A Programmable DSP Architecture for Hearing Aids

Ron Scicluna[1] and Leonard E. Cornelisse[2]
[1]ArgosyElectronics, Inc., 10300 West 70th Street, Eden Prairie, MN 55344, USA
[2]Dspfactory, Ltd., 611 Kumpf Drive, Unit 200, Waterloo, N2V 1K8, Canada

Abstract

The problem of providing beneficial hearing aid performance in multiple listening situations requires signal processing approaches which are impractical for analog technology. An X-band, Y-channel digital signal processing architecture lends itself to an acoustic transformation approach allowing the input dynamic range of real world sounds to be mapped into the available output dynamic range of the hearing impaired listener. A new programmable DSP-based system with a configurable FFT coprocessor is well suited to this application based on its size, power requirements, and raw processing speed. The software architecture is built around a system of Input/Gain (I/G) tables ensuring accurate matching to prescribed amplification targets. Gain application decisions are based upon intelligent assessment of the incoming signal using a multi-discriminant logic structure within pre-defined frequency regions or channels. Channelization addresses intelligibility issues while insuring that the demands of frequency response matching are independently met by the banding structure.

1. Introduction

The goal in fitting a hearing aid is an improvement in the overall quality of life. Unfortunately, industry statistics indicate that overall satisfaction with hearing instruments is low (53%) while nearly 12% of hearing instrument owners do not even use their instruments. In addition, roughly 80% of the eligible consumers - approximately 20 million people - do not own hearing instruments. The most often cited deficiency is a general lack of benefit across multiple listening environments. It follows then that the primary design goal is to provide obvious benefit in a variety of listening situations. Is there an opportunity here for DSP technology to be gainfully employed?

Technical improvements in amplification devices, over the last 20 years, have resulted in movement away from the very simple Class-A amplification to very complex multi-channel, programmable, multi-memory instruments (with input-adaptive frequency shaping and duration dependent AGC). These technologies have resulted in sound quality improvements, however, improvements in speech intelligibility in a variety of “real-world” environments generally have not followed.

Several signal processing algorithms, hosted on benchtop DSP’s, have shown promise, in clinical settings, with regard to achieving the stated design goal. To date, it has not been practical to implement these schemes using existing analog and even some digital technologies. The major product design hurdle is that the end user wants to hear better using an unobtrusive device! This requirement dictates the packaging limitations for the electronics and power supply. Affordability coupled with perceived value provide the other limitations.

Before we begin to evaluate the potential of DSP technology to address the performance issues described earlier, let’s consider the audiological issues, target prescription and instrument fitting methods, and high fidelity speech processing.

2. Loudness perception

The goal of amplification is to provide desired audibility for a range of normal input levels without producing perceptible distortion or exceeding uncomfortable listening levels. In general, hearing impairment is first and foremost a loss in loudness perception. This is characterized as an inability to respond to and employ the full range of sound level inputs. This loss in sensitivity to low level sounds is typically tied to damage in the outer hair cells (OHC) of the cochlea. Greater losses would tend to involve both the inner and outer hair cells.

Normal and impaired loudness perception is illustrated in Figure 1. The graph shows categorical ratings of loudness as a function of signal level for a normal hearing (red line) and a hearing-impaired (blue line) listener. For "Normal Loudness Growth" the full range of input levels maps to the various qualities of loudness - Too Soft to Too Loud.

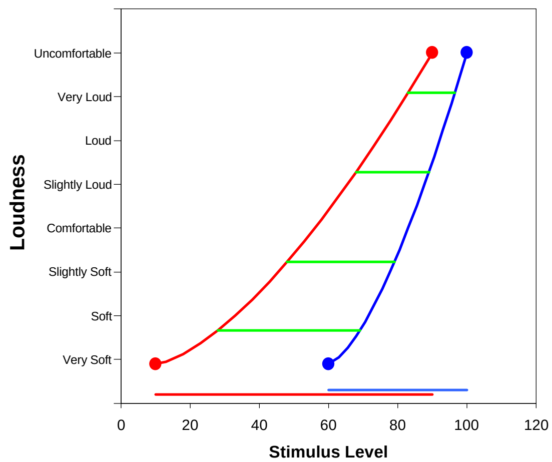

Figure 1. Loudness perception.

For the impaired auditory system, there is typically a loss in perception of the lower level sounds. For example, for the hearing-impaired listener, the signal level must be 50 dB greater than that for a normal-hearing listener to achieve loudness rating of "Very Soft". Signals that are just audible are considered to be at the listener's threshold for that particular frequency. Note that different frequencies can and typically do have different loudness growth curves.

