LINEAR ACOUSTIC
Audio Under Control.

Effective Wideband Processing: Magic or Myth?

Why Simple is Not Always Better

Tim Carroll
Linear Acoustic Inc.
April 2009

Recently there has been a resurgence in the marketing of audio processing designed around the principle of a single gain element controlling all frequencies of audio signals simultaneously. Known as wideband processing, these types of systems have been in use since the earliest days of telephony and broadcasting when frequency response and level variations were both restricted.

Times have certainly changed. Is wideband processing effective at controlling the loudness of modern, full bandwidth audio channels with the sometimes significant variations in level and quality?

Wideband Processing Structures

As shown in Figure 1, wideband processing in its simplest form is a detector controlling an attenuator.

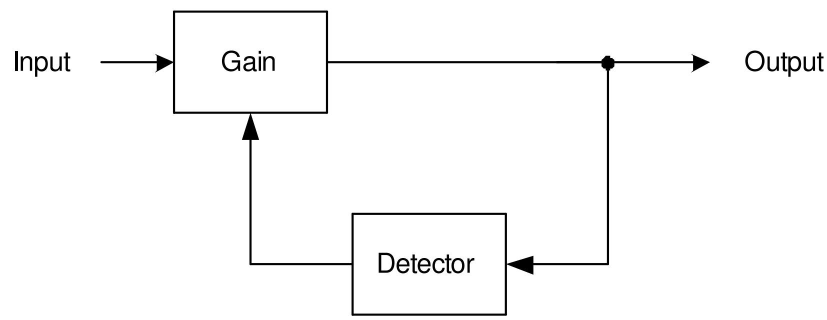

Figure 1 – Wideband AGC.

Tasked with providing protection or additional smoothing for relatively narrow bandwidth audio controlled by trained operators or with controlling the rather narrow dynamic range of spoken voice via telephone systems, wideband signal processing can produce acceptable results.

Modern versions of wideband processors have added more complex control loops that can adjust the attenuation at differing speeds depending on the input audio. The control signals can also be weighted to be less affected by very low frequencies and thus minimize the modulation of the entire spectrum by strong signals in the lowest part of the spectrum. This is known as intermodulation distortion. Weighting and control loop speed switching improve performance, but invariably there will be overshoots so wideband processing is often followed by wideband peak limiting. A more modern wideband processor with peak limiting is shown below in Figure 2.

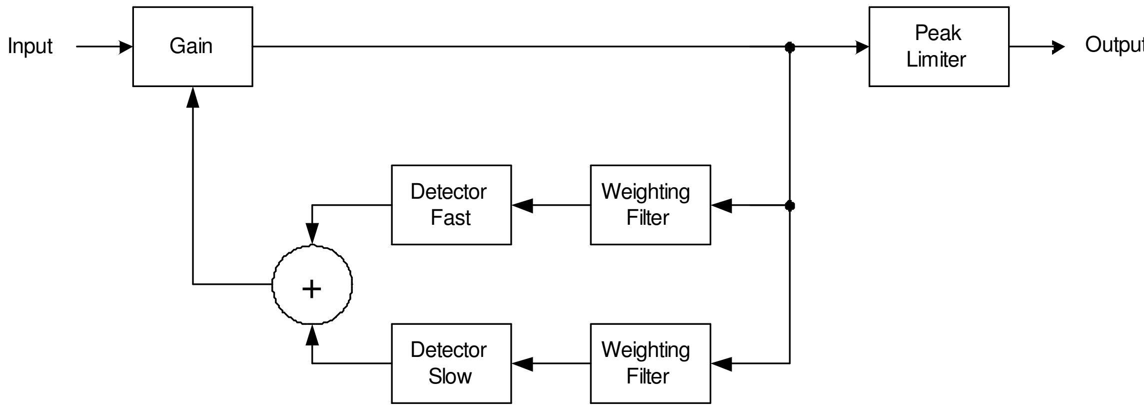

Figure 2 – Multi-loop Wideband AGC with Weighting and Peak Limiter

Modern Versions

The Dolby Digital (AC-3) system employs wideband gain control through its metadata DRC system. Control words are generated in the encoder and applied to the gain element in the decoder, with the result being a wideband processor. This system uses detection and gain control across all channels[1] and has the advantage of many milliseconds of look-ahead thanks to the inescapable delay of the audio coding itself.

The DRC system inside of Dolby Digital is a very complex implementation of multi-speed control signals that are carefully weighted to minimize the effects of intermodulation distortion.

Wideband and Spectral Balance

Claims of maintaining the original spectral balance must be carefully considered. If the original spectral balance matched the standards of the broadcaster, all is fine. If however the original spectral balance was inferior due to rushed, unskilled, or otherwise compromised production, a wideband processor cannot help and can actually make the situation worse.

