A MULTICHANNEL COMPRESSION STRATEGY FOR A DIGITAL HEARING AID

Todd Schneider *Robert Brennan*
Unitron Industries Ltd., Kitchener
Ontario, Canada N2G 4X1

ABSTRACT

Multi-channel compression schemes are a practical method of mapping the wide dynamic range of speech signals into the reduced dynamic range of hearing impaired listeners. These systems address two of the shortcomings of single-channel compression systems: (1) the reduction of gain as a result of narrow-band non-speech stimuli and (2) the reduction of gain that often occurs when high-frequency speech components are followed by intense low-frequency speech components. They also provide frequency-dependent compression ratios that are needed by many newer supra-threshold fitting strategies (e.g., DSL I/O). This paper presents a multichannel compression scheme that employs an oversampled, polyphase DFT filterbank. In each compressor channel, the gain is controlled by an adjustable combination of a overall, dual time-constant input signal level and the individual channel signal level that is measured with a short time-constant RMS detector. Informal listening tests have demonstrated that the design has very good audio quality and performs well in real-world listening situations. The design is suited for low-power, real-time operation.

1. INTRODUCTION

For many years, dynamic range compression has been widely used in hearing aids to compensate for the reduced dynamic range of hearing impaired listeners. Dillon [1] has recently pointed out that there are a wide range of compression schemes (many of which offer opposing rationale for their processing strategies) that have been developed in the past. For the most part, these have been single-channel, feedback compression systems.

1.1. Topologies

A feedback compression system uses the output signal from the hearing aid (taken either before or after the output volume control) to control the gain (Figure 1(a)). While feedback compression designs are easier to design because the feedback tends to linearize the control loop [2], they permit overshoots to occur at the output of the hearing aid because a signal level must reach the hearing aid output before the gain can be reduced to compensate. Feedforward designs (Figure 1(b)) overcome this limitation by using the input signal level to control the gain. They can act to suppress a transient before it reaches the output.

Feedforward designs are also superior to feedback designs because they can achieve infinite compression (limiting) with finite side-chain gain [2]. For large compression ratios, a feedback compressor must operate at almost open-loop which can lead to stability problems. Feedforward designs have the disadvantage that the transfer characteristics and stability of the side-chain must be precisely controlled. This presents a challenge for analog implementation but is relatively easy to achieve in a DSP implementation.

1.2. Single- versus Multi-channel Processing

Single-channel compression systems (especially compression limiters) are commonly used in hearing aids because they are relatively simple to implement and address the basic comfort issues

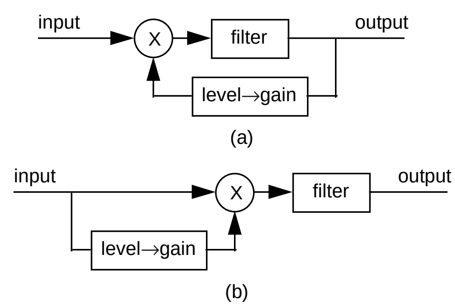

Figure 1. Basic topologies for compression systems: (a) feedback and (b) feedforward.

associated with loud sounds by limiting high-level output signals to less than a hearing aid wearer's uncomfortable loudness (UCL). Newer compression schemes [3] use an approach known as wide dynamic range compression (WDRC) that maps the input signal range (above a low compression threshold, typically 40-45 dB SPL) into the residual dynamic range of a hearing impaired listener.

When driven by a broadband, dynamic signal (e.g., speech), the amount of compression at a given frequency provided by a single-channel compression system (one that uses a single control signal to control one or more channels) is dependent on the spectral content and the signal level dynamics. Because the compression provided at a given frequency is controlled by a level representing the entire input bandwidth, an assumption must be made about the shape of the long-term spectral shape of the input signal and this presumed long-term spectral shape is mapped based on an overall level estimate. Unfortunately, the short-term spectral shape of a speech signal usually differs significantly from the long-term spectral shape (which is dominated by the low-frequency speech components because they are longer in duration and more intense than the high-frequency speech components). Thus, high-frequency energy is uncorrelated with the control signal which results in a much lower *effective* compression ratio than is desired.

Multi-band compression schemes (that employ separate control signals for each band) can overcome this difficulty because the compression in a given band can be controlled by the signal level in that band. Although the results of using multichannel compression systems have been mixed, there seems to be general agreement that suitable compression schemes can provide some improvements in speech reception thresholds both in quiet and in noise [4, 5]. Although multichannel compression schemes inherently tend to "flatten" the spectrum of the output signal (relative to the input spectrum), recent research results [6] indicate that this may not have a significant effect on speech discrimination in quiet or in noise.

