal-to-noise ratio (SNR) is estimated to be about 4-10 dB [12]. Additionally, the natural directivity of the outer ear is not effective when BTE (behind-the-ear) instruments are used. To compensate for these disadvantages, directional microphones have been used in hearing aids for several years and have proved to significantly increase speech intelligibility in various noisy environments [61].

15.2.1 First-Order Differential Arrays

In advanced hearing aids, directivity is achieved by differential processing of two nearby omni-directional microphones in endfire geometry (first-order differential array) to create a direction-dependent sensitivity. The directivity pattern of the system is defined by the ratio r of the internal delay T_i and the external delay due to the microphone spacing d (typically 7-16 mm). In this example the ratio was set to $r = 0.57$ resulting in a super-cardioid pattern also shown in Fig. 15.2. To compensate for the high-pass characteristic introduced by the differential processing, an appropriate low-pass filter (LPF) is usually added to the system.

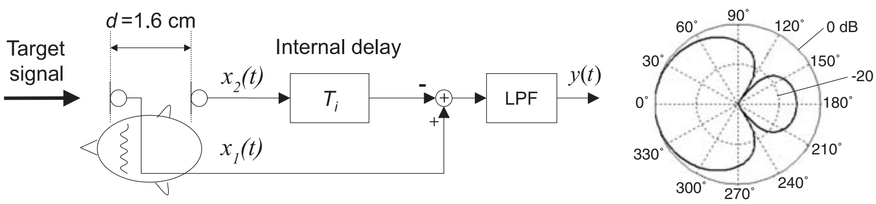

Fig. 15.2. Signal processing of a first order differential microphone.

The performance of a directional microphone is quantified by the directivity index (DI),

$$DI\left(e^{j\Omega}\right) = 10\log_{10}\left(\frac{\left|H\left(e^{j\Omega},\varphi_0,\theta_0\right)\right|^2}{\frac{1}{4\pi}\int\limits_{-\pi/2}^{\pi/2}\int\limits_{0}^{2\pi}\left|H\left(e^{j\Omega},\varphi,\theta\right)\right|^2\sin\left(\theta\right)d\varphi d\theta}\right) \tag{15.1}$$

where $H\left(e^{j\Omega},\varphi,\theta\right)$ denotes the spatial-temporal transfer function of the array depending on azimuth φ and elevation θ in a spherical coordinate system.

The DI is defined by the power ratio of the output signal (in dB) between sound incidence only from the front and the diffuse case, i.e. sound coming equally from all directions. Consequently, the DI can be interpreted as the improvement in SNR that can be achieved for frontal target sources in a diffuse noise field. The hyper-cardioid pattern ($r = 0.34$) provides the best directivity with a DI of 6 dB, which is the theoretical limit for any two-microphone array processing [4]. However, in practical use these DI values cannot be reached due to shading and diffraction effects caused by the human head. Fig. 15.3 illustrates the impact of the human head on the directivity of a BTE hearing aid with a two-microphone array. The most remarkable point is that the direction of maximum sensitivity is shifted aside by approximately 40°, if the device is mounted behind the ear of a KEMAR (Knowles Electronic Manikin for Acoustic Research). Consequently, the DI which is related to the 0° front direction, decreases typically by 1.5 dB compared to the free-field condition.

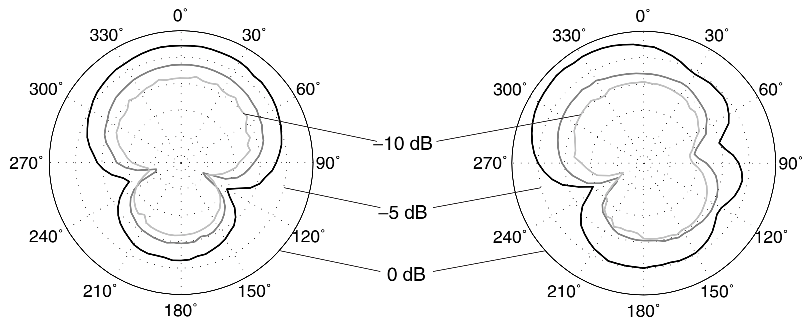

Fig. 15.3. Impact of head shadow and diffraction on the directivity pattern of a BTE hearing aid with a two-microphone differential array in free field (left plot) and mounted behind the left ear of a KEMAR (right plot). The black, dark gray and light gray curves show the directivity pattern for 2 kHz, 1 kHz, and 500 Hz, respectively (10 dB grid).

The performance related to speech intelligibility is quantified by a weighted average of the DI across frequency, commonly referred to as the AI-DI. The weighting function is the importance function used in the Articulation Index (AI) method [46] and takes into account that SNR improvements in different frequency bands contribute differently to the speech intelligibility. As shown in Fig. 15.4 for a hyper-cardioid pattern, the AI-DI (as measured on KEMAR) of two-microphone arrays in BTE instruments ranges from 3.5 to 4.5 dB. For speech intelligibility tests in mainly diffuse noise the effect of directional microphones typically leads to improvements of the Speech-Reception-Threshold (SRT) in the range from 2 to 4 dB (e.g. [53]).

In high-end hearing aids, the directivity is normally adaptive in order to achieve a higher noise suppression effect in coherent noise, i.e. in situations with one dominant noise source [12, 50]. As depicted in Fig. 15.6, the primary direction from which the noise arrives is continually estimated and the direc-

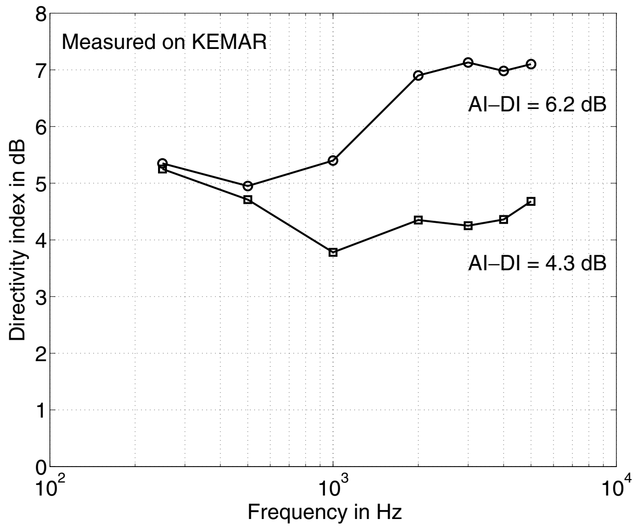

Fig. 15.4. DI and AI-DI for a fist-order array (Siemens Triano S) and the combination with a second-order array (see Sec. 15.2.2) in the upper frequency range (Siemens Triano 3).

tivity pattern is automatically adjusted so that the directivity notch matches the main direction of noise arrival. Instead of implementing computationally expensive fractional delay filters, the efficient method proposed by Elko and Pong [15] can be used. In this approach, the shape of the directivity pattern is steered by a weighted sum of the output signals of two cardioid patterns, one facing to the front (0°), the other one facing to the back (180°). The position of the directivity notch is monotonically related to the weighting factor. Great demands are made on the adaptation algorithm. The steering of the directional notch has to be reliable and accurate and should not introduce artifacts or perceivable changes in the frequency response for the 0°-target direction, which would be annoying for the user. The adaptation process must be fast enough (< 100 ms) to compensate for head movements and to track moving sources in common listening situations, such as conversation in a street-cafe with interfering traffic noise. To ensure that no target sources from the front hemisphere are suppressed, the directivity notches are limited to the back hemisphere ($90° - 270°$). Finally, the depth of the notches is limited to prevent hazardous situations for the user, e.g. when crossing the street while a car is approaching.

Fig. 15.5 shows a measurement in an anechoic test chamber with an adaptive directional microphone BTE instrument mounted on the left KEMAR ear. A noise source was moved around the head and the output level of the hearing aid was recorded (dashed line). Compared to the same measurement for a non-adaptive super-cardioid directional microphone (solid line), the higher

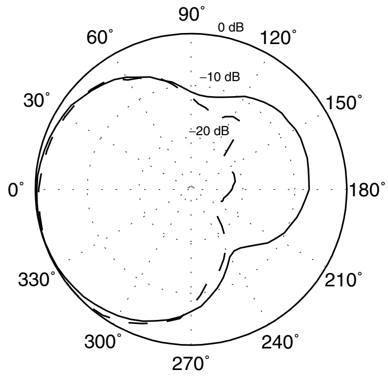

Fig. 15.5. Suppression of a noise source moving around the KEMAR for a BTE instrument (mounted on left ear) with directional microphone in adaptive mode (dashed line) and non-adaptive mode (solid line).

suppression effect for noise incidence from the back hemisphere is clearly visible.

In order to achieve optimum performance also for natural sound fields with non-diffuse spatial and non-white frequency distribution, it is advantageous to use a frequency specific implementation of the adaptive directional microphone principle. Fig. 15.6 shows the principle of a four-channel adaptive differential microphone with four directional characteristics that can independently adapt to the main direction of incidence of the interferer within the corresponding frequency band.

Studies have shown the advantage of using adaptive directional processing instead of static directivity. For a situation with three interferers from 90, 180 and 270°, as shown in Fig. 15.7 an improvement of 1.5 dB for the SRT could be achieved, see Fig. 15.8. The speech reception threshold SRT is a measure from speech audiometry, which determines the lowest intensity level of speech, presented in noise, at which the patient can correctly identify 50% of the words.

15.2.2 Second-Order Differential Arrays

Using second order arrays that are using three omnidirectional microphones [4] instead of two, generally a significantly higher DI can be achieved. (6-8 dB instead of 4-5 dB for hearing aids worn on the head.)

Unfortunately, this increase in DI has to be payed by a higher self induced noise with second order differential processing compared to first order processing. Fig. 15.9 and Fig. 15.10 compare directivity and the self induced noise for first vs. second order processing.

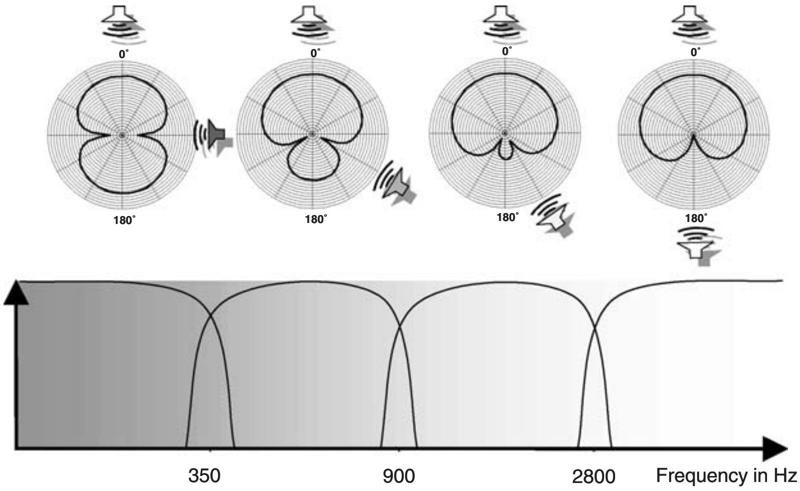

Fig. 15.6. Principle of a 4-channel adaptive directional microphone.

