Audio Engineering Society

Convention e-Brief 484

Presented at the 145th Convention
2018 October 17 – 20, New York, NY, USA

Adaptive Ballistics Control of Dynamic Range Compression for Percussive Tracks

David Moffat[1] and Mark B. Sandler[1]

[1] *Centre for Digital Music, Queen Mary University of London, London, UK*

Correspondence should be addressed to David Moffat (`d.j.moffat@qmul.ac.uk`)

ABSTRACT

Dynamic range compression (DRC) is a very commonly used audio effect. One use of DRC is to emphasise transients in an audio signal. The aim of this paper is to present an approach for automatically setting dynamic range compression timing parameters, adaptively, allowing parameters to adapt to the incoming audio signal, with the aim of emphasising transients within percussive audio tracks. An implementation approach is presented.

1 Introduction

Dynamic range compression (DRC) is a one of the most commonly used audio effects within a music mixing context. Stasis et al. [1] showed that DRC is the second most used audio effect, after an equaliser. DRC is the most common example of an adaptive audio effect, where a side-chain process calculates some changing gain parameter which is then applied to a signal based on an initial signal. An overview of different compressors was performed by Giannoulis et al. [2].

DRC are typically used to perform one of three different actions: reduce transients, emphasise transients or reduce overall dynamic range. DRC are a tool that can often be overused, to create a louder perception of a song, though reducing the dynamic range [3]. Both the range and loudness are vital aspects for a mix engineer to control [4]. Typically in the context of drums, a compressor will be used to emphasise the punch and transient of a drum [5]. The aim of this work, is to present and develop an automated compressor that can adaptively control and change parameters, based on audio signal analysis, with the aim of emphasising the transients of a drum hit.

2 Background

There is a body of research in automatic audio engineering tools, starting from analysis and evaluation of mix engineers within the studio [6], to automatic gain and fader control [7, 8], to more complex effects, such as reverberation [9, 10] and subgrouping structure [11, 12]. Typical automatic mixing approaches rely on extracting audio features [13], which are then used to drive traditional control parameters.

Scott et al. [14] produced an automatic mixing system, which mixed together a range of drums into a single drum stem, before combining it into the original track. Terrell et al. [15] demonstrated the use of noise gates in improving audio quality by performing noise removal in drums. It has been shown that in the context of drums, there is a difference in preferred compression parameters, dependant on the audio content and style of the music [16].

Massberg [17] produced an overview of compressors and provided a method for automating 4 different parameters based on audio signal analysis, the attack, release, knee and makeup gain. The ratio was set to a constant $1:\infty$ and the threshold parameter was user controlled. This work was extended by [18], who compared two different compressor setups, *temporal mode*, taken from [17] and *spectral mode*, where attack and release parameters were control by the spectral flux, rather than the crest factor. Maddams et al. [19] also produced two different semi-automatic compressors, where the attack and release parameters were the *spectral mode* attack and release parameters from [18]. A *threshold mode* and a *ratio mode* were presented, where users could control either the threshold or the ratio, thus making it semi-automatic. In the *threshold mode*, ratio was set to $1:\infty$, and in *ratio mode*, the threshold was set to the incoming signal $RMS(x) - 12dB$. Only one other compressor in literature automated all parameters, making an atonomous compressor. Ma et al. [20] performed user studies where participants were asked to apply compression to a set of audio tracks, and curve fitting of parameter values to audio feature spaces was performed. The same attack and a very similar release and knee parameter to other work was used.

All existing work, with the exception of [20], set the compressor ratio to infinity, or made it a user controllable parameter. As such, it can be justified that only [20] produced an automated compressor, [19] produced a single control compressor and all other work is a single control limiter [17, 19, 18]. A full table of control parameters and the audio features they are proportional to, are presented in Table 1. In both modes, [19] controlled the knee of the compressor, either as a fixed parameter, or as the user controlled parameter, the input threshold.

With the attack and release parameters of compressors, it is common to have a short attack and long release, or long attack and short release [21]. Particularly when using a compressor to emphasise transients of a signal, an engineer would typically tune the attack time to just past the fullness of the transient, and the release time is tied to the rhythm of the song. Currently all existing automated attack control methodologies will set a shorter attack for more transient signals and slower attack for more smooth signals, as can be seen in Figure 1. A higher crest factor or higher spectral flux indicates a more transient signal. Similarly, with the release times, existing automated compressors apply a