What is frequently overlooked is the time-proven psychoacoustic research of Fletcher, Munson, and others from which the equal loudness curves shown in Figure 3 were developed. It should be immediately apparent that the ear is not linear and does not react equally in frequency response to signals of differing loudness.

[1] Only the main channels are used for detection, the LFE is ignored, but gain control is applied to all channels including the LFE.

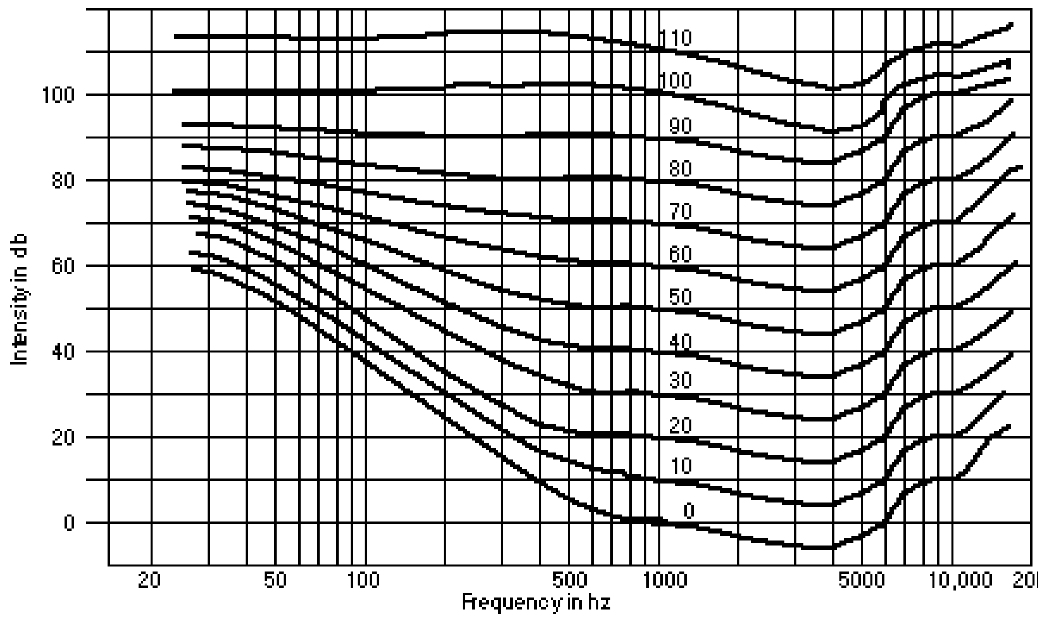

Figure 3 – Equal Loudness Contours (from Fletcher-Munson)

It can be seen that a signal reproduced at lower levels will be spectrally shaped by the sensitivity of the human auditory system. The quietest curve (bottom line) shows that the ear is 20dB less sensitive at about 200Hz and 15kHz than it is at 3.5kHz. Most people are familiar with the common loudness control found on receivers for decades. These controls were an effort to minimize the perceived change in frequency response caused by listening at lower levels.

Wideband processing will have an identical impact on the audio it is controlling. When gain is reduced by many dB, the audio will tend to sound dull and thin compared to the original. There is no escaping this. Sometimes large reductions of level are required and humans will always hear this the same way: dull.

Adding ITU to the Mix

With the development of ITU BS.1770, a simple and effective method now exists for measuring loudness in an objective and repeatable manner. A seemingly logical next step might be to use this new method to replace or augment the detection methods used by wideband processors. Essentially weighted and integrated RMS in nature, it is well-suited for measurement over time.

The key to accurate measurement is enough integration time. Research has shown that three seconds is the practical minimum for a valid result, with longer times producing better results.[2] Obviously, this is far too slow to be used to control the loudness of a program stream with content such as commercial advertisements that may be instantly much louder than the surrounding programs. In a structure where the ITU-controlled AGC was followed by a peak limiter, this would mean that content would be heavily peak limited until the integrated measurement was able to lower the gain appropriately.

[2] [13] Soulodre, G.A. & Norcross, S.G. (2003) "Objective Measures of Loudness", AES 115th Convention

The audible results of this type of processing is content that varies substantially in density as the peak limiter strains to manage the content that is too loud but decreased in gain only as fast as the integrator responds. The solution is diametrically opposed to the problem: the integrator needs time to be accurate while content that is instantly too loud needs to be adjusted quickly to avoid excessive limiting.

