Digital Dynamic Range Compressor Design—A Tutorial and Analysis

DIMITRIOS GIANNOULIS, MICHAEL MASSBERG, AND JOSHUA D. REISS, *AES Member*
(Dimitrios.Giannoulis@eecs.qmul.ac.uk) (michael@massberg.org) (josh.reiss@eecs.qmul.ac.uk)

Queen Mary University of London, London, UK

Dynamic range compression, despite being one of the most widely used audio effects, is still poorly understood, and there is little formal knowledge and analysis of compressor design techniques. In this tutorial we describe several different approaches to digital dynamic range compressor design. Digital implementations of several classic analog approaches are given, as well as designs from recent literature, and new approaches that address possible issues. Several design techniques are analyzed and compared, including RMS and peak–based approaches, feedforward and feedback designs, and linear and log domain level detection. We explain what makes the designs sound different and provide metrics to analyze their quality. Finally, we provide recommendations for high performance compressor design.

0 INTRODUCTION[1]

Dynamic Range Compression (DRC) is the process of mapping the dynamic range of an audio signal to a smaller range [1-2], i.e., reducing the signal level of the higher peaks while leaving the quieter parts untreated. DRC is used extensively in audio recording, production work, noise reduction, broadcasting, and live performance applications.

Considering the classic audio effects (equalization, delay, panning, etc.), the dynamic range compressor is perhaps the most complex one. Design choices involve the compressor topology, the static compression characteristic, placement, and type of smoothing filters, sidechain filtering, etc. This explains why every compressor in common usage behaves and sounds slightly different and why certain compressor models have become audio engineers' favorites for certain types of signal. The analysis of compressor designs is difficult because they represent nonlinear time-dependent systems with memory. The gain reduction is applied smoothly and not instantaneously as would be the case with a simple static nonlinearity. Furthermore the large number of design choices makes it nearly impossible to draw a generic compressor block diagram that would be valid for the majority of real world compressors. "No two compressors sound alike . . . each one is inaccurate in its own unique way." [3] Some differ in topology, others introduce additional stages, and some simply differ from the precise digital design since these deviations add character to the compressor. However we can describe the main parameters of a compressor unit and specify a set of standard stages and building blocks that are present in almost any compressor design.

In general, compressors are quite an unexplored field in academia. The authors know of no prior publication that collectively derives and discusses the most common compressor design options in depth. A better understanding of the different design choices would make compressor design considerably easier and less of a trial-and-error process. It would also shed light on how the different design choices affect the perceived sonic characteristics of the compressor. The goal of this paper is to provide such analysis, considering both technical and perceptual aspects.

The tutorial is structured as follows. Section 1 provides an overview of the most common controls used to operate a dynamic range compressor. In Section 2, the principles of its operation are described. Particular attention is paid to the level detection and use of attack and release times, especially their digital implementation based on classic analog designs. Section 3 discusses the different methods and implementation for a complete design. Section 4 compares the behavior of several different designs using both example compression curves and formal objective measures. Conclusions and recommendations for designs are provided in Section 5.

1 COMPRESSOR CONTROLS

A compressor has a set of controls directly linked to compressor parameters through which one can set up the effect. The most commonly used compressor parameters may be defined as follows.

[1] *Accompanying audio samples and source code (in Matlab and as VST plug-ins) for the compressor designs and analysis are available from www.elec.qmul.ac.uk/digitalmusic/audioengineering/*

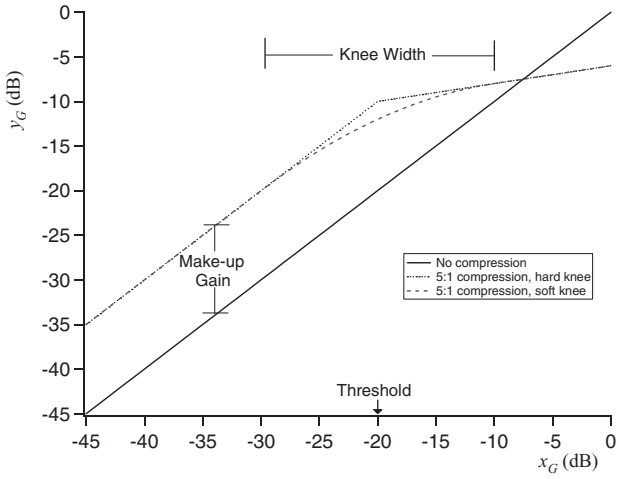

Fig. 1. Static compression characteristic with make-up gain and hard or soft knee.

Threshold defines the level above which compression starts. Any signal overshooting the threshold will be reduced in level.

Ratio controls the input/output ratio for signals overshooting the threshold level. It determines the amount of compression applied.

Attack and release times provide a degree of control over how quickly a compressor acts. They are also known as time constants, although the latter is usually referenced to dB, denoting the gain decrease in dB that the compressor will apply for the given attack time and the opposite for the release time. Instantaneous compressor response, as described in [4], is not sought because it introduces distortion on the signal.

The attack time defines the time it takes the compressor to decrease the gain to the level determined by the ratio once the signal overshoots the threshold. The release time defines the time it takes to bring the gain back up to the normal level once the signal has fallen below the threshold.

A *Make-Up Gain* control is usually provided at the compressor output. The compressor reduces the level (gain) of the signal; therefore, applying a make-up gain to the signal allows for matching the input and output loudness level.

The *Knee Width* option controls whether the bend in the compression characteristic (see Fig. 1) has a sharp angle or has a rounded edge. The Knee is the threshold-determined point where the input-output ratio changes from unity to a set ratio. A sharp transition is called a Hard Knee and provides a more noticeable compression. A softer transition where the ratio gradually grows from 1:1 to a set value in a transition region on both sides of the threshold is called a Soft Knee. It makes the compression effect less perceptible. Depending on the signal one can use hard or soft knee, with the latter being preferred when we want less obvious (transparent) compression.

A compressor has a set of additional controls most of which are found in most modern compressor designs. These include a Hold parameter, Side-Chain filtering, Look-Ahead, and many more.