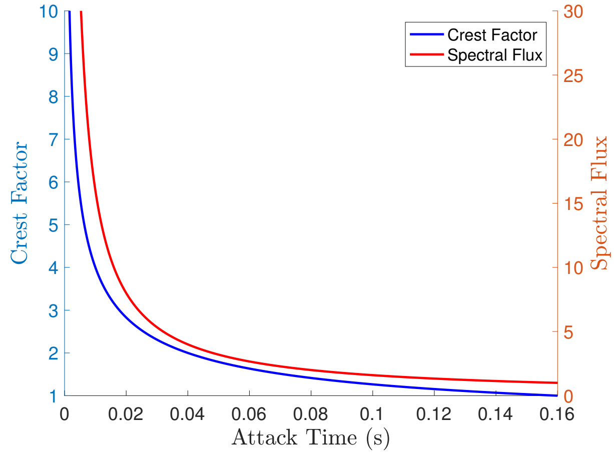

Fig. 1: Attack times of existing compressor parameter automation

quicker release for more transient signals and slower release for smoother less transient signals, as shown in Figure 2. One of the potential issues with these settings, is that the attack and release are fairly highly correlated with each other, as can be seen in Figure 3.

3 Compressor Implementation

For the purposes of this implementation, we use a standard feedforward, peak detection compressor. In this particular case, the purpose of our compressor is to emphasise the transients and suppress transient gaps within a drum track. The compressor was implemented using the Matlab audio plugin framework, for easy transferal to VST [22]. The Matlab implementation is available online[1].

3.1 Attack

To calculate the attack time parameter τ_a of our drum, a short 3 second envelope of the audio track was used. From this, an estimate of the decay envelope from each individual impacts was extracted. The T_{20} is defined as the duration of time taken for a signal peak to decay by 20dB. The mean T_{20} was taken for the attack parameter.

$$E = H_{db}(x) - max(H_{db}(x)) \tag{1}$$

where $H_{db}(x)$ is the Hilbert envelope of the signal in dB, E is the normalised signal envelope and E'_{win} is the

[1] https://github.com/djmoffat/intelligentCompressor

Table 1: Summary of compressor parameters within existing work

Reference	Attack	Release	Threshold	Ratio	Knee	Makeup
Massberg [17]	$\frac{160ms}{C^2}$	$\frac{2000ms}{C^2} - A$	User Control	∞	$\propto T$	$\propto -T$
Maddams et al. [19] - *Threshold*	$\frac{160ms}{SF}$	$\frac{2000ms}{SF^{0.8}} - A$	User Control	∞	$\lvert T \rvert$	L_i
Maddams et al. [19] - *Ratio*	$\frac{160ms}{SF}$	$\frac{2000ms}{SF^{0.8}} - A$	RMS(x) - 12dB	User Control	3dB	L_i
Giannoulis et al. [18] - *Temporal*	$\frac{160ms}{C^2}$	$\frac{2000ms}{C^2} - A$	User Control	∞	$\propto T$	$\propto -T$
Giannoulis et al. [18] - *Spectral*	$\frac{160ms}{SF}$	$\frac{2000ms}{SF^{0.8}} - A$	User Control	∞	$\propto T$	L_i
Ma et al. [20]	$\frac{160ms}{C^2}$	$\frac{2000ms}{C^2}$	$\propto RMS(x) - e^{(C-C_m)}$	$\propto e^{(C-C_m)} + LER$	$\frac{\lvert T \rvert}{2}$	L_i

C = crest factor, C_m = mean crest factor of all tracks, SF = Spectral Flux, LER = low frequency energy ratio, L_i = Integrated Loudness, A = attack, T = threshold and R = ratio

