Temporal-envelope based related processing in advanced digital hearing aids

Volkmar Hamacher, Eghart Fischer, and Henning Puder
Siemens Audiologische Technik GmbH, Gebbertstraße 125, 91058 Erlangen, Germany
email: {volkmar.hamacher; eghart.fischer; henning.puder}@siemens.com

Introduction

Temporal envelopes either of the full-band or of the respective sub-band audio signals exhibit typical characteristics which are directly related to the signal type and also related to the perceived sound quality. Additionally, these characteristics contain information which can be exploited for detection purposes. In this paper three different types of algorithms, noise reduction, compression, and classification are described. It will be shown that temporal envelope processing plays a major role in the optimum design of each of these algorithms.

The paper is organized as follows: In the first section compression methods are shortly motivated, their realization in modern hearing aids and tests with hearing impaired people are described. For noise reduction methods which will be presented in the next section the level of the signal envelopes can be utilized to determine the amount of noise reduction to be optimal inserted. In the succeeding section it will be shown that, besides other criteria, signal envelopes can be utilized for classification purposes, i.e. the distinction between different hearing situations. Finally, the last section concludes the paper.

Compression Algorithms

Compression algorithms have become standard methods for today's hearing aids. Their aim is to compensate for the recruitment [3] of hearing impaired people, i.e. especially soft sounds cannot be perceived, whereas loud sounds usually are observed comparably to normal hearing people. The compensation of this recruitment requires level and frequency dependent amplification, i.e. the hearing aid gain is reduced for high input signal levels. For this, compression methods in sub-bands are generally used which usually contain two processing stages: First, the levels of the different sub-bands of the input signal are determined and then the gains of the sub-bands have to be adjusted dependent on the input signal level.

Generally, three parameters determine the behaviour of compression algorithms:

1. The number of sub-bands for which independent compression ratios are determined,
2. the smoothing-time constant for the estimation of the input levels, and
3. the compression ratio and knee-points which characterize the input-output level relation.

From the physiological point of view [6] the best compensation for the recruitment problem should be an instantaneous compression, i.e. fast smoothing constants are utilized for the level estimation. An often made experience with hearing impaired people is, however, that compression methods are the better accepted the stronger, i.e. the slower is the smoothing of the estimated input signal power levels and the lower is the number of sub-bands in which the compression is independently performed .

A common interpretation of this phenomenon [5] is that the modification of the signal envelope, which is a result of the fast compression, gives a negative subjective quality impression. This concerns both, the envelopes of the different sub-bands and the envelope of the full-band signal. The sub-band envelopes are the stronger modified, the faster is the signal power smoothing. The full-band signal envelope is additionally modified the larger the number of sub-band signals is.

Noise Reduction

Signal envelopes also play an important role for the design of noise reduction algorithms. The signal envelopes can be characterized with the help of the modulation spectrum [4], which is – generally speaking – the spectrum of the signal envelope. It is well known that speech signals exhibit the highest modulation spectrum at 4 Hz whereas typical noise shows the strongest components at higher modulation frequencies. Early noise reduction approaches exploited this relation and determined the modulation spectrum around 4 Hz. Performing this analysis in different sub-bands of the audio signal, the sub-band signals can be independently attenuated based on the level of the modulation spectrum (s. Fig. 1). Thus, frequency components with a low SNR are attenuated which results in an SNR improvement of the full-band signal.

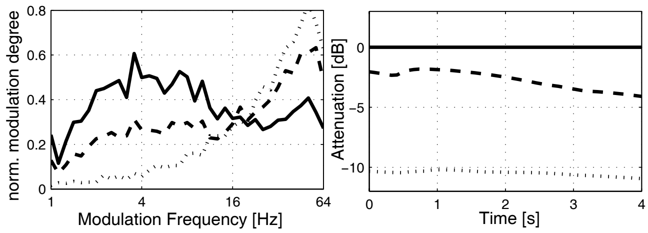

Figure 1: Modulation spectra (left) and corresponding attenuation in dependence of time (right) for one sub-band. The three graphs correspond to clean speech (solid), noisy speech (dashed), and noise (dotted).

However, these early methods show the disadvantage that only long-term smoothed attenuation is applied to the different sub-band signals. The widely spread solution to this problem is to utilize fast acting filters such as Spectral Subtraction or Wiener filtering [2] which try to reconstruct the time-dependent power spectral densities of the usually non-stationary desired signals.

Properly designed, the signal attenuation directly follows the signal envelope. The larger the signal envelope exceeds the generally slowly changing noise power, the lower is the applied attenuation of the noise reduction filter what is depicted in Fig. 2.