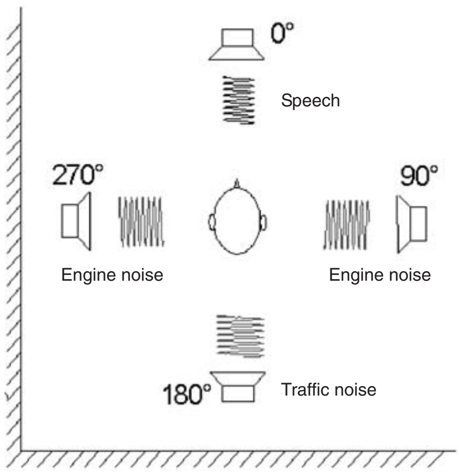

Fig. 15.7. Setup with one speaker and three interferers.

There are different possibilities to deal with the noise problem for higher order differential microphones.

One is the realization of a combined first- and second-order directional processing in a hearing aid with three microphones [50], which is shown in Fig. 15.11. Due to the high sensitivity to microphone noise especially in the low frequency range the second-order processing is limited to the frequencies above approx. 1 kHz which are most important for speech intelligibility.

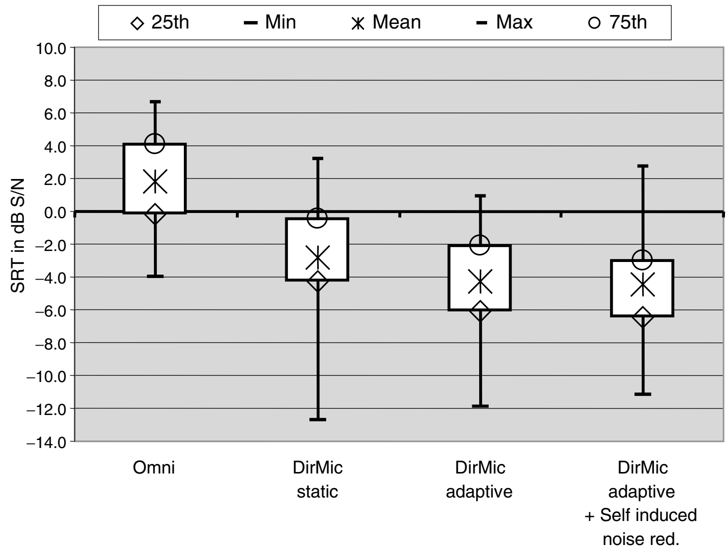

Fig. 15.8. Adaptive directional processing is showing an improvement in SRT of about 1.5 dB compared to static directional processing.

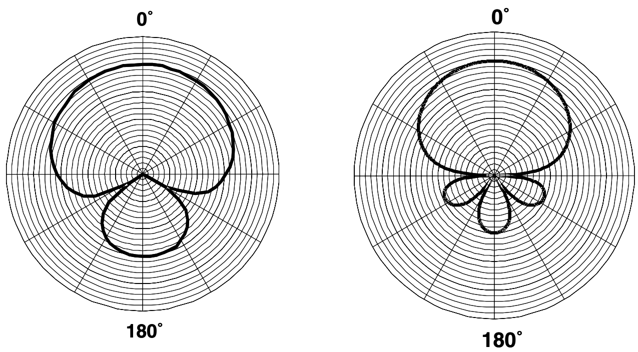

Fig. 15.9. Polar plots of first (left) and second (right) order differential microphones.

As shown in Fig. 15.4 calculation of the AI-DI leads to values of 6.2 dB, i.e. an improvement in AI-DI of about 2 dB compared to a first-order system worn at the head. It should be noted that for many listening situations, improvements of 2 dB in the AI-DI can have a significant impact on speech intelligibility [54].

Another possibility to handle the noise problem is to make the directional microphone system not only adaptive in terms of spatial shape of the directivity but also adaptive in terms of adaptation to the level of the target input

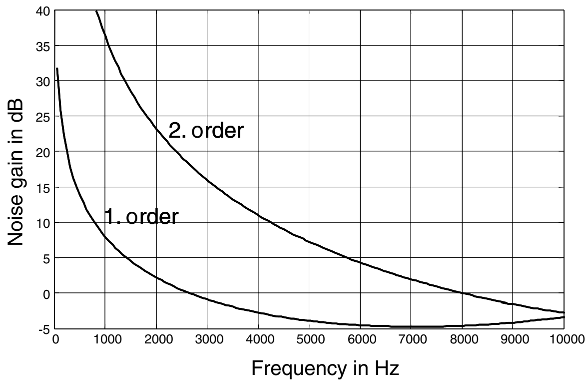

Fig. 15.10. Self induced microphone noise gain for first and second order processing with flat frequency response for the zero degree direction.

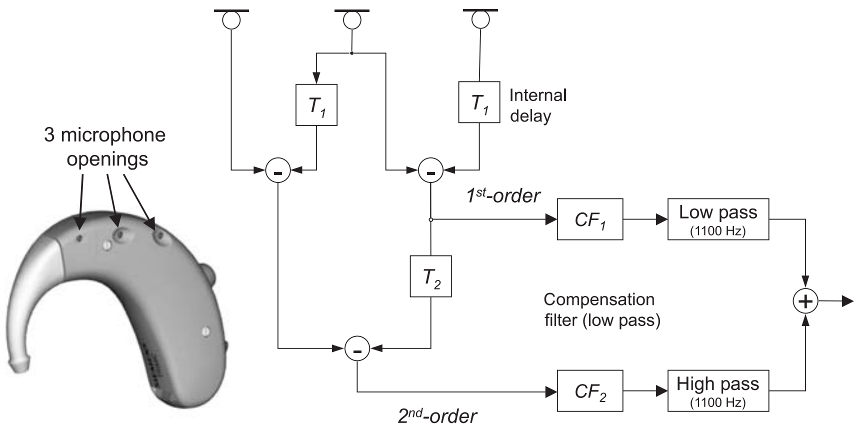

Fig. 15.11. Combined first- and second-order processing in a BTE hearing aid with three microphones.

signal: If the input level is high enough, more directivity with more self induced noise may be applied in the considered frequency band as the noise is to a large extent masked by the target signal. If the input level is low, less directivity with less self induced noise would be used in order to avoid audible noise. Fig. 15.12 illustrates the principle of this approach.

The differential approach for directional microphones as described above is of course just one - though very effective - method of generating directivity. There are several other ways with their specific advantages and disadvantages to build directional systems in hearing aids, e.g. adaptive beamformers (e.g. [13, 29, 30, 33, 34, 62]), beamformers, taking head shadow effects into account [44] and blind source separation techniques (e.g. [1, 14]).

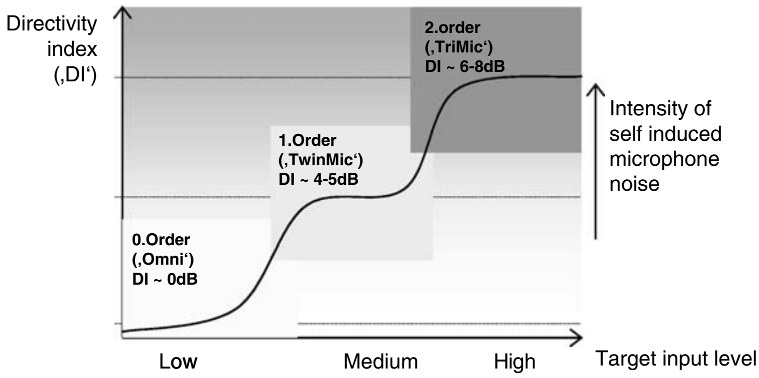

Fig. 15.12. Principle of the input level dependent directivity: Order of directivity is increasing with increasing input level.

15.3 Noise Reduction

Directional microphones, as described in the preceding section are usually not applicable to small ear canal instruments for reasons of size constraints and the assumption of a free sound field which is not met inside the ear canal. Consequently, one-microphone noise reduction algorithms became an essential signal processing stage of today's high-end hearing aids. Due to the lack of spatial information, these approaches are based on the different signal characteristics of speech and noise. Usually, despite of the fact that these methods may improve the SNR, they could yet not prove to enhance the speech intelligibility.

In the following, we will focus on two noise reduction methods which both showed their suitability for hearing aids. The first method is also one of the early ones in the field. It decomposes the noisy signal into many subbands and applies a long-term smoothed attenuation to those subbands for which the average SNR is very low. The second, a Wiener-filter based method applies a short-term attenuation to the subband signals and is thus able to enhance the SNR even for those signals for which the desired signal and the noise cover the same frequency range. Both methods can also be well combined: During speech activity, the Wiener filter approach exhibits the stronger impact whereas during speech pauses or for frequency bands with a very low long-term SNR the long-term noise reduction shows the stronger impact. Both effects are desired and by choosing the maximum noise reduction of both methods they can be both achieved.

At the end of this section, we will have an outlook to the application of Ephraim-Malah based short-term noise reduction approaches for hearing aids.

15.3.1 Long-Term Smoothed, Modulation Frequency Based Noise Reduction

Supposing a noisy input signal which has been decomposed into several frequency bands, the task of this noise reduction unit is to apply a long-term smoothed attenuation to frequency bands with a very low SNR which do not contain any remarkable speech components.

15.3.1.1 Theoretical Basis

The theoretical basis for distinguishing speech components from others is that speech signals exhibit a characteristic modulation frequency at 4 Hz [47].

For calculating this characteristic modulation frequency,

- first the envelope of a speech signal is determined according to Fig. 15.13,

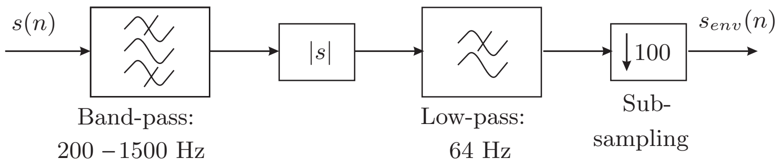

Fig. 15.13. Processing for determining the signal envelope.

- then DC component is removed by an IIR filter of first order:

$$s_{\mathrm{env,AC}}(n) = \frac{1+\beta}{2}\Big[s_{\mathrm{env}}(n) - s_{\mathrm{env}}(n-1)\Big] + \beta\, s_{\mathrm{env,AC}}(n-1) \quad (15.2)$$

  with a typical value for $\beta = 0.995$.
- The power spectral density (PSD) of the envelope has to be calculated, normalized by the mean power $m_s^{(2)} = \mathrm{E}\{{s_{\mathrm{env}}}^2(n)\}$:

$$S_{ss}(\Omega) = PSD\big\{s_{\mathrm{env,AC}}(n)\big\}. \quad (15.3)$$

$$S_{ss,\mathrm{norm}}(\Omega) = \frac{S_{ss}(\Omega)}{m_s^{(2)}}, \quad (15.4)$$

- and finally the PSD is summed over Terz-bands for determining the modulation spectrum at a logarithmic scale:

$$S_{\mathrm{mod\_spec}}(i) = \frac{1}{2\pi}\int\limits_{\Omega=\Omega_i}^{\Omega_{i+1}} S_{ss,\mathrm{norm}}(\Omega)\, d\Omega, \quad (15.5)$$

  where Ω_i are the limits of the Terz-bands.