2 PRINCIPLE OF OPERATION

The signal entering the compressor is split in two copies. One is sent to a variable-gain amplifier (the gain stage) and the other to the side-chain where the gain computer, a circuit controlled by the level of the input signal, applies the required gain reduction to the gain stage. The copy of the signal entering the sidechain has its bipolar amplitude converted into a unipolar representation of level. When the level of the signal is determined by its absolute value (instantaneous signal level), this is known as peak-sensing. An alternative is to use RMS-sensing and determine the signal level by its root-mean-square (RMS) value.

2.1 The Gain Stage

The gain stage is responsible for attenuating the input signal by a varied amount of decibels (dB) over time. The heart of every compressor is the element that applies this gain reduction: a voltage controlled amplifier (VCA) that attenuates the input signal according to an external control voltage (hereafter denoted c) coming from the side chain. Building a VCA with analog components is non-trivial and thus many approaches have been applied. Optical compressors like the Teletronix LA-2A use a light-dependent resistor (LDR) as the bottom leg of a voltage divider located within the signal path.[2] The control voltage is used to drive a light source which will then—with increasing brightness—lower the resistance of the LDR and therefore, apply the necessary gain reduction. A similar approach is taken by the Universal Audio 1176 compressor, but instead of using an LDR, they used the drain-to-source resistance of a field effect transistor (FET compressor).[3] The control voltage could then be applied to the gate terminal in order to lower the FET's resistance [5]. An even earlier approach used in so-called variable-mu compressors like the Fairchild 670[4] exploited the fact that altering the grid-to-cathode voltage would change the gain of a tube amplifier. More modern compressor designs make use of specialized integrated VCA circuits. These are much more predictable than the earlier approaches and offer improved specifications (such as less harmonic distortion and a higher usable dynamic range).

In a solely digital design, one can model an ideal VCA as multiplying the input signal by a control voltage coming from the sidechain. If $x[n]$ denotes the input signal, $y[n]$ the output signal and $c[n]$ the control voltage then $y[n] = c[n] \cdot x[n]$. In addition, make-up gain is often used to add a constant gain back to the signal in order to match output and input levels. In a digital compressor we can easily implement a make-up gain by multiplying the compressor's output by a constant factor corresponding to the desired make-up gain value. So, representing the signals in decibels, where M is the make-up gain,

$$y_{dB}[n] = x_{dB}[n] + c_{dB}[n] + M. \quad (1)$$

[2] www.recproaudio.com/diy_pro_audio/diy_files/la2a/la2a_1968.jpg

[3] www.gyraf.dk/gy_pd/1176/1176sch.gif

[4] www.audio.kubarth.com/fairchild/670_schematics_redrawn.png

2.2 The Gain Computer

The gain computer is the compressor stage that generates the control voltage. The control voltage determines the gain reduction to be applied to the signal. This stage involves the Threshold T, Ratio R, and Knee Width W parameters. These define the static input-to-output characteristic of compression. Once the signal level exceeds the threshold value, it is attenuated according to the ratio.

The compression ratio is defined as the reciprocal of the slope of the line segment above the threshold, that is:

$$R = \frac{x_G - T}{y_G - T} \quad for \quad x_G > T \tag{2}$$

the static compression characteristic is described by the following relationship:

$$y_G = \begin{cases} x_G & x_G \le T \\ T + (x_G - T)/R & x_G > T \end{cases} \tag{3}$$

In order to smooth the transition between compression and no compression at the threshold point, we can soften the compressor's knee. The width W of the knee (in decibels) is equally distributed on both sides of the threshold. Fig. 1 presents a compression gain curve with a soft knee.

To implement this, we replace the hard knee characteristic used in Eq. (3) with a soft knee characteristic, giving the following piecewise, continuous function,

$$y_G = \begin{cases} x_G & 2(x_G - T) < -W \\ x_G + (1/R - 1)(x_G - T + W/2)^2/(2W) & 2|(x_G - T)| \le W \\ T + (x_G - T)/R & 2(x_G - T) > W \end{cases} \tag{4}$$

When the knee width is set to zero, the smooth knee is identical to the hard knee.

2.3 Level Detection

The level detection stage is used to provide a smooth representation of the signal's level and may be applied at various places in the side-chain. The gradual change of gain is due to the attack and release times. The process of properly setting up these times is crucial to the performance of the compressor since unpleasant artifacts are often associated with the choice of these compressor parameters.

The attack and the release times are usually introduced through a smoothing detector filter. We can simulate the time domain behavior of the filter in the digital domain with a digital one-pole filter,

$$s[n] = \alpha s[n-1] + (1 - \alpha) r[n], \tag{5}$$

where α is the filter coefficient, $r[n]$ the input, and $s[n]$ the output. The step response of this filter is:

$$s[n] = 1 - \alpha^n \quad \text{for} \quad x[n] = 1, n \ge 1. \tag{6}$$

The time constant τ is defined as the time it takes for this system to reach 1-1/e of its final value, i.e., $s[\tau f_s] = 1 - 1/e$. Thus from (6), we have

$$\alpha = e^{-1/(\tau f_s)}. \tag{7}$$

Alternate definitions for the time constant are often used [11]. For instance, if one considers rise time for the step response to go from 10% to 90% of the final value

$$0.1 = 1 - \alpha^{\tau_1 f_s}, 0.9 = 1 - \alpha^{\tau_2 f_s} \rightarrow \tau_2 - \tau_1 = \tau \ln 9 \tag{8}$$

In analog detectors the filter is typically implemented as a simple series resistor capacitor circuit, with $\tau = RC$. Eqs. (5) and (7) may then be found by digital simulation using a step invariant transform. The advantage of this approach over other digital simulation methods like the bilinear transform is that we preserve the analog filter topology with the capacitor's voltage as the state variable. Therefore, we will not experience any clicks and pops once we start varying the filter coefficients over time.