2. DESIGN GOALS

The primary goal of our multichannel compression scheme is to ensure that speech is always audible and comfortable, even in difficult listening environments. The design also seeks to

- process speech so that a minimum amount of distortion and other artifacts are introduced,
- not degrade (and perhaps even improve) the audibility of speech in noise, and
- provide a flexible architecture that can target the different needs of users with different hearing losses.

The rationales behind these goals can be grouped into concerns that address audibility, comfort and sound quality. Finally, we wish to minimize processing requirements, memory usage and group delay.

2.1. Audibility and Comfort

Ensuring that speech is audible is an important first step in improving the intelligibility of a speech signal. For speech in quiet, audibility can be improved by mapping the dynamic range of the input signal in the residual dynamic range of the hearing impaired listener. This implies that the multichannel compressor must have amplification that varies as a function of input level and compression ratios that are a function of frequency.

As well, it is important to ensure that

- low-level, high-frequency consonants are audible,
- speech following a loud transient is not made inaudible, and
- the levels of frequency bands that do not contribute to speech understanding are suppressed.

Comfort can be ensured by never allowing the output level of the hearing aid to exceed the wearer's UCL, even for a short period of time. The suppression of narrow-band noise signals that can cause listener fatigue is also an important factor in the comfort of a compression system.

2.2. Sound Quality

Compression systems purposely distort the input signal to reduce the dynamic range. The goal of a good compression system is to minimize the *perceived* distortion while still meeting the compression goals. Factors that influence the perceived quality of a compression system are

1. **holes**: "drop-outs" in the processed signal that occur as a result of short, intense sounds having too much influence on the gain control;
2. **modulation distortion**: distortion that is caused by the gain control changing too rapidly;
3. **pumping and breathing**: noticeable changes in background noise that occur when the gain changes too rapidly;
4. **spectral ducking**: broadband gain reductions in the processed signal that occur as a result of a narrow-band interfering signal;
5. **SNR reduction**: reduction in the actual or perceived signal-to-noise ratio caused by boosting low-level signals that do not contribute to speech understanding.

In our design, items 1–3 are addressed by using a fuzzy-logic detector and mixed overall/channel level gain control. Using a multiband compression system with channel level control addresses item 4. Item 5 is currently addressed by adjusting the compression threshold to lie above the noise floor. We are working to improve this by making an unbiased estimate of the input speech level (instead of assuming that the overall input level represents the speech signal level).

3. IMPLEMENTATION

Our design is based on an 16-channel oversampled DFT polyphase filterbank that operates at a sampling rate of 16 kHz. Level measurements are made at the input to the hearing aid (feedforward topology). The level used to control each channel gain is computed as a combination of a dual time-constant overall signal level and the short-term channel level. For each band, a single parameter is used to adjust how much control is given to the channel level or the overall level. A separate static characteristic (thresholds and compression ratios) can be specified for each band.

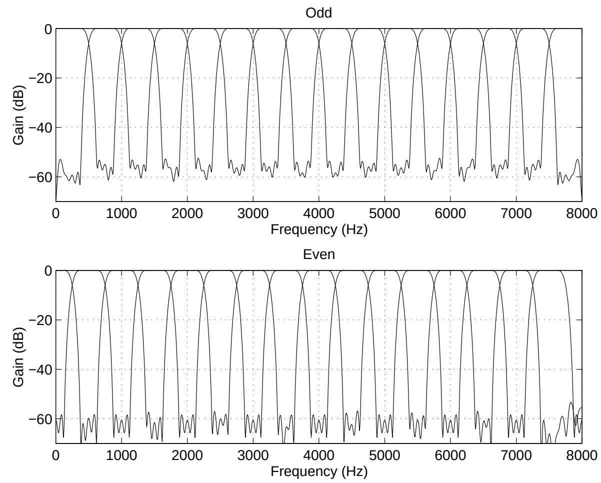

Figure 2. Frequency response of filter bank channels for odd and even channel stacking arrangements.

3.1. Filterbank

A filterbank provides a flexible means of processing signals for hearing aid applications because band gains can be adjusted independently or in combination as a function of both the overall or band levels to implement a particular processing scheme. Other digital hearing aid designs have used similar approaches [7, 8].