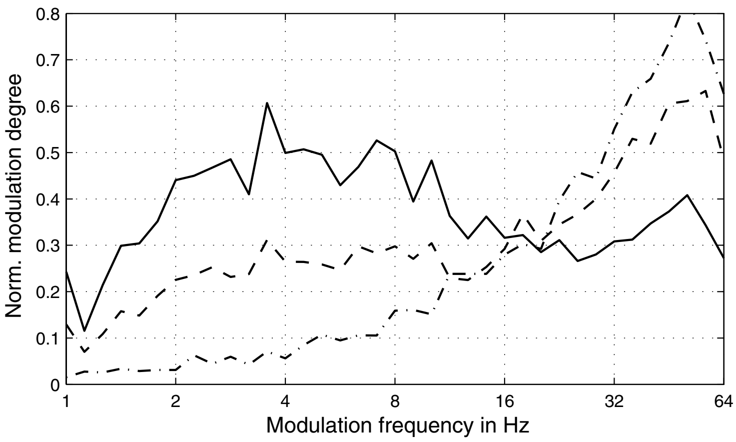

Fig. 15.14. Modulation spectrum of clean speech (solid), noisy speech (dash) and noise (dash-dot).

Such a spectrum is depicted in Fig. 15.14 for three types of signals: clean speech, noisy speech and noise.

One clearly observes that the quantity of the modulation spectrum at 4 Hz is directly related to the SNR of the corresponding signal: For the given example this values decreases form 0.6 for clean speech to 0.3 for noisy speech and to nearly zero for pure noise.

Based on the discussed properties of the modulation spectrum, a long-term noise reduction method can be designed: After the decomposition of the noisy input signal into several frequency subbands, the modulation spectrum at 4 Hz is determined for each subband. Then, this value has to be mapped to a noise reduction gain value, e.g. by

$$g = \max\left[\min\left\{v \cdot \left[S_{\mathrm{mod\_spec}}(4\ \mathrm{Hz}) - b\right], 1\right\}, spfl\right]. \tag{15.6}$$

Here, the additive constant b and the gain v map the time-frequency dependent 4 Hz-modulation spectrum to the range of the noise-reduction gains, limited between the Spectral Floor ($spfl$) and 1. The Spectral Floor assures that the attenuation does not exceed an adjustable maximum attenuation of approximately 10 to 15 dB, i.e.

$$max\_atten = -20\,\log_{10}(spfl). \tag{15.7}$$

15.3.1.2 Computational Efficient Realization

Since the procedure for determining the modulation spectrum around 4 Hz is computationally expensive, it would be advantageous to provide an alternative calculation method which also shows the desired relation without explicitly determining the modulation spectrum.

A very simple method which fulfills these requirements mainly consists of two short-term average magnitude (SAM) units, which perform a calculation according to:

$$s_{\mathrm{SAM}}(n) = \begin{cases} \alpha_r \, s_{\mathrm{SAM}}(n-1) + (1-\alpha_r)\,|s(n)| \, : |s(n)| > s_{\mathrm{SAM}}(n-1)\,, \\ \alpha_f \, s_{\mathrm{SAM}}(n-1) + (1-\alpha_f)\,|s(n)| \, : |s(n)| \leq s_{\mathrm{SAM}}(n-1)\,. \end{cases} \tag{15.8}$$

For the two SAM units different settings α_r and α_f are chosen. One unit estimates the long-term smoothed average magnitude by setting $\alpha_r = \alpha_f$, whereas the other estimator is parametrized by $\alpha_r < \alpha_f$, i.e. the output follows a raising signal power faster than a falling signal power.

With an appropriate choice of the smoothing parameters α_r and α_f for both units, the ratio of these two SAM units is equivalent to the quantity of the modulation spectrum around 4 Hz, but computationally clearly less consuming. The equivalence of the approach utilizing SAM units and the modulation spectrum around 4 Hz is shown in Fig. 15.15. Here the ratio of these two SAM units is depicted in dependence of the modulation frequency of the input signal. It can be well observed that this ratio reaches its maximum around 4 Hz.

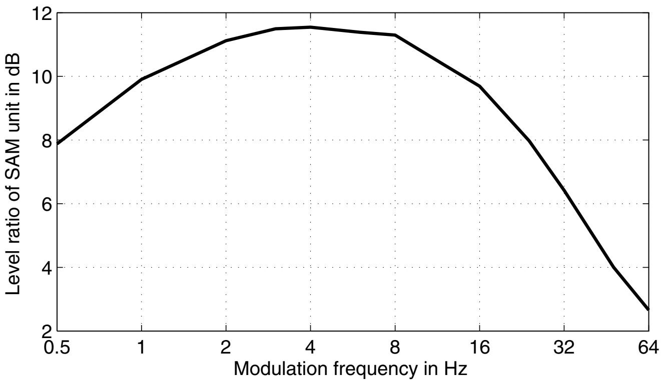

Fig. 15.15. Ratio of two SAM units with different parameter settings.

That the computational efficient approach can be well utilized for determining the long-term modulation-based noise reduction is also obvious by the results depicted in Fig. 15.16. Here, a clean and noisy speech, as well a pure noise signal are depicted in the top. Below, the corresponding modulation spectra and the applied attenuation is depicted, determined based on the

SAM-unit approach. The desired dependence of the applied attenuation on the the modulation spectrum around 4 Hz is obvious.

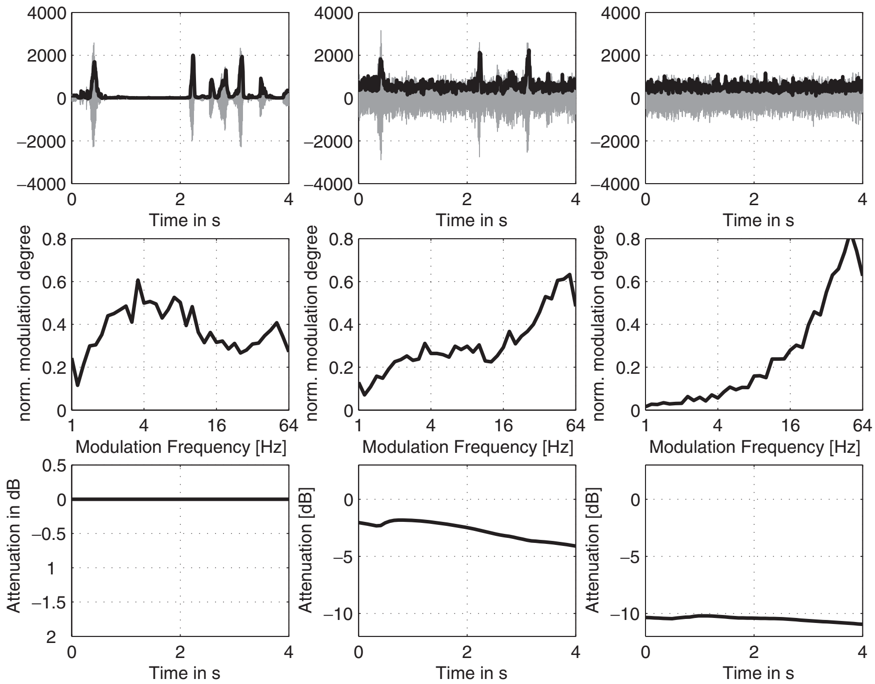

Fig. 15.16. Above: Clean speech (left), noisy speech (mid) and noise (right); Mid: Corresponding modulation spectrum; Below: Long-term noise reduction gain.

15.3.2 Wiener-Filter Based, Short-Term Smoothed Noise Reduction Methods

The aim of these noise reduction procedures is to obtain significant noise reduction performance even for signals whose desired signal and noise components are located in the same frequency range.

Applying the Wiener-filter attenuation:

$$H(\Omega, n) = \frac{S_{ss}(\Omega, n)}{S_{ss}(\Omega, n) + S_{bb}(\Omega, n)} = 1 - \frac{S_{bb}(\Omega, n)}{S_{yy}(\Omega, n)} \tag{15.9}$$

where n denotes the time indices and Ω the normalized frequency. Utilizing short-term estimates for the required power spectral densities $S_{ss}(\Omega, n)$,

$S_{bb}(\Omega, n)$ and $S_{yy}(\Omega, n)$ of speech, noise, and noisy speech, respectively, noticeable noise reduction can be obtained. In these cases, the filter coefficients $H(\Omega, n)$ directly follow short-term fluctuations of the desired signal.

However, a high audio quality noise-reduced signal cannot be easily obtained with this method. The main reason is the non-optimal estimation of power spectral densities which are required in Eq. 15.9. Here, especially the estimation of the noise power spectral density poses problems since the noise signal alone is not available.

In order to obtain reliable estimates, despite of these problems, well-known methods can be utilized. These are:

- Estimating the noise power spectral density in pauses of the desired signal which requires an algorithm to detect these pauses.
- Estimating the noise power spectral density with the Minimum Statistics Method [38] or its modifications [39].

Both methods, however, exhibit a major disadvantage: They only provide long-term smoothed noise power estimates.

However, for power spectral density estimation of the noisy signal, $S_{yy}(\Omega, n)$, which can easily be obtained by smoothing the subband input signal power, short-term smoothing has to be applied in order that the Wiener-filter gains can follow short-term fluctuations of the desired signal.

Calculating the Wiener-filter gain with differently smoothed power spectral density estimates causes the well-known Musical Tones phenomenon [5].

To avoid this unpleasant noise, a large number of procedures have been investigated of which the most widely used are

- Overestimating the noise power spectral density and
- Lower-limiting the Wiener-filter values to a minimum, the so-called Spectral Floor.

With the overestimation of the noise power spectral density, short-time fluctuations of the noise no more provoke a random "opening" of the Wiener-filter coefficients – the cause of Musical Tones.

However, this overestimation reduces the audio quality of the desired signal since especially low power signal components are more strongly attenuated or vanish due to the overestimation. Limiting the noise reduction to the Spectral Floor reduces this problem but, unfortunately, also reduces the overall noise reduction performance. Nevertheless, this reduced noise reduction performance is generally preferred against strong audio quality distortion. More sophisticated methods utilize, e.g., speech characteristics [51] or masking properties [21] of the ear to limit the Wiener attenuation and thus reduce the signal distortion without compromising the noise reduction effect too much.

The noise reduction gain one obtains with the Wiener-filter approach are depicted in Fig. 15.17 for the same signal section which had been chosen to show the long-term noise reduction in Fig. 15.16. One clearly observes that the signal attenuation follows the short-term signal power variations of the input signal: The attenuation is only reduced when short-term speech signal components are present.

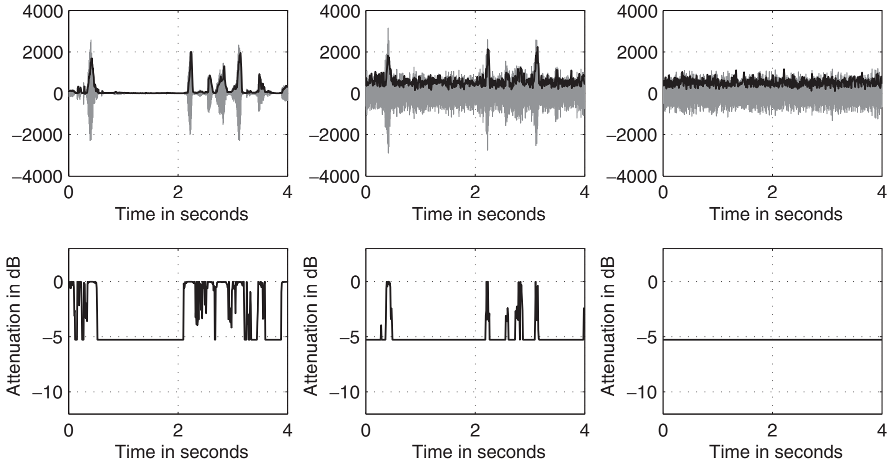

Fig. 15.17. Above: Clean speech (left), noisy speech (mid) and noise (right); Below: Short-term Wiener-filter based noise reduction gain.