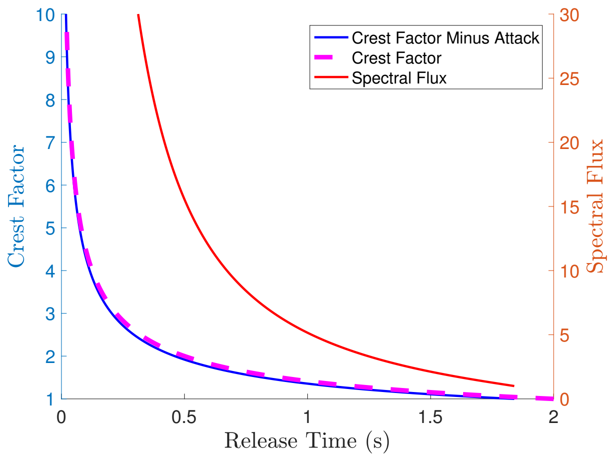

Fig. 2: Release times of existing compressor parameter automation

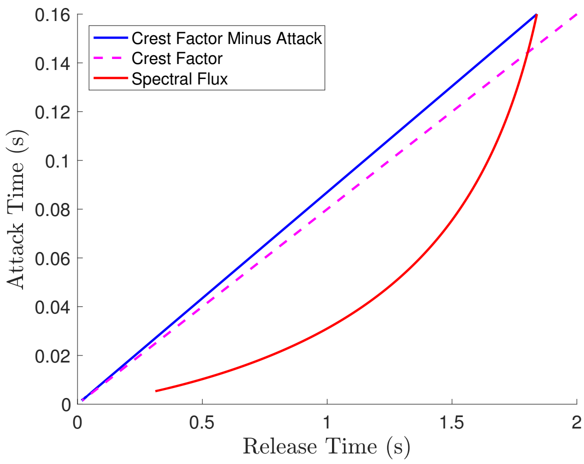

Fig. 3: Attack Time vs. Release Time

derivative of the envelope over a specific window range. Peak selection and smoothing was performed, and the T_{20} was calculated, based on work by [23], as

$$\tau_a = T_{20} = \frac{-20dB}{E'_{win}} \tag{2}$$

As such, the attack was driven by the envelope of the signal, with the intention of allowing the entire transient through, and kicking in just at the tail end of the transient.

3.2 Release

As the release function was designed to be driven by the meter of the audio track, a basic beat tracking method was used to detect the beat of the song. The spectral flux was used to determine onsets [24],

$$SF = \sum_{k=0}^{N} (\rho(|X_k(n)| - |X_k(n-1)|))^2 \tag{3}$$

$$\rho(x) = \frac{x + |x|}{2} \tag{4}$$

such that $\rho(x)$ is the half-wave rectification function, where $X_k(n)$ is the fourier transform signal of x, for frequency bin k at timestep n. Peak picking was then performed, and Inter-Onset Interval (IOI) values were smoothed to estimate a short time varying beat, and the mean IOI was calculated as $\overline{IOI}$. The release time τ_r was then calculated as

$$\tau_r = \overline{IOI} - \tau_a \tag{5}$$

4 Conclusion

Adaptive control of timing characteristics of a dynamic range compressor was presented. The timing control parameters were defined and set based on the approach that a mix engineer might take when trying to emphasise the transients of a percussive track.

There is demand for assistive mixing tools, and through developing approaches that can create a more punchy drum track. Once an automated approach can be setup, it is then possible to create semantically meaningful control interfaces, allowing for new and more intuitive approaches to moving through a search space.

There is further potential for extending this work into a fully automated adaptive compressor. There is a need for in depth evaluation as to the effectiveness of our compressor, both through large scale subjective evaluation [25], with more participants and more use cases. Further, more detailed and specific analysis of mix evaluation would also offer further insight [26, 27]. Further evaluation of the compressor, including and rigorous comparisons of each of the individual parameter settings would assist in producing a better understanding as to when the proposed compressor works well, and where it fails in comparison to the state of the art. There is also scope for objective evaluation of the effectiveness of making audio signals more punchy, through a perceptual model [28].

Analysis of large scale data of users controlling parameters settings, could present a data-driven approach to parameter settings, and could produce a rich and interesting opportunity for development of further intelligent audio effects.

Acknowledgments

This paper is supported by EPSRC Grant EP/L019981/1, Fusing Audio and Semantic Technologies for Intelligent Music Production and Consumption.

References

[1] Stasis, S., Jillings, N., Enderby, S., and Stables, R., "Audio processing chain recommendation," in *Proceedings of the 20th International Conference on Digital Audio Effects*, Edinburgh, UK, 2017.

[2] Giannoulis, D., Massberg, M., and Reiss, J. D., "Digital dynamic range compressor design—A tutorial and analysis," *Journal of the Audio Engineering Society*, 60(6), pp. 399–408, 2012.

[3] Vickers, E., "The loudness war: Background, speculation, and recommendations," in *Audio Engineering Society Convention 129*, Audio Engineering Society, 2010.

[4] Wichern, G., Wishnick, A., Lukin, A., and Robertson, H., "Comparison of loudness features for automatic level adjustment in mixing," in *Audio Engineering Society Convention 139*, Audio Engineering Society, 2015.

[5] Owsinski, B., *The Mixing Engineer's Handbook*, Nelson Education, 2013.

[6] Pestana, P. D. L. G., *Automatic mixing systems using adaptive digital audio effects*, Ph.D. thesis, Universidade Católica Portuguesa, 2013.

[7] Dugan, D., "Automatic microphone mixing," *Journal of the Audio Engineering Society*, 23(6), pp. 442–449, 1975.