Multiband Processing

Primarily driven by processing developments in the 1970s for radio broadcasting, splitting the audio into multiple bands minimized intermodulation distortion and allowed the constraints imposed by required pre-emphasis to be better managed. For the first time, loudness could be better controlled with fewer audible artifacts, and as a helpful side benefit, spectral balance could be managed as well. The only unfortunate part is that broadcast began to push for higher and higher density in the battle for ultimate loudness, and the multiband techniques employed as weapons in the battle gained an unsavory reputation. This heavy-handed processing also found its way into television thanks to the efforts of manufacturers involved in radio processing pushing the same or similar products into TV. Like any processing technique that is overused, the myriad benefits of multiband structures were overshadowed by the overly dense results from pushing it beyond what was necessary. Like many things, there is no need to throw out the baby with the bathwater.

Apart from the very wide dynamic range and frequency response capabilities offered by modern digital broadcast systems, the removal of the thorn of pre-emphasis is arguably the best audible improvement. It immediately allows for a dramatic reduction in processing ratios and allows many of the benefits of multiband processing to once again shine through.

Figure 4 below shows a typical modern multiband processing structure. Note the arrangement of processing from left to right is also from slowest to fastest with speed being directly determined by the degree to which a given section can make an adjustment without producing audible side effects.

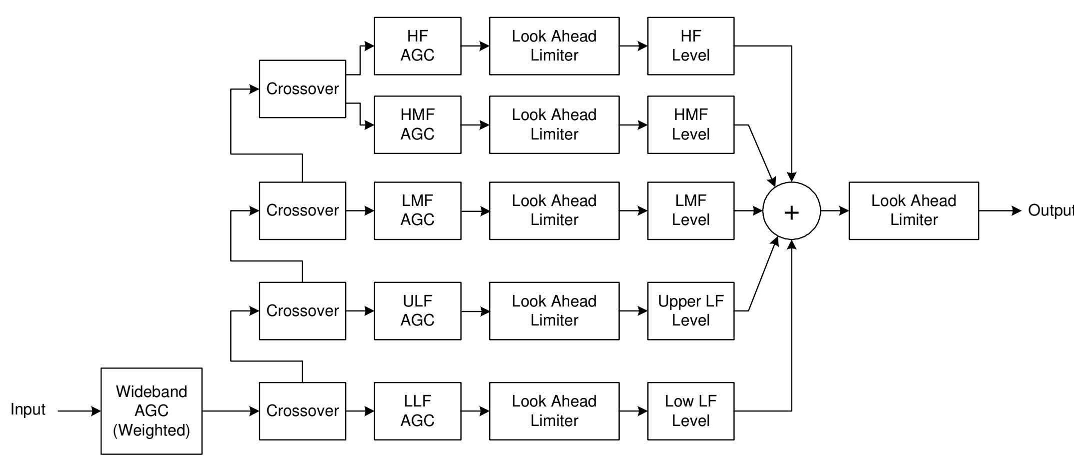

Figure 4 – Multiband Processing Structure

The process starts with an ITU-like wideband weighted AGC which acts like a smooth hand on the incoming audio levels, adjusting them to the so-called sweet spot of the following multiband section. This helps minimize over-processing, and is also the section where gating can be applied to prevent bringing up background noise.

A crossover with carefully chosen points feeds individual AGCs which are tuned to respond best to a select range of frequencies. These AGCs act faster than the wideband AGC and again position the audio in each band appropriately for the following multiband look-ahead limiters.

Look-ahead limiters are able to control peaks very accurately because they measure the audio a few samples ahead of when they apply gain control. Look-ahead correctly implies that they act faster than real time and so are extremely effective at making changes inaudibly, and this is helped by the fact that it is done on a band-by-band basis.

The entire process is followed by a final peak limiter to catch any overshoots caused by preceding stages. This final section does very little in most cases but will prevent overload if programs with dramatic variations are applied to the processor.

The somewhat complex signal path is made more complex by the fact that the processes applied in each frequency area vary in time. Some processors use adaptive timing so that response to incoming audio is constantly optimized. Unlike a simple wideband processor though, each stage does only what it is good at doing. Some stages are fast but operate over a restricted gain control area, while others can operate over a much wider range but do so only at slow speed.

Summary

It should be fairly obvious by this point that it is not possible for wideband processing to handle complex audio that varies greatly in loudness without causing audible side effects. Claims that wideband is the only way to preserve original content are simply not founded in the reality of the science of the human auditory system, which itself is a multiband system and responds in a non-linear manner to sound.

A simple test will quickly reveal the truth. Wideband processing, no matter how sophisticated, will either sound good and not control abrupt loudness problems or it will control them and cause audible side effects.

While it would be far less complex to design wideband processing, experience has shown that it is simply not effective enough to solve problems and still sound good. It is very important to remember that a sophisticated multiband processor can easily be adjusted to perform as if it were wideband, but the reverse is never true.