However the noise attenuation has to be limited to a smaller value than the long-term noise reduction in order not to reduce speech quality. By combining the noise reduction methods one can profit by the advantages of both: the short-term selective noise reduction during speech presence of the Wiener-filter approach and the stronger noise reduction of the modulation frequency based approach during speech pauses and for frequency bands with negligible SNR. The combination is simply possible by choosing the minimum noise gain which, for the selected signal samples, is shown in Fig. 15.18.

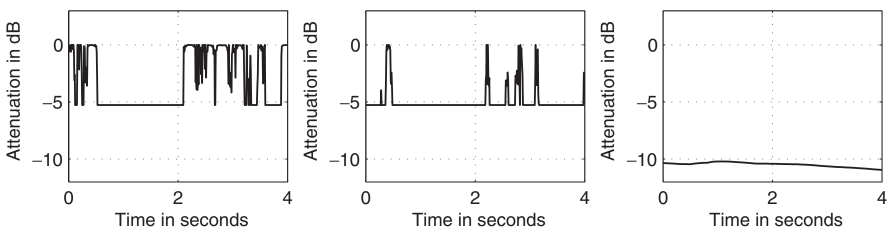

Fig. 15.18. Combined short and long-term noise reduction gain.

15.3.3 Ephraim-Malah Based, Short-Term Smoothed Noise Reduction Methods

An alternative approach to the above outlined Wiener-based noise reduction procedures is the MMSE (Minimum Mean Square Estimation) spectrum amplitude estimator which was initially proposed by Ephraim and Malah [16].

This single channel noise reduction framework is depicted in Fig. 15.19.

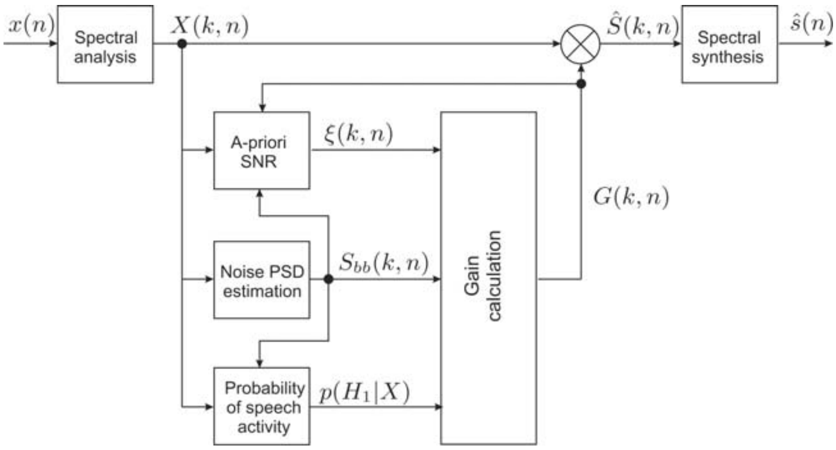

Fig. 15.19. Structure of an Ephraim-Malah based noise reduction method. After the spectral analysis, first the noise power spectral density $S_{bb}(k,n)$ has to be estimated. Then, the a-priori SNR $\xi(k,n)$ is estimated. Optionally, also the probability of speech activity $p(H_1|X)$ may be considered.

First the power spectral density $S_{bb}(k,n)$ of the background noise has to be estimated, e.g., by the Minimum Statistics approach. Then the a-priori SNR is estimated, e.g. by the *Decision directed* approach. Additionally, according [36,43,59] the probability for speech activity may be incorporated, by the additional factor $p(H_1(k,n)|X(k,n))$.

Based on these three estimates the noise reduction gain $G(k,n)$ is determined according to

$$\begin{aligned} G(k,n) &= \frac{\xi(k,n)}{1+\xi(k,n)} \exp\left[\frac{1}{2}\int\limits_{v(k,n)}^{\infty} \frac{\exp(-z)}{z}\,dz\right] \\ &\quad \cdot p\big(H_1(k,n)\big|X(k,n)\big), \qquad (15.10) \\ \text{with: } v(k,n) &= \frac{\xi(k,n)}{1+\xi(k,n)}\gamma(k,n);\;\; \gamma(k,n) = \frac{|X(k,n)|^2}{S_{bb}(k,n)}. \end{aligned}$$

For the deriving the calculation formula for the filter weights $G(k,n)$ according to Eqn. 15.10, the knowledge of the distribution of the real and imaginary parts of the speech and noise components is required. They are often assumed as Gaussian [16].

This assumption holds for many noise signals in everyday acoustic environments, but it is not exactly true for speech. A performance investigation for the application in hearing aids can be found, e.g., in [42]. More appropriate models for speech are mentioned in the next section.

15.3.4 Future Trends

So far, the application of well-known noise reduction methods for hearing aids has been explained. Now, we want to outline some methods and ideas for further enhancing the quality of noise reduction.

A big problem of noise reduction procedures is addressed by the first proposal: The estimation of the noise PSD. The basis of this proposal is to utilize both hearing aids on each side of the head for obtaining a more reliable noise PSD estimate, in particular during speech activity. The theoretical basis is the cross-correlation property of the signals of both hearing aids [14]. It is different for speech and noise components. Due to the diffuse character of noise, its components are less correlated than speech components, especially for high frequencies.

The calculation of the cross-correlation requires a full rate audio signal transmission between both hearing aids. When only lower data rate transmission is possible, also some binaural enhancements are possible: Supposing a voice activity detector is utilized for determining the time instances when the noise PSD is preferably estimated, a combined and more reliable activity detector can be obtained by logically combining the detection results of both sides.

As mentioned before, another possibility for a better noise reduction methods is to further advance the Ephraim-Malah noise reduction method by utilizing more appropriate models for the probability density of speech than the Gaussian model. One possibility is to utilize supergaussian statistical modelling for the speech DFT coefficients [32, 40, 41]. Noise reduction algorithms based on this modified estimator outperform the classical approaches using the Gaussian assumption. The noise reduction effect can be increased at an equal target signal distortion level. A computationally efficient realization has been published [32, 33] which allows a parametrization of the probability density function for speech spectral amplitudes so that an implementation in hearing aids is feasible in the near future.

Also model based noise reduction methods such as proposed in Chapter 10 are a promising idea in particular for the enhancement of speech. Since the proposed approach is optimized for car noise as disturbing signal, it has to be further generalized for other kinds of noise signals.

However, independent of the different rules for calculating the filter weights, the estimation quality of the power spectral density shows the strongest impact on the noise reduction quality. Since hearing impaired people wear their hearing aids during the whole day they are very sensitive to signal distortion which is therefore a more critical issue compared to noise reduction for hands-free telephones. For strictly avoiding desired signal distortion, for all noise reduction methods the noise attenuation has to be strongly limited. Unfortunately, for most short-term noise reduction approaches, alternative to the Wiener-filter, the gain of the acceptable noise reduction limit is not very high but has to be paid by a strongly increased computational complexity.

15.4 Multi-Band Compression

Whereas most signal processing algorithms in hearing aids can also be useful for normal hearing (e.g. noise reduction in telecommunications), multi-band compression directly addresses the individual hearing loss. A phenomenon typically observed in sensorineaural hearing loss is "recruitment" [60], which can be measured by categorical loudness scaling procedures (e.g. "Würzburger Hörfeld" [25]) and also could be demonstrated in physiological measurements of basilar membrane velocity [55]. Fig. 15.20 shows the growth of loudness as a function of level for a typical hearing impaired listener in comparison to the normal hearing reference.

With increasing frequency the level difference between normal and hearing-impaired listeners for soft sounds (< 10 CU; CU = Categorical Loudness Unit) increases, whereas curves cross at high levels. The arrows in the right bottom graph indicate the necessary level dependent gain to achieve the same loudness perception at 4 kHz for normal and hearing-impaired listeners. Thus, this measurement directly calls for the need of a frequency specific and level dependent gain - if loudness shall be restored to normal. Since more gain is needed for low input levels than for high input levels, the resulting input-output curves of an appropriate automatic gain control (AGC) system have a compressive characteristic.

Restoration of loudness - often also called "loudness normalization" - has been shown, both theoretically [10] and empirically [56], to be capable of also restoring temporal and spectral resolution (as measured by masking patterns) to normal. However, despite many years of research related to loudness normalization [31, 60], the benefits of this approach are difficult to prove [45]. Thus, over the years, many alternative rationales and design goals have been developed resulting in a large variety of AGC systems.

15.4.1 State-of-the-Art

Practically every modern hearing aid employs some form of AGC. The first stage of a multi-band AGC is a spectral analysis. In order to restore loudness,

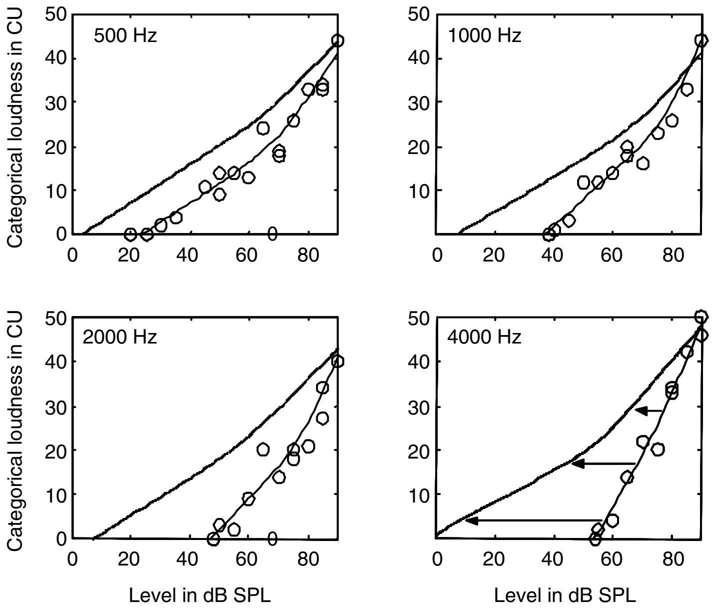

Fig. 15.20. Loudness as a function of level for a hearing-impaired listener (right curve, surrounded by circles) and normal listeners (left curve).

this spectral analysis should be similar to the human auditory system (for details see [65]). Therefore, often non-uniform filterbanks are used: constant bandwidth of about 100 Hz up to 500 Hz and approximately 1/3-octave filters above 500 Hz. In each channel the envelope is extracted as input to the nonlinear input-output function.

Depending on the time constants used for envelope extraction, different rationales can be realized. With very slow attack and release times (several seconds) the gain is adjusted to varying listening environments. These systems are often referred to as *automatic volume control* (AVC), whereas systems with fast time constants (several milliseconds) are called "syllabic compression" as they are able to adjust the gain for vowels and consonants within a syllable. For loudness normalization (also of time varying sounds) gains must be adjusted quasi-instantaneously, i.e., the gains follow the magnitude of the complex band pass signals. Moreover, combinations of both slow and fast time constants ("dual compression") have been developed [57].

