

MUSIKVERB: A HARMONICALLY ADAPTIVE AUDIO REVERBERATION

João Paulo Caetano Pereira

University of Porto, Faculty of Engineering,
MIEEC
Porto, Portugal
ee12035@fe.up.pt

Gilberto Bernardes

INESC TEC and University of Aveiro
Portugal
gba@inesctec.pt

Rui Penha

INESC TEC and University of Porto, Faculty
of Engineering
Porto, Portugal
rui.penha@inesctec.pt

ABSTRACT

We present MusikVerb, a novel digital reverberation capable of adapting its output to the harmonic context of a live music performance. The proposed reverberation is aware of the harmonic content of an audio input signal and ‘tunes’ the reverberation output to its harmonic content using a spectral filtering technique. The dynamic behavior of MusikVerb avoids the sonic clutter of traditional reverberation, and most importantly, fosters creative endeavor by providing new expressive and musically-aware uses of reverberation. Despite its applicability to any input audio signal, the proposed effect has been designed primarily as a guitar pedal effect and a standalone software application.

1. INTRODUCTION

Adaptive digital audio effects (ADAFx) are a class of audio effects, whose control parameters are mapped to attributes of the audio input signal to be transformed [1]. This level of symbiotic information exchange between an input signal and control parameters of the transformation effect has attracted the attention of academia and industry over the last decade as a new strategy for music creation [2].

The mappings between audio input attributes and effect parameters are central to ADAFx [3]. In this context, we can understand the emergence of ADAFx in light of the breakthroughs in audio-content processing for audio signals description, which have been proposed by the signal processing and music information retrieval communities.

Within the academic literature several ADAFx studies and prototype applications have been proposed [1, 4, 5]. These contributions focus mostly on mapping strategies between signal attributes and effect parameters [1]. Within industry and for the specific case of the guitar, the target instrument of our study, the following three commercial ADAFx have been recently identified in [3]: ‘TE-2 Tera Echo’, ‘MO-2 Multi Overtone’ and ‘DA-2 Adaptive Distortion’ [6, 7, 8].

In this paper, we extend existing guitar ADAFx by proposing a harmonically adaptive audio reverberation as a guitar pedal effect and a standalone software application. To the best of our knowledge, the sole existing application that implements such an ADAFx is Zynaptiq’s Adaptiverb [9], for which no technical descriptions is known to be available.

In contrast to traditional digital reverberation, which models the physical phenomena of sound waves reflecting on enclosed space surfaces [4], MusikVerb aims at controlling the tonal clarity (understood as levels of consonance/dissonance) and harmonic richness of a reverberation tail. To this end, MusikVerb transforms the output of a traditional audio reverberation by filtering its output according to a ranked list of pitch classes (i.e., the twelve notes of

the chromatic scale) computed from the perceptual-inspired Tonal Interval Space space [10]. Given this ranked list of pitch classes, the user can then ‘tune’ the reverberated signal to the harmonic context of an audio input signal.

The remainder of this paper is organized as follows. Section 2 presents the architecture of the MusikVerb system and the information flow between its component modules. Section 3 presents the extraction of harmonic attributes from an audio input signal to create a ranked list of pitch classes according to their perceptual distance to an input audio signal. Section 4 details how a ranked pitch class list is mapped to a frequency-domain representation (i.e., spectrum). Section 5 describes an algorithm which filters an audio reverberation tail to ‘fit’ the harmonic context of a performance. Section 6 provides an overview of the user control parameters of MusikVerb in both hardware and software instantiations of the system. Section 7 details the creative applicability of MusikVerb as highlighted by expert musicians when interacting with the system. Finally, Section 8 states the conclusions of our work and future directions.

2. MUSIKVERB ARCHITECTURE

Fig. 1 shows the architecture of MusikVerb, which follows the threefold typical ADAFx structure: 1) extraction of audio attributes from an input signal; 2) mappings between audio attributes and effect parameters; and 3) the processing of the effect transformation [3].

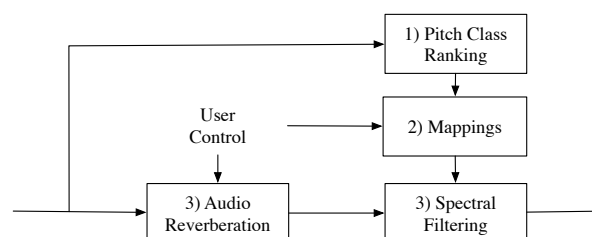


Figure 1: MusikVerb architecture. The audio signal flux flows from left to right between the (squared) component modules.