For the abnormal loudness growth curve, the mapping of SPL values to loudness perception qualities is non-linear and considerably compressed. In Figure 1 the horizontal red line indicates the dynamic range of sound levels that are audible and comfortable for a normal hearing listener. The horizontal blue line indicates the auditory dynamic range of the hearing-impaired listener. In this example, the dynamic range is reduced by a factor of two.

2.1. Amplification Strategy for Loudness

The amplification goal of a "Loudness Growth" compensation strategy is to provide sufficient gain to make low-level sounds audible while progressively reducing the amplification so that high-level sounds receive little or no amplification (see green horizontal lines in **Figure 1**). The amplifier (or gain) curve of such a system is typically referred to as a compression curve. Proper specification of the compression parameters allows the large dynamic range of real-world sounds to be compressed into the residual dynamic range of the impaired listener.

The amplification prescription strategy, (i.e., fitting strategy), is based upon audiometric data (i.e., hearing thresholds and uncomfortable levels). These data are “fitted” to empirically derived amplifier target curves using either published equations or proprietary software. These targets are circuit-independent and relate to a high speech-to-noise environment for specific classes of loss. They define the frequency response which will perform best for that individual in that type of environment. There are some clinically-derived enhancements which are employed by some methods to improve the probability of achieving a “good fit” (a.k.a. amplification benefit). The role of the clinician here is to make the necessary adjustments to the hearing instrument allowing its response to match the target as closely as possible. This can be difficult to impossible with some types of instruments due to limitations in circuit performance and available adjustments. Note that the prescriptive method focuses only on predicting amplification targets specific to the hearing loss. This effectively solves the “static problem”. Problems associated with signal dynamics and overall processing fidelity are <u>not</u> addressed by the prescriptive method. Therefore, strict adherence to any specific prescriptive method does not by itself guarantee a desirable or beneficial outcome. The more advanced circuits available today can be effectively adjusted to restore normal loudness perception in a high signal-to-noise environment, however, effectiveness in a low signal-to-noise environments is limited by the prescribed compression ratio and preset AGC dynamics.

3. Speech Intelligibility

Compensation for abnormal loudness perception is essentially a prerequisite for achieving the real goal of hearing aid amplification - improved speech intelligibility (or recognition) in a variety of listening environments. In general, speech intelligibility relies on three factors - the overall sound quality, the audibility of key speech cues, and the speech-to-noise ratio. In order to achieve optimal speech intelligibility in a wide variety of environments, the hearing aid electronics are required to provide high-fidelity amplification within the constraints of dynamic range mapping across a wide variety of sounds. Therefore, signal processing mechanisms should not add perceptible distortion, over-compress the temporal envelope, or perceptibly modify the signal dynamics. This implies a high degree of adaptive and/or intelligent processing.

Note that "good fit" is achieved when a proper target is coupled with an appropriately capable circuit or algorithm. In other words, a complex, well-researched fitting strategy coupled with a simplistic amplifier structure (which at least adequately matches the prescribed target) may not sound as good as that same strategy coupled with an instrument comprising more capable amplifier elements. Obviously the marriage of the prescriptive method with the hearing aid electronics is critical to "fitting" success.

3.1 Sound quality

The are a few pseudo-universal qualities that people think of when describing sound quality. They include loudness, clarity, background noise, brightness, and ease-of-listening. A sound quality "score" or measure has both an objective and a subjective side to it - and the two types of measures don't always agree. Unfortunately, instrument designs must address these "difficult to quantify" subjective requirements. For the hearing impaired listener, ease-of-listening is the characteristic most often mentioned when describing the quality of a hearing instrument. These are expressed in terms of the degree of strain, concentration, and overall mental effort they must go through to both hear and understand speech.

3.2 Speech audibility

The brain's recognition of sounds is linked to how well the pattern of that sound matches the pattern stored in memory. Sometimes a subtle change in the pattern of a sound can change it into a slightly different, incorrectly interpreted sound. Such subtle changes can occur due to imprecision on the part of the talker, distortion affects of the media, background noise, and/or frequency or temporal impairment in the listener's auditory system. We rely on the audibility of speech cues for our understanding of conversations we're involved in. Speech understanding can be improved using an amplifier structure designed to accommodate achievement of required and/or desired audibility in selected narrow bands of frequency.

3.3 Competing sounds

Speech intelligibility is degraded significantly, even for normal hearing listeners, when the background sounds are loud enough to compete with the primary talker. Typical difficult listening environments are restaurants, bars, sporting events, and automobiles. The various types of competing sounds - several simultaneous conversations, road noise, overhead fans, machinery, etc have different temporal and spectral characteristics thereby offering different challenges to speech recognition.