2.3.1 RMS Detector

Level detection may be based on a measurement of the Root Mean Squared (RMS) value of the input signal [6–7], which is defined as

$$y_L^2[n] = \frac{1}{M} \sum_{m=-M/2}^{M/2-1} x_L^2[n-m] \tag{9}$$

The RMS detector is useful when we are interested in a smoothed average of a signal. Another motivation for using RMS detection is that it is more closely related to perceived signal loudness. However, this is generally unsuitable for real-time implementations since it enforces a latency of M/2 samples. In the implementation of real-world effects this measurement is often approximated by filtering the squared input signal with a first-order low-pass IIR filter and taking the square root of the output [8]. This is also commonly found in analog RMS-based compressors [9–10]. The difference equation of the RMS detector with smoothing coefficient becomes:

$$y_L^2[n] = \alpha y_L^2[n-1] + (1 - \alpha) x_L^2[n] \tag{10}$$

The RMS detector is closely related to the volume unit (VU) meter. However the VU meter does not measure RMS values directly. Instead it multiplies the average rectified value of the input signal by the sine wave's form factor of $\pi/\sqrt{2}$ [11]. Therefore the VU meter's output is only equal to the true RMS value for sine waves. Furthermore the averaging filter of the VU meter is slightly under-damped, resulting in an over-swing of the needle after transients.

We can find RMS detectors in some compressors at the beginning of the sidechain and before the decibel conversion. In [12], it was shown that Eq. (10) produces behavior generally equivalent to Eq. (5), save for a scaling of the time constant. Therefore, we will focus on the various options for the peak detector.

2.3.2 Peak Detector

The analog peak detector circuit, as commonly found in analog dynamic range controls [13], is given in Fig. 2. If we are not required to simulate a particular type of diode, we can idealize it by assuming that it can supply infinite current once the voltage across the diode becomes

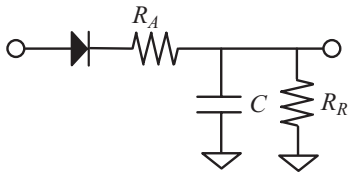

Fig. 2. Peak detector circuit

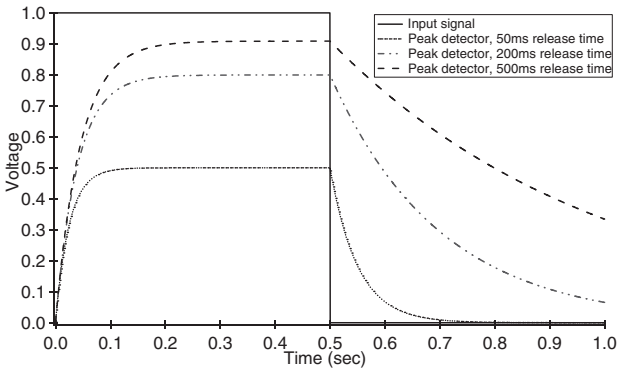

Fig. 3. Output of the peak detector circuit for different release time constants. Attack Time = 50 ms.

positive and completely blocks when reverse biased. This significantly simplifies the calculation [12].

$$\frac{dV_C}{dt} = \frac{\max(V_{in} - V_C, 0)}{R_A C} - \frac{V_C}{R_R C} \tag{11}$$

The capacitor is charged through resistor R_A according to a positive voltage across the diode but continually discharged though R_R. Taking α_A as the attack coefficient and α_R as the release coefficient, calculated according to Eq. (7) from the attack and release times $\tau_A = R_A C$ and $\tau_R = R_R C$), we can simulate the ideal analog peak detector with:

$$y_L[n] = \alpha_R y_L[n-1] + (1-\alpha_A)\max(x_L[n] - y_L[n-1],\ 0) \tag{12}$$

Although used in many analog compressors, and some digital designs [14], this circuit has a few problems. When $x_L[n] \geq y_L\ [n-1]$, the step response is

$$y[n] = (1-\alpha_A)\sum_{m=0}^{n-1}(\alpha_R + \alpha_A - 1)^m \rightarrow \frac{1-\alpha_A}{2-\alpha_R-\alpha_A} \approx \frac{\tau_R}{\tau_R+\tau_A} \tag{13}$$

where we used the series expansion of the exponential function. Eq. (13) implies that we will get a correct peak estimate only when the release time constant is considerably longer than the attack time constant. Another side effect is that the attack time also gets slightly scaled by the release time: there will be a faster attack time than expected when we use a fast release time. Both problems are illustrated in Fig. 3.

In order to accomplish a program-dependent release behavior (auto release), some analog compressors use a combination of two release time constants in their peak detectors. One such design is found in the famous SSL Stereo Bus Compressor.[5] It uses two release networks stacked on top of each other. The much earlier Fairchild compressors and similar tube compressors used similar designs. When used in a compressor, the peak detector with dual time constant automatically increases the release time once the compression continues for a longer time period. This gives the desirable property of a shorter release time after compressing transients and a longer release time for steady state compression. However, this type of peak detector suffers from the same level problems.

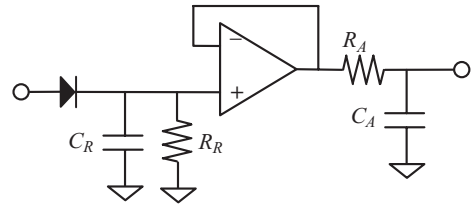

Fig. 4. Decoupled peak detector circuit.

2.3.3 Level Corrected Peak Detectors

Fig. 4 depicts an alternative peak detector where the subcircuits for attack and release are completely decoupled. The input signal is first fed through a peak detector with instantaneous attack. The result is a maximally fast peak estimation that is then buffered and smoothed by a first-order low-pass filter. The advantage of this is that the detector does not suffer from the level differences caused by different time constants exhibited by the standard peak detector circuit.

Using the derivations for the analog peak detector, a digital implementation of the decoupled peak detector is straightforward.

$$\begin{aligned} y_1[n] &= \max(x_L[n], \alpha_R y_1[n-1]) \\ y_L[n] &= \alpha_A y_L[n-1] + (1-\alpha_A) y_1[n] \end{aligned} \tag{14}$$

However the attack envelope is now also impressed upon the release envelope. A good estimate of what the actual release time would be is given by simply adding the respective attack time constant to the release time constant.