The harmonic content of an audio input signal is 1) analyzed to extract a ranked list of pitch classes according to a perceptual distance measure. 2) Then, a mapping between the ranked pitch class list and a frequency-domain audio representation is created to 3) draw a filtering shape to be applied to a reverberated audio input signal. While the choice of digital reverberation is critical to the sounding result of MusikVerb, the model can incorporate any algorithm of this class, while preserving its main characteristics.

3. PERCEPTUAL PITCH CLASS RANKING

We adopt the Tonal Interval Space [10] in MusikVerb to compute the perceptual distance between two given sonorities driven from both symbolic music representation and musical audio. Ultimately, these perceptual distances support the creation of a ranked list of pitch classes from an audio input signal. The choice of such a perceptually-guided space over other related tonal pitch spaces (e.g., Spiral Array [11] and Tonal Pitch Space [12]) is due to its possibility: i) to process both symbolic music representations and audio input signals without the need for a error-prone audio-to-score transcription; ii) to represent the most common pitch levels, i.e., pitch, chord, and key, in a single space; and iii) to efficiently compute the perceptual distance between tonal pitch.

The Tonal Interval Space uses the fast Fourier transform to convert a given sonority, represented as the L_1 normalized Harmonic Pitch Class Profile (HPCP) vector [13], $c(n)$, expressing the energy of the 12 pitch classes, into a Tonal Interval Vector (TIV), $T(k)$, expressing musical interval periodicities, such that:

$$T(k) = w_a(k) \sum_{n=0}^{N-1} \bar{c}(n) e^{-\frac{j2\pi kn}{N}}, \quad k \in \mathbb{Z}, \quad (1)$$

where $N = 12$ is the dimension of the chroma vector. $w_a(k) = \{3, 8, 11.5, 11.5, 15, 14.5, 7.5\}$ are weights derived from empirical ratings of dyads consonance used to adjust the contribution of each interval, k , thus making the space perceptually relevant [14]. We set k to $1 \leq k \leq 6$ for $T(k)$ since the remaining coefficients are symmetric. $T(k)$ uses $\bar{c}(n)$ which is $c(n)$ normalized by the DC component $T(0) = \sum_{n=0}^{N-1} c(n)$ to allow the representation and comparison of music at different hierarchical levels of tonal pitch [10].

The resulting spatial location of TIVs, $T(k)$, ensures that tonal pitch understood as perceptually related within the Western music context correspond to small Euclidean distances. For example, at the pitch class level, it places intervals that play an important role in the tonal system (e.g., octaves, fifths, and thirds) at smaller distances. At the key level, the Tonal Interval Space represents our expectancy of proximity between the 24 major and minor keys by placing the dominant, subdominant and their relative minor keys at close distances as well as the diatonic pitch class and chord sets of a particular key in its neighborhood [10]. Mathematically, the Euclidean distance between two given TIVs, $T_i(k)$ and $T_j(k)$, is given by:

$$P_{i,j} = \sqrt{\sum_{k=1}^M |T_i(k) - T_j(k)|^2}, \quad (2)$$

where $M = 12$ is the dimension of a TIV, $T(k)$.

By interpreting $T_i(k)$ and $T_j(k)$ in Eq. (2) as an audio input TIV and a pitch class TIV, respectively, and repeating the operation for the 12 pitch classes (i.e., 0-11), we compute the distances of an input TIV from the 12 pitch classes, which we then concatenate into a single list. Finally, the list values are reordered by increasing distance and a list with ranked pitch class indexes is created. Fig. 2 shows the various steps involved in the creation of a ranked list of pitch classes from an audio input TIV of the C major chord (i.e., the pitch class set {0,4,7}).

To control the output rate of the ranked pitch class vectors, we compute mean values per TIV bin from a user-defined number of $W_s = 4096$ sample window TIVs with 50% overlap. This adaptation parameter, A , is further detailed in Section 7 and has been

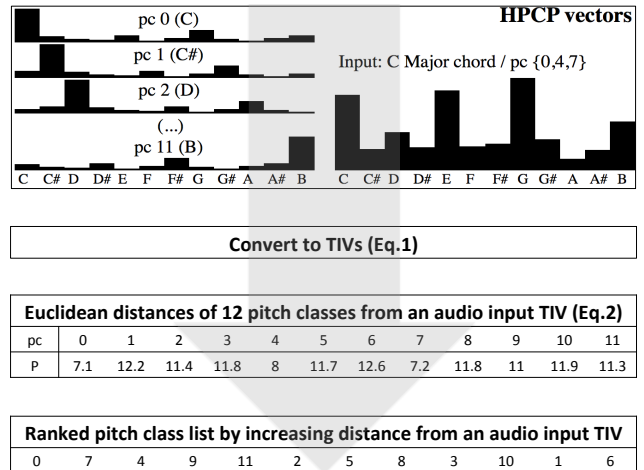


Figure 2: Illustration of the main algorithmic steps involved in the creation of a ranked list of pitch class distances from an audio input TIV of the C major chord.

shown to have a critical importance in the applicability scenarios of MusikVerb by expert musicians.