Initial experiments were conducted with critically sampled filterbanks. However, we discovered that to achieve high-quality reproduction, it was necessary to oversample by at least a factor of two to reduce the level of uncanceled aliasing that was generated when there were large gain differences between adjacent bands. Our design uses an oversampled, polyphase DFT filterbank [9] to split the input signal into 16 frequency bands. At a sampling rate of 16 kHz, the bands are 500 Hz wide. The bank can be evenly or oddly stacked (all channels shifted in unison by 250 Hz) to provide better approximations to the frequency responses required for precipitous or sloping low-frequency losses. Figure 2 shows the channel frequency responses for the even and odd stacking arrangements.

A combined analysis/synthesis filter is used to reduce the memory requirements of the filterbank. The synthesis filter is generated by decimating the analysis filter. The filter was designed using the equi-ripple, Nyquist eigenfilter method outlined in [10]. Combining the filters reduces the spurious-free dynamic range of the bank by approximately 20 dB; however, the memory requirements are also reduced by almost a factor of two. The filter bank requires approximately 46 multiply-accumulates per output point and has a group delay of 12.5 ms.

3.2. Static Characteristics

Static characteristics describe the system response for steady-state input signals. For hearing aid applications, audibility and comfort considerations usually dictate the static characteristics. In our design, the compression threshold is set to the greater of the normal hearing threshold or the estimated noise floor. This ensures that low-level noise is not overly amplified. Below the compression threshold, expansion is used to suppress low-level noise.

The limiting threshold maps normal UCL to the hearing aid user's UCL. Typical input/output and input/gain mappings are shown in Figure 3. Separate static characteristics consisting of

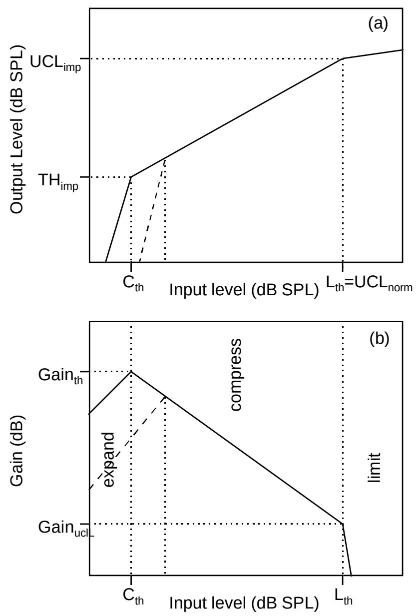

Figure 3. Static characteristics showing input versus output (a) and input versus gain (b).

thresholds and ratios for the compression and expansion regions are specified for each of the 16 bands.

It should be noted that a wide range of compression schemes can be implemented using this architecture. For example, the scheme employed by Moore [11] which uses a front-end compression-limiter to put the overall signal level at listener's most comfortable level (MCL) and syllabic compression in the high-frequency bands could be implemented within our framework.

3.3. Dynamic Characteristics

The level measurement scheme used by a compression system determines its dynamic characteristics. Usually, the dynamic characteristics are a compromise: we would like to have a level measurement that reflects the instantaneous level of the input signal. However, controlling the gain with a signal that changes rapidly can introduce modulation distortion and other artifacts like pumping and breathing. Using a long-term level measurement means that the output signal of the hearing aid may exceed the user's UCL and/or intense isolated transients many cause "holes" in the output signal. Our design seeks to

- get the overall level in the "right" place,
- not allow short-duration, intense sounds to dominate the gain control,
- adjust gain so as to avoid modulation distortion and other artifacts like pumping and breathing, and
- boost low-level, high-frequency components of speech that are preceded or followed by intense, low-frequency speech components.

The overall signal level is measured using an approach known as the *fuzzy detector* that was found to be successful in a high-quality audio application [12]. This method uses fuzzy control rules with trapezoidal membership functions to compute the overall level as a weighted combination of fast peak-level ($T_a = 4$ ms, $T_r = 150$ ms) and slow peak-level ($T_a = 325$ ms, $T_r = 4$ s) measurements. The weights applied to each level are adjusted as a

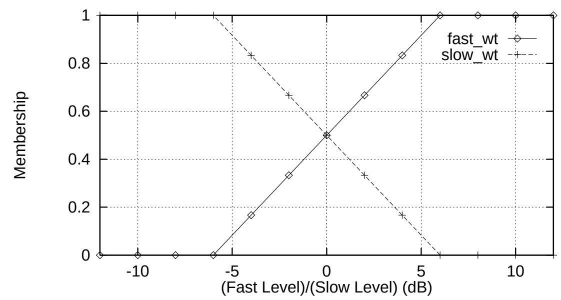

Figure 4. Fuzzy membership functions for transient (fast weight) and non-transient (slow weight) signals.