In a solely digital environment, we can efficiently fix this by adding a branch to the difference equation of the standard peak detector.

$$y_L[n] = \begin{cases} \alpha_A y_L[n-1] + (1-\alpha_A)x_L[n] & x_L[n] > y_L[n-1] \\ \alpha_R y_L[n-1] & x_L[n] \leq y_L[n-1] \end{cases} \tag{15}$$

The added branch ensures that the state variable is only discharged during the release phase. Such a peak detector was described in [10, 15].

Fig. 5 depicts the output of the decoupled and branching peak detector circuits, Eqs. (14) and (15) respectively, for different release time constants. The branching peak detector produces the intended release time constant, whereas the decoupled peak detector produces a measured time constant of approximately $\tau_R + \tau_A$. Ideally then, the decoupled peak

[5] www.ka-electronics.com/images/SSL/ssl_82E27.pdf

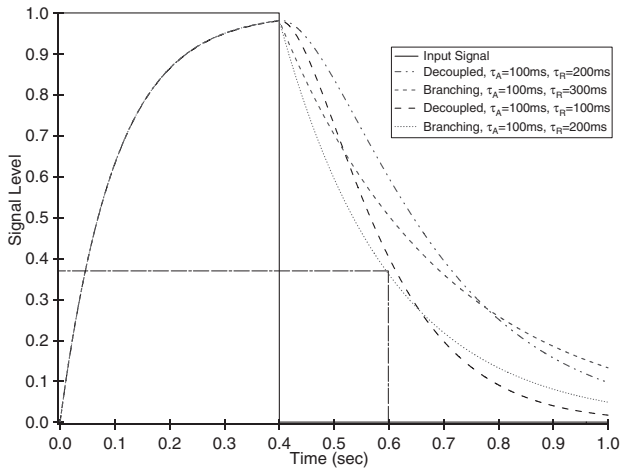

Fig. 5. Output of the decoupled and branching peak detector circuits for different release time constants. The branching peak detector produces the intended release time constant, whereas the decoupled peak detector produces a measured time constant of approximately $\tau_A+\tau_R$.

detector should be used with τ_R replaced by $\tau_R+\tau_A$. Instead, the attack time is often added to the release time for the branching peak detector. In which case, neither peak detector displays accurate time constants, but the release trajectories are roughly in agreement and it guarantees that the release time can never be shorter than the attack time.

2.3.4 Smooth Peak Detectors

All the analog designs only make use of the full release time if the input returns back to zero after a peak. However, when the signal settles on an intermediate plateau instead, the release envelope will simply stop at this point and the release time is much shorter than we would expect. The sudden stop creates a discontinuity in the envelope output. We can adjust the branching peak detector in order to ensure that it will always make use of the full release time.

$$y_L[n] = \begin{cases} \alpha_A y_L[n-1] + (1-\alpha_A)x_L[n] & x_L[n] > y_L[n-1] \\ \alpha_R y_L[n-1] + (1-\alpha_R)x_L[n] & x_L[n] \le y_L[n-1] \end{cases} \quad (16)$$

This peak detector, also used in [1, 15], now simply switches coefficients between the attack and the release phase.

The same modification can be introduced to the release mechanism of the decoupled peak detector, Eq. (14), if a continuous release mechanism is desired. So instead of having the capacitor discharge toward ground by the release resistor, it can release to the input signal, so:

$$\begin{aligned} y_1[n] &= \max(x_L[n], \alpha_R y_1[n-1] + (1-\alpha_R)x_L[n]) \\ y_L[n] &= \alpha_A y_L[n-1] + (1-\alpha_A)y_1[n] \end{aligned} \quad (17)$$

3 SYSTEM DESIGN

3.1 Feedback and Feedforward Design

There are two possible topologies for the side-chain, a feedback or feedforward topology. In the feedback topology the input to the side-chain is taken after the gain has been applied. This was traditionally used in early compressors and had the benefit that the side-chain could rectify possible inaccuracies of the gain stage.

The feedforward topology has the side-chain input before the gain is applied. This means that the side-chain circuit responsible for calculating the gain reduction, the gain computer, will be fed with the input signal. Therefore it will have to be accurate over the whole of the signal's dynamic range as opposed to a feedback type compressor where it will have to be accurate over a reduced dynamic range since the side-chain is fed with the compressors output. This bears no implications to a digital design but it should be taken under consideration if designing an analog feedforward compressor. Most modern compressors are based on the feedforward design.

By combining Eqs. (1) and (3) we can calculate the control voltage either from the input or from the output of the compressor, leading to the two topologies of feedforward and feedback compression.

Both feedforward and the feedback compressors can be implemented by subtracting the threshold from the signal level (which is either derived from the input or from the output of the compressor), applying halfwave rectification and multiplying the result by the slope variable (which is different for feedforward and feedback compressors).

The feedback design has a few limitations, such as the inability to allow a look-ahead function or to work as a perfect limiter due to the infinite negative amplification needed. From Eqs. (1) and (3), when, $x_L > T$, the control voltage for the feedback compressor is calculated as:

$$c_{dB}[n] = (1-R)(y_G[n-1] - T) \quad (18)$$

where we have assumed a hard knee and no attack or release. A limiter (with a ratio of ∞: 1) would need infinite negative amplification to calculate the control voltage. This is why feedback compressors are not capable of perfect limiting.

In contrast, the control voltage of the feedforward compressor, for $x_L > T$, is:

$$c_{dB}[n] = (1/R - 1)(x_G[n] - T) \quad (19)$$

Here, since the slope for a limiter approaches –1, implementing limiting is not a problem using a feedforward compressor. The feedforward compressor is also able to smoothly go into over-compression (with $r < 0$); the slope variable simply becomes smaller than –1.