These situations are tricky but tolerable for those of us with normal loudness perception because of our ability to process the entire dynamic range of the input signal. Individuals with a hearing impairment on the other hand typically wear amplifiers, which use non-linear amplification to compress the existing real-world dynamic range into their smaller residual dynamic range. The prescribed compression ratio may be inappropriate for certain of these situations serving to obliterate information in the lower speech-to-noise signal. Intelligent selection of the compression ratio and the attack and release characteristics of the AGC functions plays a big role in preserving speech information in the presence of noise.

4. The DSP Solution

Given the above requirements, what algorithmic approach should we pursue? Three things appear to be true: (1) any replication of existing, traditional hearing instrument processing logic is going to provide limited success, (2) DSP technology generally lends itself quite well to processing decisions involving multiple discriminants, and (3) there aren't enough MIPs available in the currently applicable technology to fully simulate the cochlea. The solution then is a compromise which leverages industry experience regarding loudness restoration, a practical understanding of high-fidelity sound processing, and an overall appreciation for psycho-acoustics as they relate to sound quality and speech intelligibility.

In addition to the foregoing, the processing architecture we have chosen fits well within the power requirements and processing latency limitations specific to hearing instruments. Our X-band, Y-channel structure is unique in that it can fully and accurately accommodate the latest in prescriptive fitting methods without compromising its ability to properly recognize and dynamically process a wide range of psychoacoustic information in a wide variety of noise backgrounds. Our algorithm uses multi-discriminant logic (e.g. acoustic cues) to not only select but to continually and transparently fine tune the processing performed within the instrument to fit the dynamics of the listening situation. The algorithm includes a multilevel detector structure designed to individually recognize frequency, amplitude, and duration cues and process them in accordance with rule-based logic. This processing architecture effectively minimizes the compromises to sound quality which typically result when the list of discriminants and/or processing remedies is too short for the prevailing environmental conditions. This level of intelligent processing is well beyond the scope of analog and simple digital hearing instruments.

4.1 Hardware

The processing solutions discussed in this paper are implemented on the Toccata Application Specific Signal Processor (ASSP) system, a. new 3-chip hybrid (1). The heart of this system is a 16-bit DSP core fabricated on 0.18 micron semiconductor technology. The key element of the DSP architecture is an oversampled, weighted overlap-add (WOLA) DFT filterbank operating as a co-processor. The individual chip architectures and the DSP instruction set have been optimized for use in hearing instruments where size, speed, and power consumption are critical. The DFT coprocessor along with instruction set and related architectural efficiencies allow this chipset to host complex and effective hearing aid algorithms well within tolerable audio delays.

4.2 Block diagram

The simplified system block diagram (Figure 2) shows the basic elements of this DSP-based hearing instrument. Conditioned input signals – single microphone, dual microphone, and/or telecoil data are fed to an n-band FFT processor. Those bands are then combined into channels. The Discriminant Detectors analyze several statistics within this channelized data and estimate the input level. The input levels are, in turn, used by the I/G Tables to select the gain, which is applied prior to the signal synthesis process. Information about the spectral shape of the amplification target is stored in the Freq Shape Table. Gain adjustments are facilitated via an external Volume Control. In addition, the user can choose to exercise certain operational