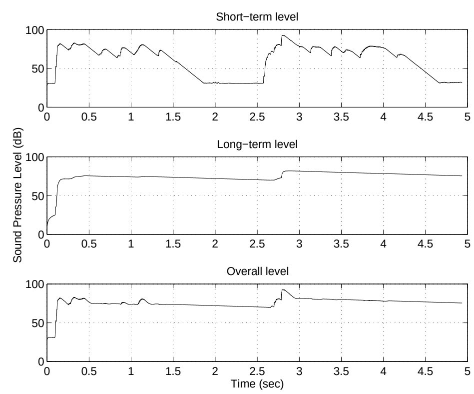

Figure 5. Level measurements versus time for a loud transient in speech showing short-term level, long-term level and the fuzzy detector "Overall" level.

function of the ratio between the fast and slow level measurements as shown in Figure 4.

Within each channel, a fast RMS level ($\tau = 50$ ms) is computed using a first-order recursive average. The gain applied to each channel is computed based on a level measurement that is a weighted combination of the overall level and the difference between the overall and channel levels. This causes the overall level to set the operating point and the band levels to cause deviations around this point. A similar approach was used by Schmidt in a high-quality, multichannel compressor [8].

In each compressor band, the weighting that is applied to the overall level and the difference between the overall and channel levels is adjustable so that at one extreme the channel gain is controlled by the overall level and at the other it is controlled by the channel level. Intermediate settings give gain that is controlled by combined overall and channel levels. This arrangement is very flexible; for example, it can be configured to provide overall level control in the low-frequency bands and syllabic compression in the high-frequency bands.

4. RESULTS

The measurement of overall level with a fuzzy detector provides the desired dual time-constant performance. Figure 5 shows a graph of the gain recovery for a loud transient (a door slam at 2.75 sec) during the second of two sentences (in quiet). The overall level (and hence the gain) tracks the long-term level of the speech signal, changing over to the short-term level only when it is necessary to suppress a transient (the door slam or the onset of a vowel). Informal listening tests revealed that there are no audible "holes" in the processed signal.

Figure 6(a) shows how the dynamic range of a processed speech signal (measured over two sentences using the 10% and 90% levels in 1/3rd-octave bands with a 46 ms window [13]) maps into the

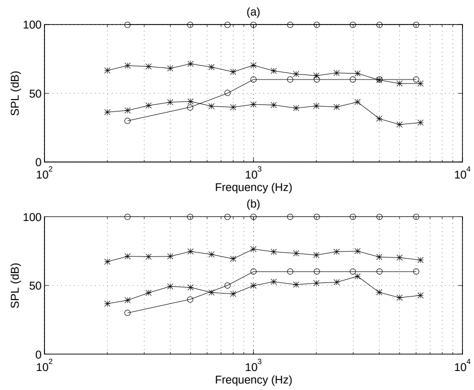

Figure 6. Mapping of 10% and 90% speech spectrum levels (*) into the residual dynamic range (O) for (a) overall level control and (b) 97 % channel level control.

residual dynamic range of a hypothetical hearing loss when the gain in each band is controlled by the the overall signal level. Figure 6(b) shows the same mapping when the gain is controlled by a level that is computed as $overall - 0.97(overall - channel)$ (i.e., 97% channel level control). Clearly, the use of combined channel/overall level control improves the mapping of the short-term dynamic range into the residual dynamic range, especially at high frequencies. For this parameter setting the processed signal has good audio quality and does not suffer from the modulation and pumping that occur when the gain in each channel is controlled by the channel level alone.

Because of the time constants used, overall level control results in a very small reduction in the short-term dynamic range of the processed signal: less than 5 dB at low frequencies and no change at high frequencies. The mixed overall/channel level control gives a short-term dynamic range reduction of 5 to 10 dB over most of the input frequency range.

It is difficult to objectively evaluate the audio quality of compression systems because their performance is very signal dependent. For steady-state signals, the gain is fixed and almost no artifacts are introduced. It is only when dynamic signals (like speech) are applied that the "real-world" performance of a compression system can be determined. We have conducted informal subjective listening tests with speech in quiet and found that the audio quality is very good. As expected, using channel level control alone (with $\tau = 50$ ms) results in some modulation distortion and breathing/pumping artifacts. Using combined control gives improved audio performance and compression that is very close to that of channel control alone (Figure 6(b)). For a flat hearing loss (60 dB SPL thresholds), speech processed with the mixed overall/channel level gain control has a noticeable high-frequency emphasis and higher overall level than signals processed with overall level gain control. This indicates a more effective mapping of the signal spectrum into the residual dynamic range.