3.1.1 An Alternate Digital Feedback Compressor

As an alternative, we can apply the previous gain value to the current input sample in order to estimate the current output sample. Since the compressor's gain is likely to change only by small amounts from sample to sample, the estimation is a fairly good one. Since the side-chain is now fed by the input signal, we could instead feed in an arbitrary signal

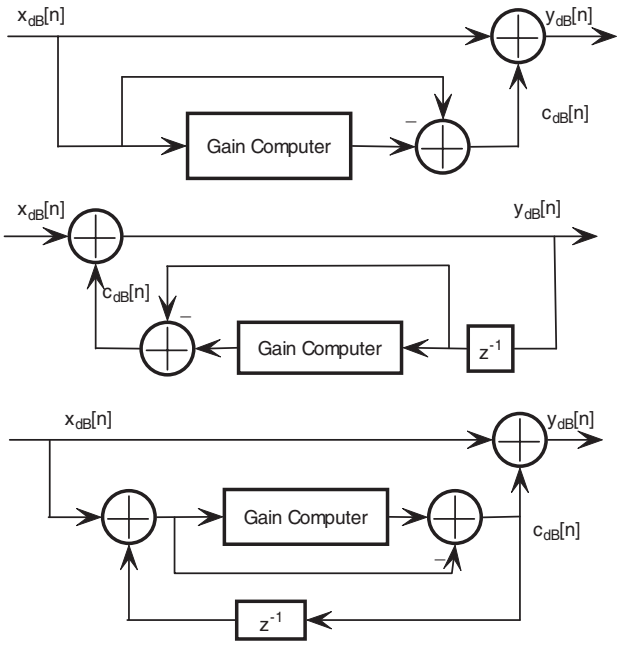

Fig. 6. Static compressor block diagrams showing operation on decibel level representations of signals. Feedforward design on top, feedback design in the middle, and alternate feedback design on bottom.

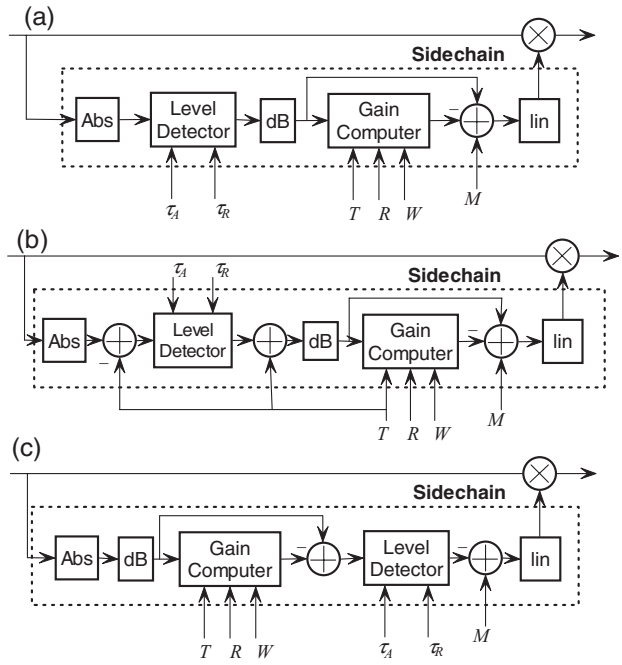

Fig. 7. Block Diagrams of the Compressor Configuration. (a) the return-to-zero detector; (b) the return-to-threshold detector; (c) the log domain detector.

in the side-chain, in order to achieve side-chaining operation (e.g., to implement a ducker or a de-esser). A similar technique has been used in analog compressors such as the SSL Bus Compressor (by using a slave VCA that mirrors the main VCA's action) and proposed for digital designs [10].

Fig. 6 presents the block diagrams for the feedforward, feedback and alternate feedback compressors.

3.2 Detector Placement

There are various choices for the placement of the detector circuit inside the compressor's circuitry, as depicted in Fig. 7 for feedforward compressors. In [8, 12, 15-17], the suggested position for the detector is within the linear domain and before the decibel converter and gain computer. That is,

$$\begin{aligned} x_L[n] &= |x[n]| \\ x_G[n] &= 20\log_{10} y_L[n] \\ c_{dB}[n] &= y_G[n] - x_G[n] \end{aligned} \tag{20}$$

However, the detector circuit then works on the full dynamic range of the input signal, while the gain computer only starts to operate once the signal exceeds the threshold (the control voltage will be zero for a signal below the threshold). The result is that we again experience a discontinuity in the release envelope when the input signal falls below the threshold. It also generates a lag in the attack trajectory since the detector needs some time to charge up to the threshold level even if the input signal attacks instantaneously.

A way to overcome this and achieve a smooth release trajectory and avoid the envelope discontinuity is by leaving the detector in the linear domain but bias it at the threshold level. That is, subtract the threshold from the signal before it enters the detector and add the threshold back in again after the signal has left it.

$$\begin{aligned} x_L[n] &= |x[n]| - 10^{T/20} \\ x_G[n] &= 20\log_{10}(y_L[n] + 10^{T/20}) \\ c_{dB}[n] &= y_G[n] - x_G[n] \end{aligned} \tag{21}$$

This turns the return-to-zero detector into a return-to-threshold type, and the envelope will smoothly fade out once the signal falls below the threshold. However, this solution becomes problematic once we start using a soft knee characteristic. Since the threshold becomes a smooth transition instead of a hard boundary, we do not have a fixed value with which to bias the detector.