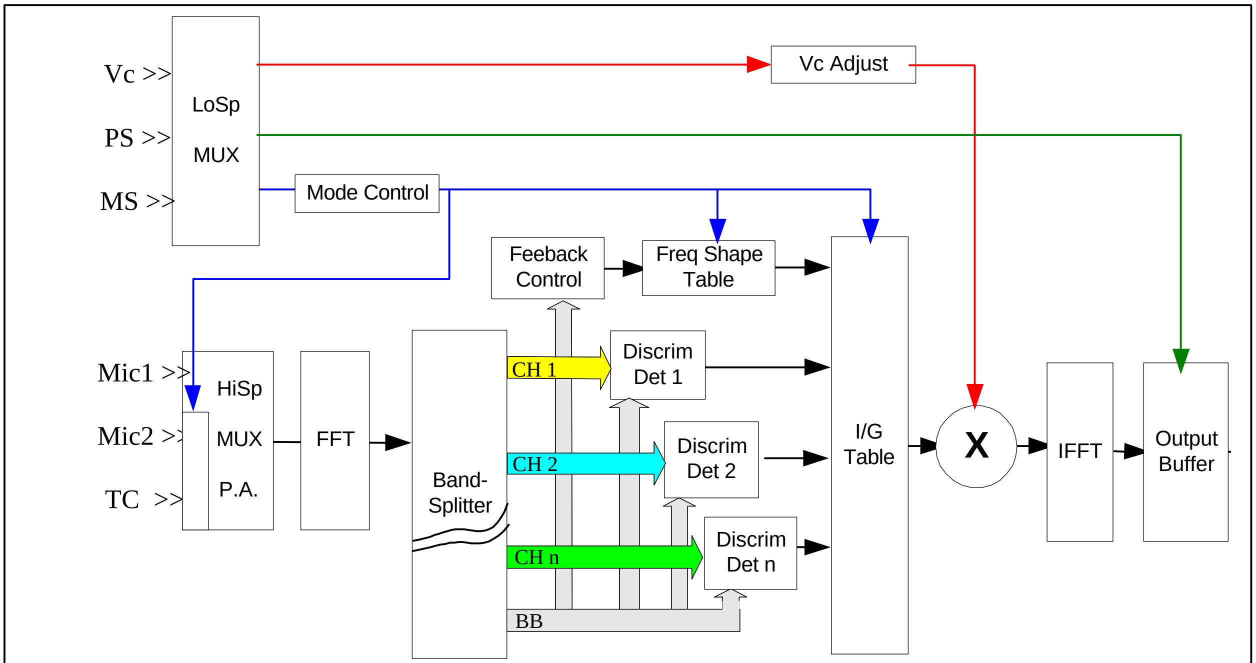

Figure 2. System block diagram.

characteristics via an external Mode Selection switch.

The contents of both the I/G Tables and the Freq Shape Table are defined by a separate Prescriptive Method algorithm and downloaded at the time of instrument fitting. This allows the gain in the individual bands to be defined and controlled separately. This table structure achieves loudness restoration by facilitating accurate mapping of the wide dynamic range for normal speech and noise information into the narrower residual dynamic range of the hearing impaired individual. This structure effectively addresses the static problem described earlier.

4.3 Dynamic processing

The challenge for the hearing instrument amplifier is to transparently accommodate the fluctuations in the input level for various environments. The key in high-fidelity sound processing is an algorithm which intelligently and accurately determines the input level and the situation and selects the gain accordingly. This approach distinguishes between transient characteristics of various speech cues and slow changes in the background statistics. It is able to adaptively "attack" and subsequently "release" from a complex sequence of variable type and duration acoustic events. Gain application changes are transparent to the instrument wearer. There are no obvious side effects of pumping, roughness, dropouts, or distortion.

In general, automatic gain control systems use continuously updated averages of the input signal to control the applied gain. If the signal processing structure uses a single broadband channel it is easy to envision circumstances when the averaged input will be dominated by predominantly low frequency or high frequency energy. This can seriously degrade audibility and intelligibility in the "out-of-band region". The remedy, included in the block diagram, is to subdivide the processing structure into a collection of amplifier chains (e.g. multiple channels) each having independent control over a specific and significant region of the audio frequency band. Each chain has its own system of multi-discriminant logic optimized for the acoustic events anticipated therein.

Our X-band, Y-channel structure allows us to place the channel boundaries in accordance with psycho-acoustic logic (rather than in accordance with target matching criteria) since the channelization governs the gain update rate and not the fitting characteristics. Information regarding target gain as a function of input level remains within each individual band. This approach allows us to achieve desired Input/Output characteristics by frequency and, at the same time, identify and protect key spectral regions from out-of-band corruption which would only serve to compromise overall speech intelligibility.

The difficulty hearing instruments have in dealing with multiple environments is reduced intelligibility when the background noise increases. One reason for this in hearing instruments today is over compression. This is a result of applying the prescribed high SNR compression ratio value to noisy, low SNR data. Our algorithm is designed to evaluate the environmental statistics and to select the most suitable I/G table. This effectively provides for adaptive compression ratio strategies. In fact, intelligibility studies have shown that although the prescribed level of compression is acceptable for the rather high signal-to-noise situations, hearing impaired listeners prefer a reduced compression ratio for the high ambient, low signal-to-noise environments.

5. Conclusion

The promise of multi-environment, high-fidelity audio processing for hearing aid wearers comes as a result of a new programmable DSP-based system. The processing hardware along with our X-band, Y-channel algorithm lend themselves well to the heretofore conflicting requirements of loudness restoration and speech intelligibility. The electronics package fits well within the size and power limitations imposed on hearing aid components and easily hosts sophisticated signal processing algorithms.

References

(1) Schneider T, Brennan R, Balsiger P, Heubi A, 1999, An ultra low-power programmable DSP system for hearing aids and other audio applications, Proc. ICSPAT, this issue.