4. MAPPINGS

The mappings module is responsible for translating the ranked pitch class distance list into a spectral representation, which is then used to control the amplitude of frequency bins in a spectral filtering algorithm.

From the 12-element ranked list of pitch classes, a set of N_{pc} user-defined pitch classes are retrieved sequentially from the first element. N_{pc} is an integer value ranging from $N_{pc} = 1$, the first element of the list, to $N_{pc} = 12$, the entire list. The greater the N_{pc} value, the more perceptually distant notes to the input audio signal are introduced. The trimmed pitch class list, $m[k]$, is then mapped to an array of $0.5 \cdot W_s$ elements, representing the entire pitch range given by Eq.(3), where f_{ref} is the tuning reference (e.g $f_{ref} = 440Hz$)

$$x[k] = f_{ref} \cdot 2^{\frac{m[k]}{12}}, \quad 0 \leq k < N_{pc}, \quad (3)$$

where $x[k]$ is a vector containing the frequency corresponding to the first octave of the notes that should be on the output. For each pitch class in Eq. (3), a user-defined number of harmonics, N_h , is added, to regulate the harmonic richness of the re-synthesized signal. We empirically defined the number of harmonics N_h to be an integer value between 1 and 20, which we compute as:

$$y_k[n] = \prod_{1 \leq n < N_h, 0 \leq k < N_{pc}} n \cdot x[k], \quad (4)$$

After obtaining the vectors y_k , containing the frequencies that correspond to the selected N_{pc} and N_h we map them to elements of the $0.5 \cdot W_s$ window-sized filtering shape, H_f , using Eq. (5) where f_{res} corresponds to the FFT frequency resolution.

$$H_f[p] = 1, \quad p = \frac{y_k[n]}{f_{res}} \quad (5)$$

5. SPECTRAL FILTERING

MusikVerb resynthesises the input signal processed by a digital reverberation using a spectral filtering algorithm, similar to the one of the phase vocoder [15]. By multiplying the equal-sized frequency-domain representations of both the reverberated signal and the spectral filter shape resulting from Eq. 4, we then regulate the amplitude of each frequency bin.

6. USER CONTROL

MusikVerb has a dual implementation as a guitar pedal and a standalone software application. The Pure Data [16] software environment was initially adopted to prototype the effect due to its flexibility in running as a standalone application, a VST plugin [17] and in embedded DSP systems, such as the low-latency audio processing BELA¹ [18].

Both hardware (guitar pedal) and software (standalone application) instantiations of MusikVerb have two main groups of control parameters. The first group includes the digital reverberation parameters, such as room size, reverberation time, and spread, to cite a few. These parameters depend on the adopted digital reverberation algorithm, and thus can change accordingly. The digital reverberation adopted in the current version of our system includes several well-known digital reverberations implemented in Pure Data by Tom Erbe [19].

The second group includes the control parameters specific to MusikVerb: adaptation, harmonicity, and richness. Adaptation regulates the rate at which the ranked list of pitch classes is computed, which the user can control using a potentiometer in the guitar pedal and a slider in the software application (see Fig. 3). The harmonicity and richness parameters regulate the number of (ranked) pitch classes which are present in the output reverberated signal and the number of harmonics assigned to each note, respectively. These two latter parameters are controlled simultaneously with a single control in both hardware and software implementation of MusikVerb. In the hardware implementation, an expression pedal is scaled logarithmically to both parameters simultaneously. The choice of a logarithmic scale allows a finer degree of control over the initial range of the scale, where the effect more significantly alters a traditional digital reverberation. In the software implementation, the control of these two parameters are done via a 2-dimensional panel, whose x and y axis are assigned to each parameter (see Fig. 3).

7. APPLICATION

We have conducted several informal sessions with expert guitarists acquainted with different musical styles to infer recurrent applicability scenarios of MusikVerb and their creative potential. Three typical parameter combinations have caught the attention of the participants. These three parameter combinations explore MusikVerb in a wide range of creative applicability scenarios from a clutter-free reverberation with control over the reverberation harmonic quality to effects which are rather situated in the accompaniment systems domain.

The first two cases adopt low degrees of harmonicity and (harmonic) richness (e.g., $N_{pc} = 3$ and $N_h = 5$) and focus on the manipulation of the adaptation and reverberation time parameters.