A similar approach is to place the detector in the linear domain but after the gain computer [18–19]. In this case the detector circuit works directly on the control voltage.

$$\begin{aligned} x_G[n] &= 20\log_{10}|x[n]| \\ x_L[n] &= 10^{y_G[n]/20} \\ c[n] &= y_L[n] \end{aligned} \tag{22}$$

For each of Eqs. (20), (21), and (22), since the release envelope discharges exponentially toward zero in the linear domain, this is equivalent to a linear discharge in the log (or decibel) domain. Although the release rate in terms of decibels per time is now constant, it also means that the release time will be longer for heavier and shorter for lighter compression. Unfortunately, this is not what the ear perceives as a smooth release trajectory and the return-to-zero detector in the linear domain is therefore mostly used

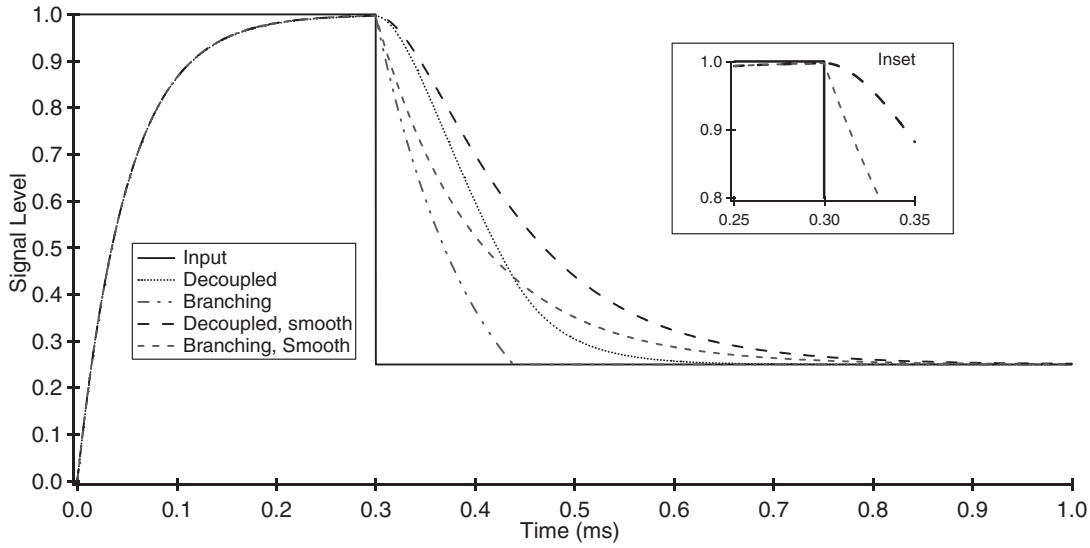

Fig. 8. Output of the various peak detector circuits. Attack Time = 50 ms, release time = 100 ms. Inset, a close-up of the decoupled smoothed and branching smoothed peak detectors as they respond to a sudden level change at 300 ms.

by compressors where artifacts are generated on purpose [20]. Therefore, the preferred position for the detector is within the log domain and after the gain computer [21–22].

$$\begin{aligned} x_G[n] &= 20\log_{10}|x[n]| \\ x_L[n] &= x_G[n] - y_G[n] \\ c_{dB}[n] &= -y_L[n] \end{aligned} \tag{23}$$

Now, the detector directly smooths the control voltage instead of the input signal. Since the control voltage automatically returns back to zero when the compressor does not attenuate, we do not depend on a fixed threshold and a smooth release envelope is guaranteed. The trajectory now behaves exponentially in the decibel domain, which means that the release time is independent of the actual amount of compression. This behavior seems smoother to the ear since the human sense of hearing is roughly logarithmic [23]. It is therefore used in most compressors that want to achieve smooth and subtle (artifact-free) compression characteristics for use with complicated signals (such as program material).

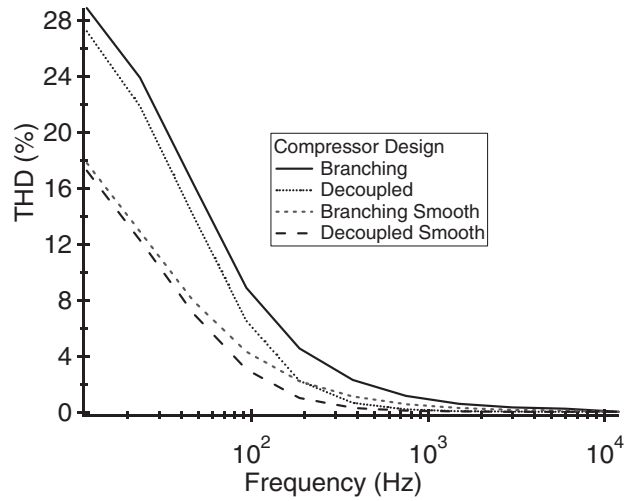

Fig. 9. Total harmonic distortion as a function of input signal. The input signal was a pure sinusoid with unity amplitude. Threshold, ratio, attack, release, and knee were set to –20 dB, 7, 1 ms, 15 ms, and 0 dB, respectively.

4. COMPARISON OF COMPRESSOR DESIGNS

4.1 Level Detector Performance

Fig. 8 shows the output of the various peak detector circuits for the same input. As can be seen, all envelopes reach the maximum peak value and feature similar attack trajectories. However, the release envelopes are far too short for the decoupled and branching peak detectors, (14) and (15), whereas the smoothed versions, (17) and (16), make full use of the release time. In the inset, a discontinuity can be seen for the smooth, branching peak detector, Eq. (16), since it produces an abrupt switch from attack to release, whereas the release envelope is continuous for the smoothed, decoupled peak detector, Eq. (17).

Most objective measures of compressor performance have come from hearing aid design research. These often ignore the effects of distortion, which are considered negligible when speech intelligibility is the fundamental concern [24]. In Fig. 9, the total harmonic distortion (THD) is given for compressors based on the four peak detector designs as a function of the input signal frequency. The effect of smoothing in minimizing harmonic distortion is clearly evident. Furthermore, the decoupled peak detector offers a slight improvement over the branching design, with and without smoothing, since it minimizes the introduction of distortion components at the transition between attack and release phases.