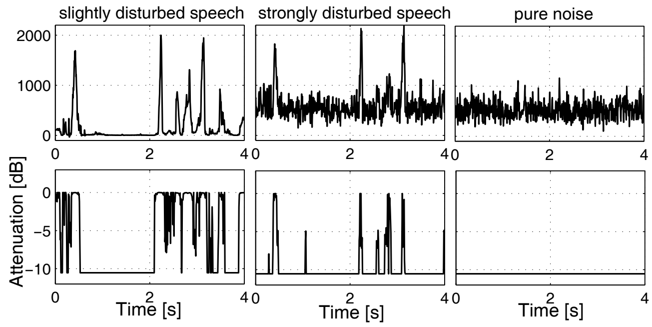

Figure 2: Above: signal envelopes for slightly (left), moderately (middle) disturbed speech, and noise. Below: corresponding attenuation values of the Wiener filter.

Thus, again the envelope plays a crucial role for the algorithm design, as the Wiener filter can also be interpreted as a method to reconstruct the signal envelope. Especially important is that the power spectral density of the input signal properly follows the signal envelope. Estimates which are too fast smoothed produce an output signal of the Wiener filter which show signal artifacts which are generally known ans "musical tones" [1]. In contrast, too strong smoothing results in reverberant output signals. Thus, the better the clean signal sub-signal envelopes are reconstructed the better is the performance of the noise reduction method.

Classification Algorithms

The aim of classification is to distinguish between different everyday-life situations such as *speech*, *speech in noise*, *noise*, *speech in car*, *music*, *telephone*, *TV*.

The classification results may then be applied for the control of adaptive signal processing algorithms especially to chose or switch between different parameter sets. Examples for this are the choice of the situation-optimum frequency response, the control of directional microphones, the activation and the choice of the strength of noise reduction.

Classification systems usually consist of complex feature extraction and decision methods. For the latter, many methods are possible, such as Bayes detection and neural networks. One feature that is already implemented as one part of a classification system being part of commercially available hearing aids is the level of the modulation frequency at e.g. 4 Hz, which was already described in Section *Noise Reduction* (s. Fig 1). To obtain a classification based on the – typically more than one – chosen features, a decision system is necessary. For this, a Bayes decision can be utilized [7]. This detection method requires the conditional probabilities of the different classed which are a-priori unknown and have to be estimated in advance. This requires a sophisticated training stage. The results which can be obtained with such a simple, only envelope-based classification system are depicted in the graph of Fig. 3.

Here, for the three classes to be detected, the probability of the correct detection (black) and the false detection probabilities of the other classes are depicted (gray).

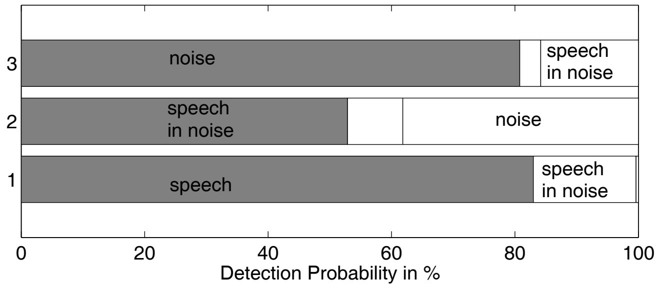

Figure 3: Classification results for the classes: 1. clean speech, 2. noisy speech, and 3. noise. The true detections are depicted in gray.

It can be observed that with this simple criterion a rather good distinction between *speech* and *noise* is possible. For the distinction between *speech in noise* and *noise*, however, a reduced reliability is observed. The reason is that with increasing noise, the level of the modulation spectrum at 4 Hz reduces continuously not instantaneously with the noise level. This makes is difficult to determine a fixed threshold for the classification. Here the additional features are necessary which can reduce the overall false detection probability.

However, the negative impact of wrong decision results, can be reduced by not instantaneously switching between different parameter sets when the detected class changes, but performing an continuous switch over time. Thus, the perceptive impact is strongly reduced.

Conclusions

In this paper three different temporal-envelope based signal processing methods were described. It could be shown that the envelopes play a major role when designing hearing aid algorithms. The overall goal for compression and noise reduction algorithms should be to preserve or reconstruct the full-band and sub-band signal envelopes whenever possible. Additionally, it could be shown that the envelope can also be utilized for signal classification purposes.

References

[1] O. Cappé: *Elimination of the Musical Noise Phenomenon with the Ephraim and Malah Noise Suppressor*, IEEE Trans. on Speech and Audio Processing, vol. 2, no. 2, pp. 345-349, 1994

[2] S. Haykin: *Adaptive Filter Theory*, Third Edition, Prentice Hall Inc., 1996

[3] B.C. Moore: *Perceptual Concequences of Cochlear Damage*, Oxford Medical Publications, 1995

[4] M. Ostendorf, V. Hohmann, B. Kollmeier: *Empirische Klassifizierung verschiedener akustischer Signale und Sprache mittels einer Modulationsfrequenzanalyse* DAGA-97, p. 608-609, 1997

[5] R. Plomb: *Noise, amplificaton, and compression: considerations of three main issues in hearing aid design*, Ear & Hearing, 15(1), 2-12, 1994

[6] 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., Vol. 11, 1057-1067, 1991

[7] H.L. Van Trees: *Detection, Estimation, and Modulation Theory*, Part I, John Wiley and Sons, Inc., New York, 1968