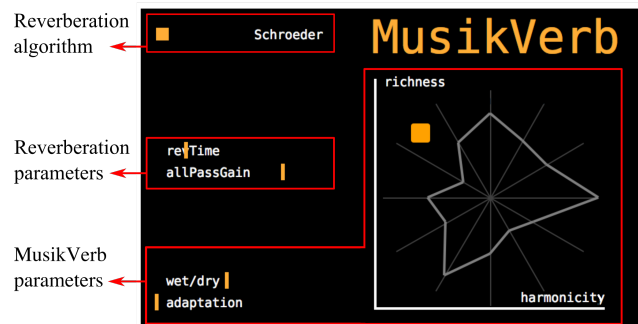


Figure 3: MusikVerb software application interface.

Adopting a low adaptation (e.g., $A = 6$) and a reverberation time typical of concert venues (e.g., around two seconds of decay time), MusikVerb significantly reduces the typical clutter of traditional reverberations, which result from the superposition of inharmonic frequencies around the frequency range of the source (as shown in Fig. 4). While this parametrization mode preserves most attributes of a reverberation without obscuring the source, it does not model the acoustic reflections of a room, as such an harmonically-tuned space does not exist.

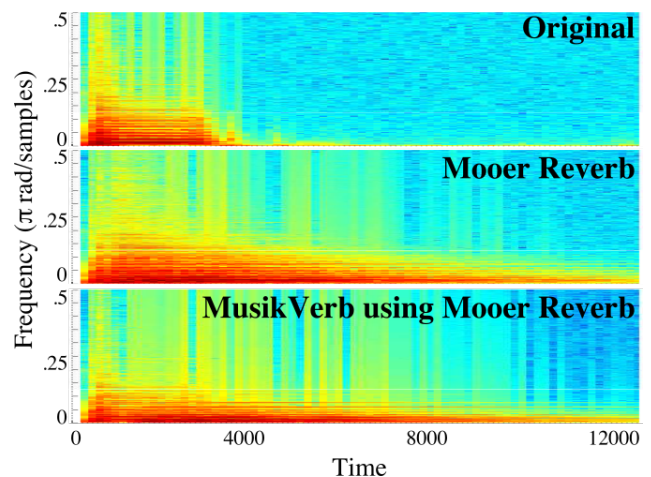


Figure 4: Three sonogram representations of an (original) audio soundfile (top), and two processed renditions of the soundfile after being processed by Mooer reverberation (middle) and MusikVerb using the Mooer reverberation (bottom).

The second case retains the low degrees of harmonicity and richness and opposes the first scenario by adopting high adaptation and reverberation time values (e.g., $A = 15$ and reverberation times around 5-10 seconds of decay time). This parameter combination creates an accompaniment close to drones or pedal tones which are predominant in the harmonic context of large sections of the input signal. Harmonicity in the context of this parameter combination can alter the density of pitch classes in the accompaniment which can range from a monophonic pedal tone to chords changes over time with variable number of notes. High adaptation

¹<https://bela.io/>

values impose a certain shift in time between the input signal and the (filtered) reverberation response to a level which no physical space can create or its digital reverberation models. This scenario provides ambient artists, film composers and sound designers with exciting new creative options for making evolving drones, organic pads, lush ambient and soundscapes.

Finally, the third parameter combination fixes the adaptation and reverberation time to average values across their range (e.g., $A = 10$ and a 1 second reverberation tail) and explore the dynamic manipulation of the linked harmonic richness parameters across the musical time. In manipulating these linked dimensions via the guitar pedal, for example, we can change the harmonic quality of the reverberation output in real-time in light of the harmonic content of the input. Manipulating the degree of harmonic proximity to the input signal, has a clear perceptual correlate with consonance (lower values) and dissonance (higher values), which can be dynamically manipulated irrespective of the performance audio content, thus promoting new strategies for creation.

The MusikVerb application, some sound examples demonstrating the three aforementioned applicability scenarios, and a demonstration video of a session with a guitarist performing with MusikVerb can be found online at: <https://bit.ly/2Jw30oP>.

8. CONCLUSIONS AND FUTURE WORK

We presented MusikVerb, a system which promotes a novel adaptive reverberation audio effect, which results from technical and artistic contributions. The system is effective in reducing the sonic clutter, commonly introduced by traditional reverberation effects, while promoting the exploration of new creative spaces, notably those close to an automatic accompaniment system, by leveraging a constant symbiosis between engineering and creativity. MusikVerb was developed as an embedded guitar pedal system using the BELA platform and as a software standalone application in the Pure Data programming language.