One of the main metrics used in assessing compressors used in cochlear implants and hearing aids is the effective

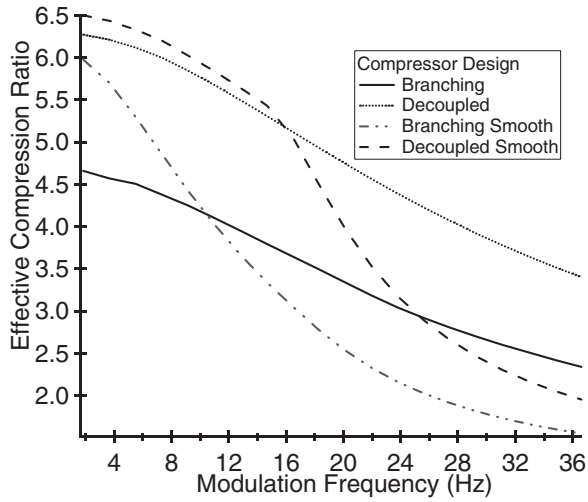

Fig. 10. Effective compression ratio as a function of modulation frequency. The input signal and compressor parameters were set as in Fig. 9.

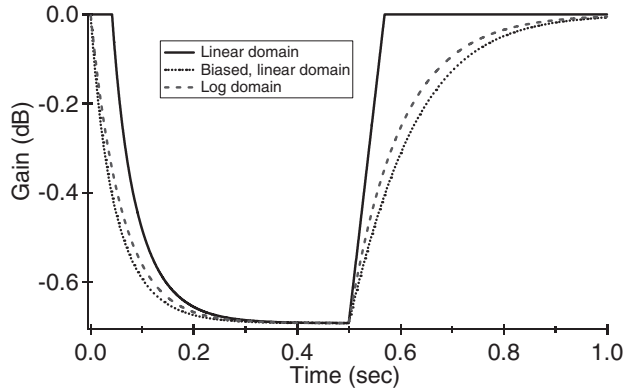

Fig. 11. Compressor envelope trajectories in the linear domain for the linear domain return-to-zero detector (solid), linear domain return-to-threshold detector (dotted), and log domain detector (dashed).

compression ratio [24–25], R_{eff}. This is found by measuring the modulation of a signal at the output given a certain modulation at the input. Using the method described in [24], an amplitude modulated sine wave,

$$x[n] = (1 + m\cos(2\pi f_m n))\cos(2\pi f_c n), \tag{24}$$

is applied at the input. The spectrum consists of a carrier and two side bands. The difference between the amplitude of the side bands and the amplitude of the carrier is ΔS_i. Compression is applied, and the difference between the amplitude of the side bands and the amplitude of the carrier of the compressed signal is found, ΔS_o. The effective compression ratio is then given by $\Delta S_i/\Delta S_o$.

The effective compression ratio as a function of the modulation frequency is given in Fig. 10. The decoupled designs generally perform closer to the target compression ratio of 7:1. This is primarily due to the lack of full use of the release time in the branching compressors. However, the nonsmooth decoupled design actually outperforms the smooth decoupled design for high modulation frequencies, since the rapid decay of the compression ratio during release prevents the low amplitude modulator from being overly compressed.

4.2 Comparison of Side-Chain Configurations

Fig. 11 compares the three different detector options in terms of the linear gain envelope they produce. The attack lag and the release discontinuity of the linear domain return-to-zero detector can be clearly seen. The linear domain return-to-threshold and log domain detector curves look similar, though the former features faster attack and slower release behavior than the latter.

One further metric from prior work on compressors for improved speech intelligibility can be used. The Fidelity of the Envelope Shape (FES) is intended to measure how well the envelope of the signal is maintained [25]. It is based on a measure of the correlation between the envelopes (in dB) of the compressed and uncompressed signals. Although it is highly dependent on the envelope estimation method, it has the added benefit that it does not require an artificial test signal.

Table 1 provides the FES measures for compressors applied to four signals. Only minor differences can be seen between the FES for the various log domain approaches. The slightly worse performance of the smooth compressors may be due to the fact that the signal is always in attack or release phase, whereas the nonsmooth compressors have times where the envelope is completely unchanged. More noticeable, however, is the difference in FES for different level detector placements.

The log domain detectors clearly outperform linear domain detectors on all signals except vocals (where there is only a minimal difference in FES).

5 CONCLUSION

Based on the analysis and design choices presented, we can propose a compressor configuration that serves the goals of ease of modification and avoidance of artifacts.

Feedforward compressors are preferred since they are more stable and predictable than the feedback type ones, and high dynamic range problems do not occur with digital designs. The detector is placed in the log domain after the gain computer, since this generates a smooth envelope, has no attack lag, and allows for easy implementation of variable knee width. For the compressor to have smooth performance on a wide variety of signals, with minimal artifacts and minimal modification of timbral characteristics, the smooth, decoupled peak detector should be used. Alternately, the smooth, branching peak detector could be used in order to have more detailed knowledge of the effect of the time constants, although this may yield discontinuities in the slope of the gain curve when switching between attack and release phases.

The analysis in this paper was limited to the design of standard compressors. Their uses and recommended parameter settings are discussed in detail in other texts. Variations

Table 1. Fidelity of the Envelope Shape for compression applied to four acoustic signals. In each case, a strong, fast acting compressor was applied; threshold –40 dB, ratio 10:1, attack time 1 ms, release time 40 ms, knee 20 dB.

		Fidelity of Envelope Shape (FES)			
Detector Placement	Detector Type	Guitar	Bass	Drums	Vocals
Log domain	Branching	0.884	0.945	0.766	0.952
	Decoupled	0.899	0.932	0.755	0.941
	Branching Smooth	0.852	0.927	0.640	0.941
	Decoupled Smooth	0.859	0.911	0.648	0.936
Linear domain	Branching	0.836	0.879	0.517	0.932
	Decoupled	0.790	0.831	0.537	0.930
	Branching Smooth	0.825	0.856	0.456	0.932
	Decoupled Smooth	0.775	0.805	0.461	0.934

and more advanced designs, such as side-chain filtering or multiband compressors, were also beyond the scope of this paper. Design of an intelligent compressor, where the signal is analyzed in order to adapt the parameter settings to the audio content, represents an interesting future direction and will be presented in related work by the authors [26].

5 REFERENCES

[1] P. Dutilleux, et al., "Nonlinear Processing, Chap. 4," in *Dafx:Digital Audio Effects*, U. Zoelzer, Ed. (2nd ed: Wiley, John & Sons, 2011), p. 554.