To avoid a flattening of the spectral structure of speech signals - which is regarded to be important for speech intelligibility - neighboring channels are coupled or the control signal is calculated as a weighted sum of narrowband and broadband level [57]. The input-output function (see component in Fig. 15.21) calculates a time-varying gain which is multiplied by the band pass signal or the magnitude of the complex bandpass signal prior to the spec-

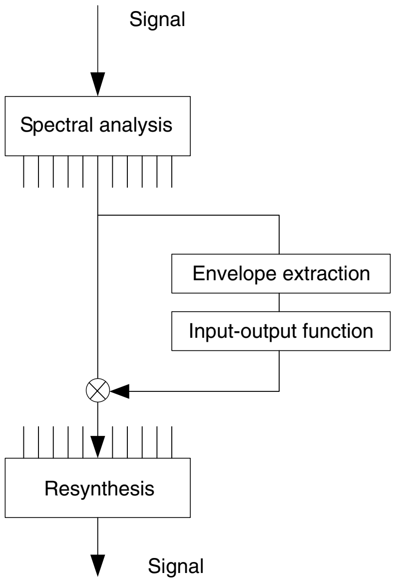

Fig. 15.21. Signal-flow for multi-band AGC processing.

tral resynthesis stage. There are many rationales to determine the frequency specific input-output functions from an individual audiogram, e.g. loudness restoration (see above), restoration of audibility (DSL i/o [11]) or optimization of speech intelligibility without exceeding normal loudness (NAL-NL1 [9]). The optimum rationale usually depends on many variables like hearing loss, age, hearing aid experience and actual acoustical situation.

Whereas the above mentioned AGC systems branch off the control signal before the multiplication of bandpass signal by nonlinear gain ("AGC-i"), output controlled systems ("AGC-o") get the control signal afterwards. AGC-o is often used to ensure that the maximum comfortable level is not exceeded and is thus typically implemented subsequent to an AGC-i. Recently, an AGC-o system has been proposed which is based on percentile levels and keeps the output not only below a maximum level but also above a minimum level in order to optimize audibility [37].

15.4.2 Future Trends

A possibility to cope with situation dependent fitting rationales is to control the AGC parameters (e.g. attack and release time, input-output function) by the classifier. In a situation where speech intelligibility is most important, e.g. a conversation in a crowded restaurant, the appropriate parameters for realizing NAL-NL1 are loaded, whereas when listening to music a setting with optimized sound quality is activated. A wireless link between hearing aids might be beneficial to synchronize the settings on both sides in order to avoid localization problems.

Another promising scenario is to implement psychoacoustic models (e.g. speech intelligibility, loudness, pleasantness) and use them for a continuous and situation dependent constrained optimization of the AGC parameters or directly of the time-varying gain. The latter can be realized by estimating the spectra of noise, speech and the composite signal block by block, similar to the Wiener-filter approach. The speech and noise spectra are used to calculate speech intelligibility (e.g. according to the SII [2]), whereas the overall spectrum is used to determine the current loudness (e.g. according to [10]). Then the channel gains are optimized for each block with the goal to maximize speech intelligibility and the constrained that the aided loudness for the individual hearing impaired listener does not exceed the unaided loudness for a normal listener. In this case, the hearing aid setting is not optimized for the average male speaker in a quiet surrounding (as is done with NAL-NL1), but for the individual speaker in the given acoustical situation.

15.5 Feedback Cancellation

Acoustic feedback (”whistling”) is a major problem when fitting hearing aids because it limits the maximum amplification. Feedback describes the situation when output signal components are fed back to the hearing aid microphone and are again amplified. In cases where the hearing aid amplification is larger than the attenuation of the feedback path, and the feedback signal is in phase, instabilities occur and whistling is provoked. The feedback path describes the frequency response of the acoustic coupling between the receiver and the microphones as depicted in Fig. 15.22.

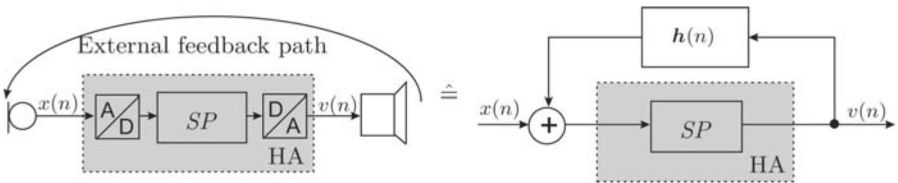

Fig. 15.22. On the left, the acoustic coupling between the hearing aid output and its microphone is shown and on the right the corresponding signal model where the acoustic path is modelled as a FIR filter with impulse response $\boldsymbol{h}(n)$. (HA: hearing aid).

Increasing the ear mold venting or even using open-fitting hearing aids, is more and more preferred by hearing aid users. The reason is that the occlusion effect [12] is usually reduced and the open fitting hearing aids are very comfortable to wear. However, increasing the vent diameter or even using open fitting hearing aids automatically increases the feedback risk and lowers the achievable amplification of the hearing aid. Therefore, well-performing feedback cancellation systems are becoming more and more important.

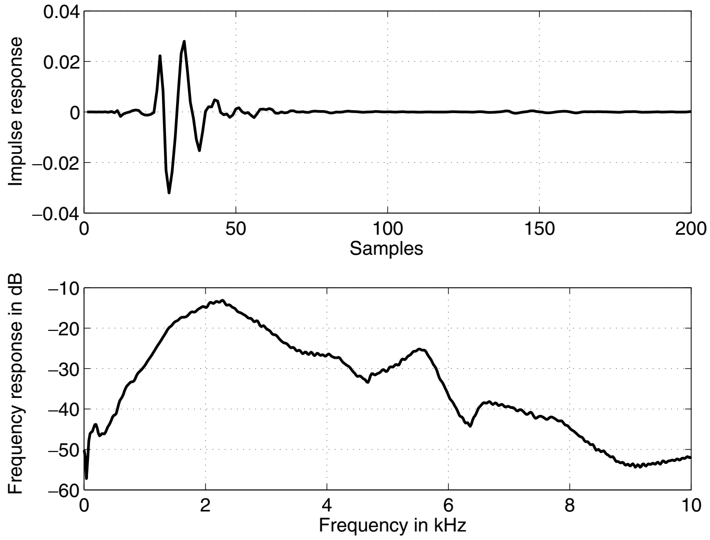

Fig. 15.23. Impulse (top) and frequency (bottom) responses of a typical hearing aid feedback path sampled at 20 kHz.

A typical hearing aid feedback path is depicted in Fig. 15.23. Here, one can observe that generally the paths exhibit a band-pass characteristic with the highest amount of coupling at frequency components between 1 and 5 kHz. The typical length of feedback paths which has to be modelled be a feedback cancellation system is approximately 64 coefficients long for a sampling rate of 20 kHz. Additionally, the current feedback path is highly dependent on many parameters of which the three most important are:

- the type of the hearing aid: BTE (behind-the-ear) or ITE (in-the-ear),
- the vent size,
- obstacles around the hearing aid (hands, hats, telephone receivers),
- the physical fit in the ear canal and leaks from jaw movements.

The first two parameters are static whereas the third is highly time-varying during the operation of the hearing aid. In Fig. 15.24 the variance of the feedback paths can be observed in dependence for the above given parameters.

Corresponding to the time-dependent or static parameters, fixed and dynamic measures are utilized in today's hearing aids to avoid feedback.

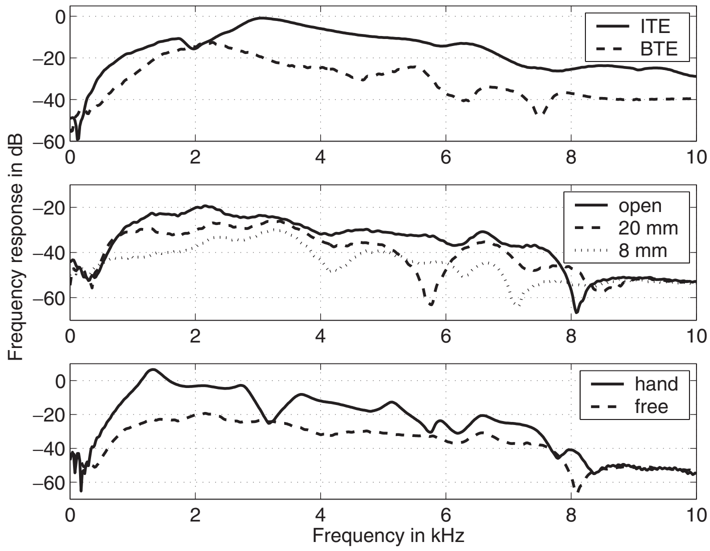

Fig. 15.24. Typical feedback paths for different types of hearing aids (top), different vent sizes (middle), and obstacles, i.e. a hand near the hearing aid compared to the normal situation (bottom).

A static method is to measure the normal feedback path (without obstacles) once after the hearing aid has been fitted. Limiting the gain of the hearing aid so that the closed loop gain is smaller than one for all frequency components, generally can prevent feedback.

Nevertheless, a totally feedback-free performance of the hearing aid can usually not be obtained without additional measures, especially when the closed-loop gain of the hearing aid in normal situations is close to one. Reflection obstacles such as a hand may then provoke feedback. To avoid this, dynamic methods are necessary for cancelling feedback adaptively when it appears.

For these dynamic measures, two methods are widely spread:

1. Selectively attenuating the frequency components for which feedback occurs is utilized in today's hearing aids. This method is normally efficient to avoid feedback. However, it is equivalent to a narrow-band hearing aid gain reduction.
2. Another method is the feedback compensation method where the feedback path is modelled with an internal filter in parallel to the feedback path and which subtracts the feedback signal. Thus, the hearing aid gain is not affected by this method. Additionally, it even allows hearing aid gain settings with closed-loop gains larger than one. This method is currently becoming state-of-the-art for hearing aids.

15.5.1 Feedback Suppression: Dynamic and Selective Attenuation of Feedback Components

An effective and selective attenuation of feedback components can be reached by notch filters. These notch filters are generally characterized by three parameters: the notch frequency, the notch width and the notch depth. It is most important to choose the appropriate notch frequency, i.e. when feedback occurs, the feedback frequency has to be determined fast and precisely.

Different methods, in the time and frequency domains, are applicable for the estimation of the feedback frequency. These are comparable to methods which can also be found for pitch frequency estimation [63]. These methods are, e.g., the zero crossing rate, the autocorrelation function and the linear predictive analysis. Most important is the fast reaction to feedback but also to apply the notch filters only where and as long as necessary in order to minimize the negative effect of the reduced hearing aid gain.

15.5.2 Feedback Compensation

The reduced hearing aid gain can be totally avoided by the compensation approach. Here, a filter is internally put in parallel to the external acoustic feedback path, as shown in Fig. 15.25. The output of the filter models the feedback signal.

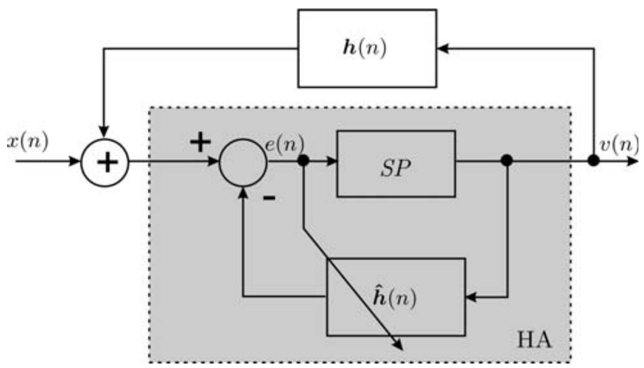

Fig. 15.25. General setup of a feedback cancellation system with SP modeling the hearing aid signal processing, $\boldsymbol{h}(n)$ the external feedback path, $\hat{\boldsymbol{h}}(n)$ the adaptive filter.