[8] Perez Gonzalez, E. and Reiss, J. D., "An automatic maximum gain normalization technique with applications to audio mixing." in *Audio Engineering Society Convention 124*, Audio Engineering Society, 2008.

[9] Chourdakis, E. T. and Reiss, J. D., "A Machine-Learning Approach to Application of Intelligent Artificial Reverberation," *Journal of the Audio Engineering Society*, 65(1/2), pp. 56–65, 2017.

[10] Benito, A. L. and Reiss, J. D., "Intelligent multitrack reverberation based on hinge-loss markov random fields," in *Audio Engineering Society Conference: 2017 AES International Conference on Semantic Audio*, Audio Engineering Society, 2017.

[11] Ronan, D., Moffat, D., Gunes, H., and Reiss, J. D., "Automatic Subgrouping of Multitrack Audio," in *Proc. 18th International Conference on Digital Audio Effects (DAFx-15)*, DAFx-15, 2015.

[12] Ronan, D., Ma, Z., Mc Namara, P., Gunes, H., and Reiss, J. D., "Automatic Minimisation of Masking in Multitrack Audio using Subgroups," *ArXiv e-prints*, 2018.

[13] Moffat, D., Ronan, D., and Reiss, J. D., "An Evaluation of Audio Feature Extraction Toolboxes," in *Proc. 18th International Conference on Digital Audio Effects (DAFx-15)*, 2015.

[14] Scott, J., Prockup, M., Schmidt, E. M., and Kim, Y. E., "Automatic multi-track mixing using linear dynamical systems," in *Proceedings of the 8th Sound and Music Computing Conference, Padova, Italy*, p. 12, 2011.

[15] Terrell, M., Reiss, J. D., and Sandler, M., "Automatic noise gate settings for drum recordings containing bleed from secondary sources," *EURASIP Journal on Advances in Signal Processing*, 2010, p. 10, 2010.

[16] Bromham, G., Moffat, D., Fazekas, G., Barthet, M., and Sandler, M. B., "The impact of compressor ballistics on the perceived style of music," in *Audio Engineering Society Convention 145 (to appear)*, New York, USA, 2018.

[17] Massberg, M., *Investigation in Dynamic Range Compression*, Master's thesis, Queen Mary University of London, London, UK, 2009.

[18] Giannoulis, D., Massberg, M., and Reiss, J. D., "Parameter automation in a dynamic range compressor," *Journal of the Audio Engineering Society*, 61(10), pp. 716–726, 2013.

[19] Maddams, J. A., Finn, S., and Reiss, J. D., "An autonomous method for multi-track dynamic range compression," in *Proceedings of the 15th International Conference on Digital Audio Effects (DAFx-12)*, pp. 1–8, Citeseer, 2012.

[20] Ma, Z., De Man, B., Pestana, P. D., Black, D. A., and Reiss, J. D., "Intelligent multitrack dynamic range compression," *Journal of the Audio Engineering Society*, 63(6), pp. 412–426, 2015.

[21] Senior, M., *Mixing Secrets*, Focal Press, 2012.

[22] DeVane, C. and Bunkheila, G., "Automatically Generating VST Plugins from MATLAB Code," in *Audio Engineering Society Convention 140*, Audio Engineering Society, 2016.

[23] Lundeby, A., Vigran, T. E., Bietz, H., and Vorländer, M., "Uncertainties of measurements in room acoustics," *Acta Acustica united with Acustica*, 81(4), pp. 344–355, 1995.

[24] Bello, J. P., Daudet, L., Abdallah, S., Duxbury, C., Davies, M., and Sandler, M. B., "A tutorial on onset detection in music signals," *IEEE Transactions on speech and audio processing*, 13(5), pp. 1035–1047, 2005.

[25] Moffat, D. and Reiss, J. D., "Perceptual Evaluation of Synthesized Sound Effects," *ACM Transactions on Applied Perception (TAP)*, 15(2), p. 19, 2018, doi:10.1145/3165287.

[26] De Man, B., Jillings, N., Moffat, D., Reiss, J. D., and Stables, R., "Subjective comparison of music production practices using the Web Audio Evaluation Tool," in *2nd AES Workshop on Intelligent Music Production*, 2016.

[27] De Man, B., Reiss, J. D., and Stables, R., "Ten Years of Automatic Mixing," in *3rd AES Workshop on Intelligent Music Production*, 2017.

[28] Fenton, S. and Lee, H., "Towards a Perceptual Model of "Punch" in Musical Signals," in *Audio Engineering Society Convention 139*, Audio Engineering Society, 2015.