Like most compression systems, the effectiveness of our compression system is reduced for speech plus noise input signals because the SNR of the processed signal is less than that of the input signal. Setting the compression threshold above the noise floor helps alleviate this problem, but it still persists in high background noise environments (e.g., a cafeteria or restaurant). We are working to improve this by making an unbiased estimate of the input speech level using a long-term level measurement in each channel and employing a simple method of speech/noise detection to determine the "speech" level and whether a particular channel level represents the speech or noise level.

5. CONCLUSIONS

Compression hearing aid designs have traditionally relied on single channel, feedback topologies. Digital technology allows for the implementation of much more complicated schemes, namely multichannel, feedforward designs. The compressor design presented here is an efficient implementation of a 16-channel feedforward compressor that utilizes a polyphase DFT filterbank with selectable even or odd stacking.

The static characteristics are selected to map the input signal dynamic range into the residual dynamic range of the hearing impaired listener. The gain in each channel is controlled by an adjustable combination of the overall signal level and the short-term channel level. The overall signal level is measured using a *fuzzy detector* that provides a dual time-constant characteristic. This combination ensures that the overall level is mapped correctly without loud, short-duration sounds causing "holes" in the processed signal. The use of mixed overall/channel level control allows for a syllabic compression in combination with overall level control.

We have simulated our design and shown that it is effective at mapping the short-term dynamic range into the residual dynamic range with good audio performance. Work on improving the performance in noise and on a real-time implementation is underway.

REFERENCES

[1] H. Dillon, "Tutorial: Compression? Yes, But for Low or High Frequencies, for Low or High Intensities and with What Response Times?," *Ear and Hearing*, pp. 287–307, Aug. 1996.

[2] A. J. Olivera, "A feedforward side-chain limiter compressor de-esser with improved flexibility," *J. Audio Engineering Society*, vol. 37, pp. 226–240, Apr. 1989.

[3] M. C. Killion, "The K-Amp Hearing Aid: An Attempt to Present High Fidelity for Persons with Impaired Hearing," *American Journal of Audiology*, pp. 52–74, July 1993.

[4] B. C. J. Moore, "Perceptial consequences of cochlear hearing loss and their implications for the design of hearing aids," *Ear and Hearing*, pp. 131–161, Apr. 1996.

[5] E. W. Yund and K. M. Buckles, "Multichannel compression hearing aids: Effect of number of channels on speech discrimination in noise," *J. Acoust. Soc. America*, vol. 97, pp. 1206–1223, Feb. 1995.

[6] E. W. Yund and K. M. Buckles, "Enhanced speech perception at low signal-to-noise ratios with multichannel compression hearing aids," *J. Acoust. Soc. America*, vol. 97, pp. 1224–1240, Feb. 1995.

[7] T. Lunner and J. Hellgren, "A Digital Filterbank Hearing Aid – Design, Implementation and Evaulation," in *Proc. ICASSP-91*, vol. 5, pp. 3661–3664, IEEE, May 1991. Toronto, ONT.

[8] J. C. Schmidt and J. C. Rutledge, "Multichannel dynamic range compression for music signals," in *Proc. ICASSP-96*, vol. 2, pp. 1013–1016, IEEE, May 1996. Atlanta, GA.

[9] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1983. Chapter 7.

[10] P. P. Vaidyanathan and T. Q. Nguyen, "Eigenfilters: A New Approach to Least-Squares FIR Filter Design and Applications Including Nyqiust Filters," *IEEE Trans. on Circuits and Systems*, vol. CAS-34, pp. 11–23, Jan. 1987.

[11] B. C. J. Moore, "Signal Processing to Compensate for Reduced Dynamic Range," in *Recent Developments in Hearing Instrument Technology* (J. Beilin and G. R. Jensen, eds.), 15^{th} Danavox Symposium, pp. 147–165, 1993.

[12] A. T. Schneider, "An Adaptive Digital Dynamic Range Controller," Master's thesis, University of Waterloo, Department of Electrical Engineering, Oct. 1991.

[13] R. M. Cox, J. S. Matesich, and J. N. Moore, "Distribution of short-term RMS levels in conversational speech," *J. Acoust. Soc. America*, no. 3, pp. 1100–1104, 1988.