The challenge of this approach is to properly estimate the external feedback path with an adaptive filter. This is hard to realize due to the correlation of the input signal and the signal which is acoustically fed back to the microphones. For reliable estimates of the feedback path, the adaptation has to be controlled by sophisticated methods.

Adaptive algorithms generally estimate the filter coefficients, based on an optimization criterion. The criterion which is very often utilized is the minimization of the mean square error signal, i.e., the signal $e(n)$ after the subtraction of the adaptive filter's output signal. Writing $e(n)$ as

$$e(n) = x(n) + \sum_{l=0}^{N-1} \Big[h_l(n) - \hat{h}_l(n)\Big]\, v(n-l), \tag{15.11}$$

where the adaptive filter is assumed to model the complete feedback path of length N, and deriving the mean square error $\mathrm{E}\{e^2(n)\}$ with respect to $\hat{h}_l(n)$, one obtains the following relation:

$$\mathrm{E}\left\{e(n)\left[v(n-\nu) - \sum_{l=0}^{N-1}\left[h_l(n) - \hat{h}_l(n)\right]\frac{\partial v(n-l)}{\partial \hat{h}_\nu(n)}\right]\right\} \overset{!}{=} 0 \tag{15.12}$$

$$\forall\, \nu \in [0, N-1] \tag{15.13}$$

Under the assumption that the adaptive filter is nearly converged to the feedback path, one obtains the well-known orthogonality theorem:

$$\mathrm{E}\big\{v(n-l)\, e(n)\big\} \overset{!}{=} 0 \qquad \forall\, l \in [0, N-1]. \tag{15.14}$$

Writing Eqn. 15.11 in vector notation as

$$e(n) = x(n) + \Big[\boldsymbol{h}(n) - \hat{\boldsymbol{h}}(n)\Big]^T \boldsymbol{v}(n) \tag{15.15}$$

with $\boldsymbol{v}(n) = [v(n), \ldots, v(n-N+1)]^T$, $\hat{\boldsymbol{h}}(n) = [\hat{h}_0(n), \ldots, \hat{h}_{N-1}(n)]^T$ and $\boldsymbol{h}(n) = [h_0(n), \ldots, h_{N-1}(n)]^T$ and deriving the mean square error with respect to $\hat{\boldsymbol{h}}(n)$, one obtains the following equation:

$$\boldsymbol{r}_{x\boldsymbol{v}}(n) + \boldsymbol{R}_{\boldsymbol{vv}}(n)\left[\boldsymbol{h}(n) - \hat{\boldsymbol{h}}(n)\right] = \left[0, \ldots, 0\right]^T, \tag{15.16}$$

with the cross-correlation vector $\boldsymbol{r}_{x\boldsymbol{v}}(n) = \mathrm{E}\{x(n)\,\boldsymbol{v}(n)\}$ and the autocorrelation matrix $\boldsymbol{R}_{\boldsymbol{vv}}(n) = \mathrm{E}\{\boldsymbol{v}(n)\boldsymbol{v}^T(n)\}$, respectively.

Resolving this equation with respect to $\hat{\boldsymbol{h}}(n)$, it becomes obvious that the optimum solution which minimizes the mean square error shows a bias compared to the true feedback path:

$$\hat{\boldsymbol{h}}_{opt}(n) = \boldsymbol{h}(n) + \boldsymbol{R}_{\boldsymbol{vv}}^{-1}(n)\,\boldsymbol{r}_{x\boldsymbol{v}}(n). \tag{15.17}$$

The second term $\boldsymbol{R}_{\boldsymbol{vv}}^{-1}(n)\,\boldsymbol{r}_{x\boldsymbol{v}}(n)$ distorts the input signal $x(n)$ as the signal $v(n)$ is filtered such that all predictable components of $x(n)$ are subtracted, i.e. $x(n)$ is whitened. For an alternative derivation of correlation effects, see e.g. [58].

To demonstrate the relations, simulations were performed where the *SP* block of Fig. 15.25 was simply set to a gain g. The filter $\hat{\boldsymbol{h}}(n)$ was adapted under three different conditions:

1. for a white input signal,
2. for a colored input signal with the external feedback path turned to zero: $\boldsymbol{h} = [0, \dots, 0]^T$, and
3. for a colored input signal with an activated model of the external feedback path.

For the feedback path, a very simple model was used with $\boldsymbol{h} = [\ 0\ 0\ 0\ 1\ -0.6\ 0.1\ -0.3\ -0.2]^T$. The colored input signal was generated by a MA (moving average) process: $x(n) = u(n) + \sum_{l=0}^{L} a(l)\, u(n-l-1)$, with a white signal $u(l)$ and $L = 20$.

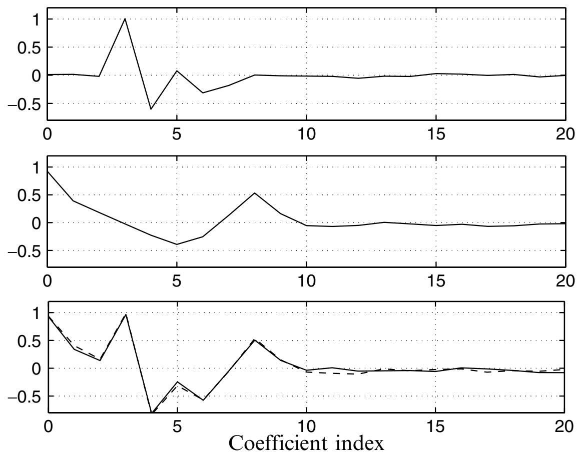

Fig. 15.26. Results for $\hat{\boldsymbol{h}}(n)$ for white (above) and colored excitation with external feedback path off (middle) and on (below). In the lower graph it is shown that the filter (solid line) nearly converges to the sum of the upper and middle graph (dashed line).

The results are depicted in Fig. 15.26. For the white input signal, the filter $\hat{\boldsymbol{h}}(n)$ adapts – as desired – to the feedback path (upper graph). When the feedback path is turned off and the colored signal is used as input, however, the filter acts as a decorrelation filter: If the *SP* block simply is a gain $g = 1$ the filter coefficients model the coefficients $a(l)$ of the input signal's model (middle graph). The result, which is obtained for the case when a colored signal is used to identify the feedback path, shows the superposition of both, the true feedback path and the FIR model of the input signal (lower graph).

Unfortunately, the last case corresponds to the general application for which the decorrelating effect of the feedback cancellation filter can hardly be avoided. This bias causes a distortion of the hearing aid output and has to be reduced as much as possible.

Thus, the main objective for enhancing the adaptation should be to reduce this correlation. Here, different methods exist [52]:

- Decorrelating the input signal with fast-adaptive decorrelation filters,
- delaying the output signal, or
- putting a nonlinear processing unit before the output stage of the hearing aid.

However, none of these methods is a straight-forward solution to the given problem, since many problems occur while implementing the proposals.

We made good experiences with three main settings:

- We reduce the step size, when music is detected as excitation signal,
- we utilize an internal feedback detector which allows a fast feedback reduction when suddenly the external feedback signal decreases, and
- we avoid gain settings of the hearing aid which provoke a closed loop gain setting strongly larger the critical gain.

The music detection is based on the decision of the classificator (see Sec. 15.6). In case when music is present as excitation signal, the risk of a correlated input and thus a biased adaptation of the filter is high. Therefore, to avoid this the step size is reduced. The drawback that the tracking of the filter is reduced can be accepted since when listening to music people usually move less and thus the risk of feedback provoked by feedback path changes is not very high.

The internal feedback detector steadily compares the input signal of the feedback cancellation system and the output signal of the adaptive filter which is the estimated feedback signal. In case the estimated feedback signal is larger than the input signal, this is a clear indication of a mis-adjustment of the adaptive filter. Either an increased step size or a complete reset of the filter coefficients can assure a fast readaptation of the filter coefficients. Usually this case occurs when a obstacle near the hearing aid (hand, telephone receiver, hat, etc.) which provokes a larger feedback path is suddenly removed.

Finally, one has to be aware of the limits of a feedback compensation system: The larger the hearing aid gain exceeds the critical gain, i.e. the gain when feedback occurs without feedback cancellation, the higher are the demands for the feedback compensator, and the more accurately the feedback path has to be estimated to avoid feedback. In other words, only slight mis-adjustment of the feedback path may already provoke strong feedback. This also has a direct impact to the adaption control: Only weak correlations of the input signal and thus a small bias of the estimated filter coefficients may provoke feedback in case of hearing aid gains that exceed the critical gain strongly, i.e. more than 10-15 dB.

15.6 Classification

Hearing aid users encounter a lot of different hearing situations in every day life, e.g. conversation in quiet or in noise, telephone calls, being in a theater or in road traffic noise. They expect real benefits from a hearing aid in each of the mentioned situations. As was shown in the previous part of this paper, modern digital hearing aids provide multiple signal processing algorithms and possible parameter settings, e.g. concerning directivity, noise reduction and dynamic compression. This portfolio of algorithms is expected to still grow with increasing IC computational power. Single algorithms and their multitude of possible parameter settings are mostly working in a situation specific way, i.e. these algorithms are beneficial in certain hearing situations whereas they have no or even negative impact in other situations. For example noise reduction algorithms as described in Sec. 15.3 reduce stationary background noise efficiently, whereas they may have some negative influence on the sound of music and should therefore be disabled in such situations. Even if the optimal signal processing algorithm for any relevant situation would be available, the problem to activate it reliably in the current specific hearing situation remains. A promising solution for this problem is to use a classification system, which can be understood as a superordinate, intelligent algorithm that continuously analyzes the hearing situation and automatically enables the optimal hearing aid setting. The alternative would be a great number of situation specific hearing aid programs, which have to be chosen manually. However, this approach would certainly overextend the mental and motor abilities of many hearing aid users, especially for the small ITE (in-the-ear) devices, and therefore, seems not to be a very attractive alternative [22].

15.6.1 Basic Structure of Monaural Classification

Fig. 15.1 shows the basic structure of a digital hearing aid with a superordinated classification system controlling the different signal processing blocks like directional microphone, noise reduction, shaping of the frequency response and dynamic compression. Classification systems consist of different functional stages:

As a first step, "features" are extracted from the microphone signal. "Features" are certain properties of the signals, whose magnitude is as different as possible for selected situation classes like "speech in quiet", "(speech in) noise" or "music" and can therefore be used to distinguish between situation classes. In literature several spectral and temporal features have been proposed, mostly in the context of separation of "speech in quiet" and "speech in noise": profile and temporal changes of the frequency spectrum [7, 17, 27], statistical distribution of signal amplitudes [35] or analysis of modulation frequencies [48].

To illustrate the principle of feature extraction, Fig. 15.27 shows the extraction of a modulation feature from three different signals belonging to the classes "speech in quiet", "speech in noise" and "music".