To further extend MusikVerb, it would be interesting to adapt it and test it with different input sources, either instruments, ambient sounds or any other sonic input. Adapting the weights, $w_a(k)$ of the Tonal Interval Space, to privilege intervals other than octaves, fifths and thirds, can extend the creative potential of the tool beyond the perceptually-inspired syntax of the Western tonal harmony. Finally, we aim to compare our system with Zynaptiq's Adaptiverb [9] to unveil their sonic and usability differences.

9. ACKNOWLEDGMENTS

This work is supported by national funds through the FCT - Foundation for Science and Technology, I.P., under the project IF/01566/2015.

10. REFERENCES

- [1] V. Verfaillie and D. Arfib, "A-dafx: Adaptive digital audio effects," in *Proceedings of COST G-6 Conference on Digital Audio Effects*, 2001.
- [2] O. Campbell, C. Roads, A. Cabrera, M. Wright, and Y. Visell, "Adept: A framework for adaptive digital audio effects," in *2nd AES Workshop on Intelligent Music Production (WIMP)*, 2016.
- [3] J. Holfelt, G. Csapo, N. Andersson, S. Zabetian, M. Castenieto, S. Dahl, D. Overholt, and C. Erkut, "Extraction, mapping, and evaluation of expressive acoustic features for adaptive digital audio effects," in *Proceedings of the Sound & Music Computing Conference*, 2017.
- [4] U. Zölzer, *DAFX: Digital Audio Effects*, John Wiley & Sons, Ltd, Sussex, UK, second edition, 2011.
- [5] V. Verfaillie, M. Wanderley, and P. Depalle, "Mapping strategies for gestural and adaptive control of digital audio effects," *Journal of New Music Research*, vol. 35, no. 1, pp. 71–93, 2006.
- [6] B. Corporation, "Te-2-tera echo," <https://www.boss.info/global/products/te-2/>, accessed April 9, 2018.
- [7] B. Corporation, "Da-2-adaptive distortion," <https://www.boss.info/global/products/da-2/>, accessed April 9, 2018.
- [8] B. Corporation, "Mo-2-multi overtone," <https://www.boss.info/global/products/mo-2/>, accessed April 9, 2018.
- [9] Zynaptiq, "Adaptiverb," <http://www.zynaptiq.com/adaptiverb/>, accessed March 30, 2018.
- [10] G. Bernardes, D. Cocharro, M. Caetano, C. Guedes, and M. E. P. Davies, "A multi-level tonal interval space for modelling pitch relatedness and musical consonance," *Journal of New Music Research*, vol. 45, no. 4, pp. 281–294, 2016.
- [11] E. Chew, "The spiral array: An algorithm for determining key boundaries," in *Music and artificial intelligence*, pp. 18–31. Springer, 2002.
- [12] F. Lerdahl, "Tonal pitch space," *Music Perception: An Interdisciplinary Journal*, vol. 5, no. 3, pp. 315–349, 1988.
- [13] E. Gómez, *Tonal Description of Music Audio Signals*, Ph.D. thesis, Universitat Pompeu Fabra, 2006.
- [14] G. Bernardes, M. E.P. Davies, and C. Guedes, "A perceptually-motivated harmonic compatibility method for music mixing," in *Proceedings of the CMMR conference*, 2017, pp. 104–115.
- [15] M. Dolson, "The phase vocoder: A tutorial," *Computer Music Journal*, vol. 10, no. 4, pp. 14–27, 1986.
- [16] M. Puckette, "Pure data: another integrated computer music environment," in *Proceedings of the second intercollege computer music concerts*, 1996, pp. 37–41.
- [17] Enzien Audio Ltd., "Heavy audio tools," <https://enzienaudio.com>, accessed March 28, 2018.
- [18] P. Brinkmann, P. Kirn, R. Lawler, C. McCormick, M. Roth, and H.-C. Steiner, "Embedding pure data with libpd," in *Proceedings of the Pure Data Convention*, 2011, pp. 291–.
- [19] T. Erbe, *Building the Erbe-Verb: Extending the Feedback Delay Network Reverb for Modular Synthesizer Use*, Ann Arbor, MI: Michigan Publishing, University of Michigan Library, 2015.
- [20] J. Sterne, "pace within space: Artificial reverb and the detachable echo," *Grey Room*, vol. 60, pp. 110–131, 2015.
- [21] X. Serra, *Musical Signal Processing*, chapter Musical Sound Modeling with Sinusoids plus Noise, pp. 91–122, G. D. Poli and A. Piccilli and S. T. Pope and C. Roads, Lisse, Switzerland, 1996.