[2] J. O. Smith, *Introduction to Digital Filters with Audio Applications* (Booksurge Llc, 2007).

[3] R. Izhaki, *Mixing Audio: Concepts, Practices and Tools* (Focal Press, 2008).

[4] B. Lachaise and L. Daudet, "Inverting Dynamics Compression with Minimal Side Information," in *International Conference on Digital Audio Effects (DAFx)* (Helsinki, 2008).

[5] B. Rudolf, "1176 Revision History," *Mix Magazine Online* (June 1 2000).

[6] D. Barchiesi and J. D. Reiss, "Reverse Engineering the Mix," *J. Audio Eng. Soc.*, vol. 58, pp. 563–576 (2010 Jul./Aug.).

[7] M. Bosi, "Low-Cost/High-Quality Digital Dynamic Range Processor," presented at the *91st Convention of the Audio Engineering Society* (1991 Oct.), convention paper 3133.

[8] R. J. Cassidy, "Level Detection Tunings and Techniques for the Dynamic Range Compression of Audio Signals," presented at the *117th Convention* of the Audio Engineering Society (2004 Oct.), convention paper 6235.

[9] F. Floru, "Attack and Release Time Constants in RMS-Based Feedback Compressors," *J. Audio Eng. Soc.*, vol. 47, pp. 788–804 (1999 Oct.).

[10] G. W. McNally, "Dynamic Range Control of Digital Audio Signals," *J. Audio Eng. Soc.*, vol. 32, pp. 316–327 (1984 May).

[11] B. E. Lobdell and J. B. Allen, "A Model of the VU (Volume-Unit) Meter, with Speech Applications," *J. Acous. Soc. Am.*, vol. 121, pp. 279–285 (2007).

[12] J. S. Abel and D. P. Berners, "On Peak-Detecting and RMS Feedback and Feedforward Compressors," presented at the *115th Convention of the Audio Engineering Society* (2003 Oct.), convention paper 5914.

[13] D. Berners, "Analysis of Dynamic Range Control (DRC) Devices," *Universal Audio WebZine*, vol. 4 (September 2006).

[14] P. Dutilleux and U. Zölzer, "Nonlinear Processing, Chap. 5," in *DAFX - Digital Audio Effects*, U. Zoelzer, Ed. (1st ed: Wiley, John & Sons, 2002).

[15] U. Zolzer, *Digital Audio Signal Processing*, 2nd ed. (John Wiley and Sons, Ltd., 2008).

[16] P. Hämäläinen, "Smoothing of the Control Signal Without Clipped Output in Digital Peak Limiters," in *International Conference on Digital Audio Effects (DAFx)* (Hamburg, Germany , 2002), pp. 195–198.

[17] S. J. Orfanidis, *Introduction to Signal Processing* (Orfanidis [prev. Prentice Hall], 2010).

[18] J. Bitzer and D. Schmidt, "Parameter Estimation of Dynamic Range Compressors: Models, Procedures and Test Signals," presented at *the 120th Convention of the Audio Engineering Society* (2006 May), convention paper 6849.

[19] L. Lu, "A Digital Realization of Audio Dynamic Range Control," in *Fourth International Conference on Signal Processing Proceedings (IEEE ICSP)* (1998), pp. 1424–1427.

[20] Sonnox, "Dynamics Plug-In Manual," *Sonnox Oxford Plug-ins*, April 1st 2007.

[21] S. Wei and W. Xu, "FPGA Implementation of Gain Calculation Using a Polynomial Expression for Audio Signal Level Dynamic Compression," *J. Acoust. Soc. Jpn. (E)*, vol. 29, pp. 372–377 (2008).

[22] S. Wei and K. Shimizu, "Dynamic Range Compression Characteristics Using an Interpolating Polynomial for Digital Audio Systems," *IEICE Trans. Fundamentals*, vol. E88-A, pp. 586–589 (2005).

[23] H. Fastl and E. Zwicker, *Psychoacoustics: Facts and Models*, 3rd ed. (Springer, 2007).

[24] J. Verschuure, et al., "Compression and its Effect on the Speech Signal," *Ear Hear*, vol. 17, pp. 162–175 (1996).

[25] M. A. Stone and B. C. J. Moore, "Effect of the Speed of a Single-Channel Dynamic Range Compressor on Intelligibility in a Competing Speech Task," *J. Acoust. Soc. Am.*, vol. 114, pp. 1023–1034 (2003).

[26] D. Giannoulis, M. Massberg and J. D. Reiss, "Parameter Automation in a Dynamic Range Compressor," submitted to *J. Audio Eng. Soc.*, 2011.

THE AUTHORS

Dimitrios Giannoulis

Michael Massberg

Joshua D. Reiss

Dimitrios Giannoulis studied physics at the National University of Athens, Greece, and he specialized, among others, on signal processing and acoustics. He received a Master's degree in music signal processing in 2010 from Queen Mary University of London. He is currently, pursuing a Ph.D. degree in the Centre for Digital Music (C4DM) in Queen Mary University of London. His main research areas of interest are machine learning, digital signal processing, audio and speech processing and music information retrieval.

●

Michael Massberg studied media technology at FH Oldenburg / Ostfriesland / Wilhelmshaven in Emden, Germany and digital music processing at Queen Mary University of London, UK, where he received a Master's degree in 2009. He has been with German pro-audio company Brainworx since 2008, working as as an R&D engineer and specializing in digital modeling of analog recording studio hardware.

●

Dr. Josh Reiss is a senior lecturer with the Centre for Digital Music at Queen Mary University of London. He received his Ph.D. in physics from Georgia Tech, specializing in analysis of nonlinear systems. Dr. Reiss has published over 100 scientific papers and serves on several steering and technical committees. He has investigated music retrieval systems, time scaling and pitch shifting techniques, polyphonic music transcription, loudspeaker design, automatic mixing for live sound and digital audio effects, among others. His primary focus of research, which ties together many of the above topics, is on the use of state-of-the-art signal processing techniques for professional sound engineering.