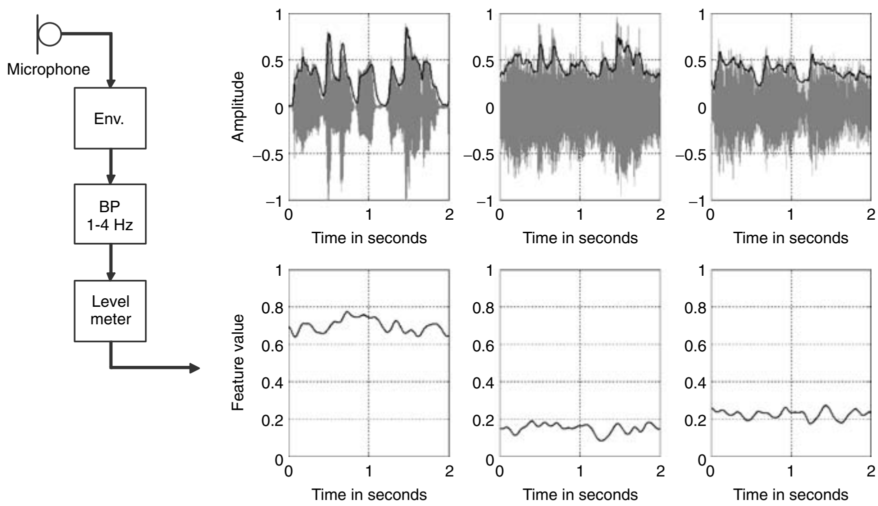

Fig. 15.27. Example for the calculation of an envelope-modulation feature.

The fluctuations of the signal envelope which are calculated by taking the absolute value and lowpass filtering are called "modulation". Typical for speech are strong modulations in the range of 1-4 Hz. The magnitude curves of this feature for the three examples as depicted in Fig. 15.27, show that values of this feature are obviously higher for "speech in quiet" than for the other signals. Consequently, the modulation feature allows to separate "speech in quiet" from "speech in noise" and "music", whereas separation of "speech in noise" and "music" is not possible due to similar feature values. Therefore, most applications of classification techniques require the simultaneous evaluation of a larger number of features to ensure sufficient decision reliability. The assignment of feature values and their combinations to the different classes can be achieved with standard approaches like the Bayes Classifier [48] or Neural Networks [17]. These algorithms learn the necessary a-priori knowledge about the relationship between feature values and situation classes in appropriate training procedures, which have to be based on large and representative databases of every-day life signals.

Fig. 15.28 demonstrates the performance of a classification system in a commercially available high-end hearing aid, which uses a Bayes classification system based on two envelope-modulation and a rhythm features to detect the four classes "speech", "speech in noise", "noise" and "music".

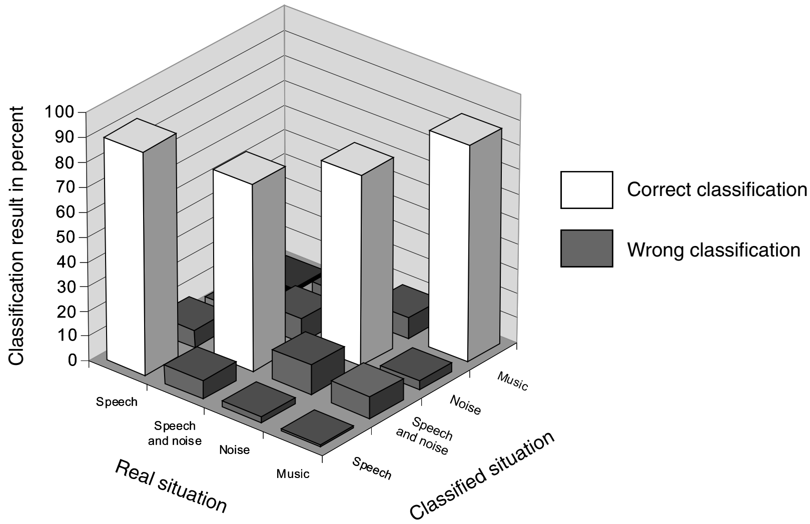

Fig. 15.28. Performance evaluation of a classification system of a modern hearing aid, based on classification of 15 hours of recorded hearing aid microphone signals comprising 500 hearing-situations in total.

Following the Bayes approach, for each detected value of the three features the probability of the four different classes is calculated. After summing up the probabilites across the three different features, the decision is made for the class with the highest cumulated probability. The underlying probability density functions, which are shown in Fig. 15.29, were derived in training with a large training database.

They can be implemented in hearing aids as look-up table or more efficiently in terms of hearing aid memory as polynomial approximations.

Every second a classification decision was made finally leading to the detection and error rates calculated for each of the four classes. Obviously, detection rates between 75 and 90 percent can be achieved, which have shown to be sufficient for a robust and beneficial control of the hearing aid signal processing. The perceptual influence of the misdetections can be reduced to a negligible level by nonlinear temporal averaging of the classification results and, as described in the next section, by smooth transitions between different operation states.

The adaption of the hearing aid signal processing to the detected listening situation is divided into two parts as shown in Fig. 15.1. The block "selection of algorithm and parameters" contains an "action matrix" describing which of the settings for the algorithms and parameters are optimal in each situation. The definition of the action matrix is based on detailed knowledge of the properties of the particular algorithms in the different situations. Extensive investigations and tests are the base for this knowledge. Every time the

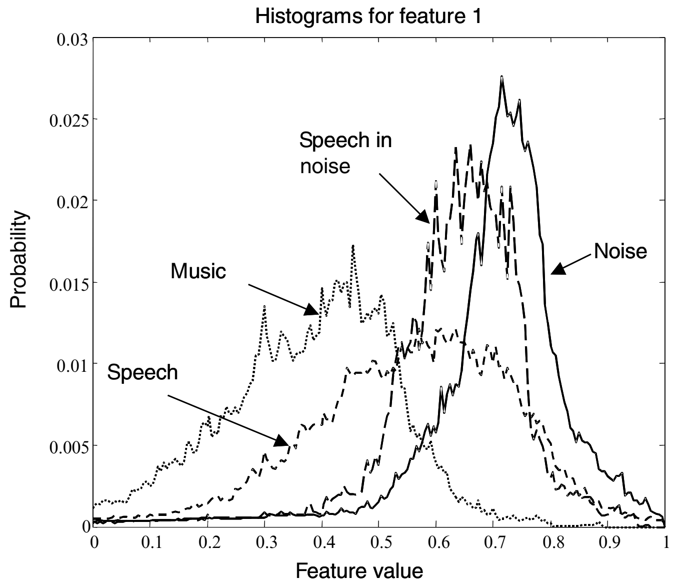

Fig. 15.29. Calculated probability density functions of an envelope-modulation feature in four different classes of hearing situations.

detected situation class is updated, the next block generates "on/off"-control signals for each hearing aid algorithm. Sudden "off/on"-switching of signal processing components like the directional microphone are considered as irritating and unpleasant. Thus, appropriate fading mechanisms which realize a gliding smooth transition from one state of operation to another are advantageous. In many cases, this can easily be achieved by low pass filtering of the control signals. Fig. 15.30 illustrates the fading from omnidirectional to directional microphone mode.

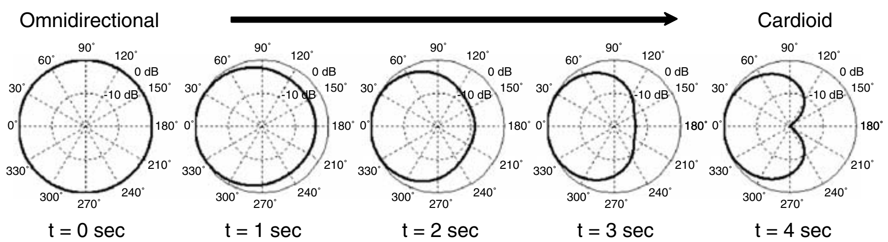

Fig. 15.30. Fading from omnidirectional to directional microphone mode

15.6.2 Binaural Classification

A problem in bilateral fittings of hearing aids with classification systems is that different classes can be detected in the left and right hearing aid resulting in different processing schemes. These differences, e.g. if the directional microphone is activated only on one side, can temporarily reduce the sound quality as well as the speech intelligibility and in addition to that, introduce artificial interaural time and level differences reducing the localization ability of the hearing impaired, which is mainly based on analyzing these signal cues [28].

Differences in classification results are mainly caused by head shading effects in asymmetrical hearing situation, e.g. a hearing situation with a music source on one side of the head and a talker on the other side, would lead to local classification decisions dominated by the ispilateral source, since the contralateral source is shaded, i.e. attenuated, by the head. Real-life evaluations with BTE (behind-the-ear) hearing aids showed that the percentage of asymmetrical classification results can reach up to 20 %. To overcome the problems described above, a binaural synchronization of the classification systems based on a bidirectional low-power wireless link between both hearing aids was introduced recently. In this realization both hearing aids first analyze the sound field independently, then exchange information of the local classification results and then follow exactly the same procedure in parallel to determine the global "binaural" class, see Fig. 15.1. Finally, both hearing aids are adapted synchronously to the signal processing and parameter settings prescribed for the common class. Doing so, the above mentioned disadvantages in unsymmetrical hearing situations can be avoided.

15.6.3 Future Trends

Using multi-microphone signals is the most important step from classification based on the statistical information of one microphone signal towards a future sound scene classification [49]. Typical situations where single-signal based classification systems fail are, for example, listening to music from the car radio while driving or conversing in a cafe with background music. To classify these situations correctly so that the algorithms can take advantage of the result requires information about the sound incidence direction, and the number, distance and type of sound sources in the room. This information can be derived from future multi-microphone localization and classification algorithms. Methods known from the Computational Auditory Scene Analysis (CASA) [6] can be used to further develop today's classification systems. For example, simultaneous speech sources in noisy environment can be recognized by pitch tracking [64].

15.7 Summary

The development of hearing aids covers a wide range of different signal processing components. They are mainly motivated by audiological questions. This chapter focuses on algorithms dealing with the compensation of the recruitment phenomenon, the improvement of speech intelligibility and the enhancement of comfort while using the hearing aid in everyday life.

As one important component of hearing aids, the directional microphone and its effect on the improvement of speech intelligibility is discussed. Directional microphones of different complexities are investigated starting with simple methods like first-order and second-order differential arrays. A description of a four-channel adaptive beamformer closes this topic.

One component which mainly focuses on the improvement of comfort is the noise reduction unit. Algorithms of different complexities, with different amounts of statistical a priori knowledge concerning the computed signal and different speeds of reaction are described. Noise reduction algorithms which exploit the binaural wireless link of future high-end digital hearing aids are discussed as well.

A significant unit in hearing aids is the AGC which compensates the recruitment phenomenon. This chapter discusses state-of-the-art systems and future trends.

Another important aspect is the feedback phenomenon which may occur at high levels of amplification in the hearing aid. This chapter presents two concepts to reduce feedback, namely the feedback compensation approach and the feedback suppression approach.

Finally, the ability of modern hearing aids to detect different hearing situations on the basis of binaurally coupled classification algorithms using a wireless link and to properly adapt to the optimal processing for the specific situation is discussed.

References

[1] J. Anemueller: *Across-Frequency Processing in Convolutive Blind Source Separation*, Univ. Oldenburg, 2001.
[2] *ANSI S3.5-1997*, Methods for calculation of the speech intelligibility index, 1997.
[3] J. Benesty, D. R. Morgan: A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation, *Transactions on Speech and Audio Processing*, **6**(2), 156-165, March, 1998.
[4] J. Benesty, S. Gay (Eds.): *Acoustic Signal Processing for Telecommunication,* Boston/Dordrecht/London: Kluwer Academic Publishers, 2000.
[5] M. Berouti, R. Schwarz, J. Makhoul: Enhancement of speech corrupted by acoustic noise, *Proc. ICASSP '79*, Tulsa, USA, 1979.

[6] A. S. Bregman: *Auditory Scene Analysis*, Cambridge, Massachusetts: MIT Press, 1990.
[7] M. Büchler, N. Dillier, S. Allegro, S. Launer: Klassifizierung der akustischen Umgebung für Hörgeräte-Anwendungen, *Proc. DAGA*, 282-283, Oldenburg, Germany, 2000 (in German).
[8] H. Buchner, R. Aichner, W. Kellermann, Y. Huang, J. Benesty: *Audio Signal Processing for Next-Generation Multimedia Communication Systems*, Blind Source Separation for Convolutive Mixtures: A Unified Treatment, Boston: Kluwer Academic Publishers, Feb., 2004.
[9] D. Byrne, H. Dillon, T. Ching, R. Katsch, G. Keidser: NAL-NL1 Procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures, *J. Am. Acad. Audiol.*, **12**, 37-51, 2001.
[10] J. Chalupper, H. Fastl: Dynamic loudness model (DLM) for normal and hearing-impaired listeners, *Acta Acustica united with Acustica*, **88**, 378-386, 2002.
[11] L. E. Cornelisse, R. C. Seewald, D. G. Jamieson: The input/output formula: A theoretical approach to the fitting of personal amplification devices, *J. Acoust. Soc. Am.*, **97**, 1854-1864, 1995.
[12] H. Dillon: *Hearing aids,* New York, Stuttgart: Boomerang Press, 2001.
[13] S. Doclo: *Multi-Microphone Noise Reduction and Dereverberation Techniques for Speech Applications*, Univ. of Leuven, ISBN 90-6582-409-0, 2003.
[14] M. Doerbecker, S. Ernst: Combination of two-channel spectral subtraction and adaptive wiener post-filtering for noise reduction and dereverberation, *Proc. EUSIPCO*, 995-998, Triest, Italy, 1996.
[15] G. W. Elko, A. N. Pong: A simple first order directional microphone, *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, 169-172, Mohonk, New Paltz, NY, 1995.
[16] Y. Ephraim, D. Malah: Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator, *IEEE Trans. on Acoustics, Speech, and Signal Processing*, **ASSP-32**(6), 1109-1121, 1984.
[17] F. Feldbusch: Geräuscherkennung mittels Neuronaler Netze, *Zeitschrift für Audiologie*, **1**, 30-36, 1998 (in German).
[18] S. Gannot, D. Burshtein, E. Weinstein: Signal enhancement using beamforming and non-stationarity with application to speech, *IEEE Trans. on Sig. Proc.*, **49**(8), 1614-1626, Aug., 2001.
[19] J. E. Greenberg, P. M. Zurek: Evaluation of an adaptive beamforming method for hearing aids, *JASA*, **91**(3), 1662-1676, 1992.
[20] L. J. Griffiths, C. W. Jim: An alternative approach to linearly constrained adaptive beamforming, *IEEE Trans. Antennas Propagat.*, **AP-30**, 2734, Jan., 1982.
[21] S. Gustafsson: *Enhancement of Audio Signals by Combined Acoustic Echo Cancellation and Noise Reduction*, RWTH Aachen, ABDN Band 11, Aachen, P. Vary (Hrsg. Verlag der Augustinus Buchhandlung), 1999.

[22] V. Hamacher, E. Fischer, I. Holube: Methods to classify listening situations, *UHA-Tagungsband*, 167-180, Nürnberg, Median Verlag, 2001.

[23] V. Hamacher: Algorithms for future commercial hearing aids, *Proc. ICA '04*, **2**, 1385-1388, Kyoto, Japan, April, 2004.

[24] V. Hamacher, J. Chalupper, J. Eggers, E. Fischer, U. Kornagel, H. Puder, U. Rass: Signal Processing in High-End Hearing Aids: State of the Art, Challenges, and Future Trends, *EURASIP Journal on Applied Signal Processing*, **2005**(18), 2915-2929, 2005.

[25] O. Heller: Hörfeldaudiometrie mit dem Verfahren der Kategorienunterteilung, *Psychologische Beiträge*, **27**, 478-493, 1985 (in German).

[26] I. Holube, V. Hamacher, M. Wesselkamp: Hearing instruments: Noise reduction strategies, *Proc. 18th Danavox Symposium: Auditory Models and Non-linear Hearing Instruments*, Kolding, Denmark, Danavox, Copenhagen, 1999.

[27] J. M. Kates: Classification of background noises for hearing-aid applications, *JASA*, 461-470, 1997.

[28] G. Keidser, K. Rohrseitz, H. Dillon, V. Hamacher, U. Rass, E. Convery, L. Carter, J. Mejia: The effect of multi-channel wide dynamic range compression, noise reduction, and directionality on horizontal localisation performance, *submitted to Ear & Hearing*, 2005.

[29] B. Kollmeier, J. Peissig, V. Hohmann: Binaural noise reduction hearing aid scheme with realtime processing in the frequency domain, *Scand. Audiol. Suppl.*, **38**, 28-38, 1993.

[30] M. Kompis, N. Dillier: Performance of an adaptive beamforming noise reduction scheme for hearing aid applications, *JASA*, **109**(3), 1123-1143, 2001.

[31] A. Leijon: Hearing aid gain for loudness-density normalization in cochlear hearing losses with impaired frequency resolution, *Ear & Hearing*, **12**, 242-250, 1990.

[32] T. Lotter, P. Vary: Noise reduction by maximum a posteriori spectral amplitude estimation with supergaussian speech modeling, *Proc. IWAENC '03*, 83-86, Kyoto, Japan, 2003.

[33] T. Lotter: *Single and Multimicrophone Speech Enhancement for Hearing Aids*, Aachener Beiträge zu Digitalen Nachrichtensystemen, **18**, ISBN 3-86130-645-X, 2004.

[34] C. Liu et al.: A two-microphone dual delay-line approach for extraction of speech sound in the presence of multiple interferers, *JASA*, **110**(6), 3218-3231, 2001.

[35] C. Ludvigsen: Schaltungsanordnung für die automatische Regelung von Hörhilfsgeräten, *Europäische Patentschrift, EP 0 732 036 B1* (in German).

[36] D. Malah, R.V. Cox, A.J. Accardi: Tracking speech-presence uncertainty to improve speech enhancement in non-stationary noise environments, *Proc. ICASSP '99*, **2**, 1761-1764, Phoenix, USA, 1999.

[37] L. F. A. Martin, P. J. Blamey, C. J. James, K. L. Galvin, D. Macfarlane: Adaptive dynamic range optimization for hearing aids, *Acoustics Australia*, **29**, 21-24, 2001.

[38] R. Martin: Spectral subtraction based on minimum statistics, *Proc. EUSIPCO '94*, 1182-1185, Edinburgh, Scotland, 1994.

[39] R. Martin: Noise power spectral density estimation based on optimal smoothing and minimum statistics, *IEEE Trans. on Speech and Audio Processing*, **9**(5), July, 2001.

[40] R. Martin: Speech enhancement using MMSE short time spectral estimation with gamma distributed priors, *Proc. ICASSP '02*, 504-512, Orlando, FL, USA, May, 2002.

[41] R. Martin, C. Breithaupt: Speech Enhancement in the DFT domain using Laplacian speech priors, *Proc. IWAENC '03*, 87-90, Kyoto, Japan, 2003.

[42] M. Marzinzik: *Noise Reduction Schemes for Digital Hearing Aids and Their Use for the Hearing Impaired*, Univ. Oldenburg, Medical Physics, Oldenburg, Germany, 2000.

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

[44] J. Meyer: *Beamforming für Mikrofonarrays unter Berücksichtigung von akustischen Streukörpern*, TU Darmstadt, 2001 (in German).

[45] B. C. J. Moore: *Perceptual Consequences of Cochlear Damage*, London: Oxford University Press, 1995.

[46] H. G. Mueller, M. Killion: An easy method for calculating the articulation index, *The Hearing Journal*, **43**(9), 14-17, 1990.

[47] M. Ostendorf, V. Hohmann, B. Kollmeier: Empirische Klassifizierung verschiedener akustischer Signale und Sprache mittels einer Modulationsfrequenzanalyse, *Proc. DAGA '97*, 608-609, Kiel, Germany, 1997 (in German).

[48] M. Ostendorf, V. Hohmann, B. Kollmeier: Klassifikation von akustischen Signalen basierend auf der Analyse von Modulationsspektren zur Anwendung in digitalen Hörgeräten, *Proc. DAGA*, 402-403, Zürich, Switzerland, 1998 (in German).

[49] V. Peltonen, J. Tuomi, A. Klapuri, J. Huopaniemi, T. Sorsa: Computational auditory scene recognition, *Proc. ICASSP '02*, **2**, 1941-1944, Orlando, FL, USA, 2002.

[50] T. A. Powers, V. Hamacher: Three microphone instrument is designed to extend benefits of directionality, *The Hearing Journal*, **55**(10), 38-45, 2002.

[51] H. Puder: *Geräuschreduktionsverfahren mit modellbasierten Ansätzen für Freisprecheinrichtungen in Kraftfahrzeugen*, TU Darmstadt, Darmstädter Dissertation D17, Fortschritt-Berichte VDI-Reihe 10(721), 2003 (in German).

[52] H. Puder, B. Beimel: Controlling the adaptation of feedback cancellation filters – problem analysis and solution approaches, *Proc. EUSIPCO '04*, Vienna, Austria, Sept., 2004.

[53] T. A. Ricketts: Impact of noise source configuration on directional hearing aid benefit and performance, *Ear and Hearing*, **21**(3), 194-205, 2000.

[54] T. A. Ricketts: Directional hearing aids, *Trends in Amplification*, **5**(4), 139-176, 2001.

[55] M. A. Ruggero, N. C. Rich: Furosemide alters organ of corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane, *J. Neurosci.*, **11**, 1057-1067, 1991.

[56] N. Sasaki, T. Kawase, H. Hidaka, M. Ogura, T. Takasaka, K. Ozawa, Y. Suzuki, T. Sone: Apparent change of masking functions with compression-type digital hearing aid, *Scand. Audiol.*, **29**, 159-169, 2000.

[57] T. Schneider, R. Brennan: A Multichannel compression strategy for a digital hearing aid, *Proc. ICASSP '97*, *1*, 411-414, Munich, Germany, 1997.

[58] M. G. Siqueira, A. Alwan: Steady-state analysis of continuous adaptation in acoustic feedback reduction systems for hearing-aids, *Transactions on Speech and Audio Processing*, **8**(4), 443-453, July, 2000.

[59] I.Y. Soon, N.K. Soo, C.K. Yeo: Improved noise suppression filter using self-adaptive estimator of probability of speech absence, *Signal Processing*, **75**, 151-159, 1999.

[60] J. Steinberg, M. Gardner: Dependence of hearing impairment on sound intensity, *Journal of the Acoustical Society of America*, **9**, 11-23, 1937.

[61] M. Valente, R. E. Sandlin: *The Textbook of Hearing Aid Amplification*, Use of Microphone Technology to Improve User Performance in Noise, 2nd edition, San Diego, California: Singular Publishing Group, 2000.

[62] J. Vanden Berghe, J. Wouters: An adaptive noise canceller for hearing aids using two nearby microphones, *JASA*, **103**, 3621-3626, 1998.

[63] H. L. Van Trees: *Detection, Estimation, and Modulation Theory*, part 1, New York: John Wiley and Sons, Inc., 1968.

[64] M. Wu, D. Wang, G. J. Brown: A multi-pitch tracking algorithm for noisy speech, *Proc. ICASSP '02*, **1**, 369-372, Orlando, FL, USA, 2002.

[65] E. Zwicker, H. Fastl: *Psychoacoustics: Facts and Models*, 2nd edition, Berlin, Heidelberg, New York: Springer, 1999.