Extraction, Mapping, and Evaluation of Expressive Acoustic Features for Adaptive Digital Audio Effects

Jonas Holfelt, Gergely Csapo, Nikolaj Andersson, Sohejl Zabetian,
Michael Castanieto, Daniel Overholt, Sofia Dahl, Cumhur Erkut
Aalborg University Copenhagen
{jholfe13,gcsapo13,na13,szabet16,mcasta16}@student.aau.dk
{dano,sof,cer}@create.aau.dk

ABSTRACT

This paper describes the design and implementation of a real-time adaptive digital audio effect with an emphasis on using expressive audio features that control effect parameters. Research in adaptive digital audio effects is covered along with studies about expressivity and important perceptual sound descriptors for communicating emotions. This project was aiming to exploit sounds as expressive indicators to create novel sound transformations. A test was conducted to see if guitar players could differentiate between an adaptive and non-adaptive version of a digital audio effect. The participants could hear a difference, especially when performing expressively. Although the adaptive effect did not seem to enhance expressive capabilities, participants did report an increased sense of control and general awareness of the effect. Overall, the preference over the two versions varied evenly between participants.

1. INTRODUCTION

Throughout history, advances in technology have played a huge role in the development and progression of music [1]. Digital audio is but another advancement built upon foundations established by the early pioneers of digital processing [2]. Today, the progression of music technology is as relevant as ever, and digital audio is at the heart of a wide range of applications, from music production and performance, to the very way in which music is consumed.

'Adaptive' effects automatically modify control parameters through mappings based on the analysis and extraction of audio features derived from the audio signal [3]. Traditionally, adaptive effects tend to be based on the dynamic properties of sound (examples of this include the compressor and noise gate) [3]. In this work, however, we will refer to adaptive effects as effects based on spectral properties of the input signal, which have a direct correlation to the notion of musical expression [4]. While adaptive effects can be applied and used for many different purposes, this paper will focus on its use with electric guitar playing.

Expression plays a crucial role in musical performances [5]. Musicians tend to communicate emotions and messages through expressive play, which often involves expressive improvisation and deviations from notated music [6, 7].The expressive intention of a performer can be investigated through the analysis of acoustic tone parameters, as these parameters carry sufficient information to identify some of the underlying emotions that the musician is trying to convey [5].

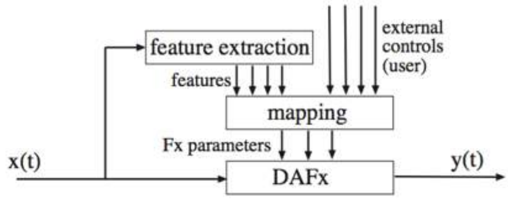

Figure 1. *"Structure of an adaptive DAFx (A-DAFx), with the input signal and the output signal. Features are extracted, and then a mapping is done between these features and user controls as input, and effect parameters as output"*. (Verfaille and Arfib, 2001 [4]).

Temporal variations in timing, dynamics and timbre, all contribute to the expressive quality of musical performances. Such variations in the acoustic signal have been previously considered in the context of modelling expressive performances for analysis [5, 6, 8] and synthesis [3, 7, 9, 10]. However, so far they have not been considered in conjunction with expressive intentions to control adaptive audio effects. An interesting approach would be to exploit such intentions with the purpose of interacting with audio effects in musical performances.

The novelty of this approach has created an opportunity to implement and evaluate a system which extracts expressive acoustical parameters, and maps them to relevant effect control parameters. A multi-effect delay engine was chosen, as it provides scope to clearly test and interpret parameter mapping relationships. In order to create this system, the general structure for adaptive effect implementation was used with the following steps (see Fig 1) [4]:

1. Analysis and feature extraction.
2. Mapping between features and effect parameters environment.
3. Transformation and Re-synthesis

2. RELATED WORK

Research on digital audio effects is highly interdisciplinary in the sense that a wide variety of topics and technical principles have to be understood for their design and implementation. This section will take a methodical approach in considering the various research paths explored during the process of this project.

2.1 Adaptive Digital Audio Effects

Verfaille et al. proposed a new type of audio transformation: 'adaptive digital audio effects' (A-DAFx). This type of effect is also referred to as dynamic processing, intelligent effects, or content based transformations since the effect parameters are altered by the input source over time [4]. The idea of altering the output sound based on the input already exists in commonly used effects such as compressors and auto-tune. However, adaptiveness introduces a new kind of control since specific sound features (or sound descriptors) can be mapped to alter control parameters of the output. The implemented echo effect uses the auto-adaptive approach, which is defined as an effect where acoustic features are extracted from the same source as to which the effect is applied.

The mapping of the sound features is an important aspect of A-DAFx. Several different mapping strategies have been suggested for adaptive effects, covering both sound and gesture inputs [3, 4, 11, 12]. However, only sound features were considered for the purpose of this project. An important consideration for mapping of sound features is the mapping strategy between extraction and manipulation parameters. This choice can be difficult to argue for, since it depends on musical intention [4]. If the mapping gets too complex, the users ability to manipulate the mapping depends on their prior knowledge about DAFx and the techniques used.

2.2 Examples of A-DAFx

At this time, only three commercial adaptive products are known to be aimed at guitarists, these are the 'TE-2 Tera Echo', 'MO-2 Multi Overtone' and 'DA-2 Adaptive Distortion' by 'BOSS' [13–15]. However, no technical descriptions of these effects are known to be available.

2.3 Analysis and Extraction of Expressive Acoustic Features

Adaptive effects are classified based on perceptual attributes, the main perceptual axes in sound being: timing, pitch, timbre, loudness and space, with the latter being the least significant from a musical performance perspective [3]. The first four perceptual attributes will now be examined in the context of expressivity.

The feature extraction portion of the project was to be used for identifying musical expression in the acoustic spectrum of the input signal, and required general insight into various music processing techniques and algorithms [3,11,16]. Acoustic tone parameters constitute of information outlining not only musical structure, but also the expressive intentions of the performer. Friberg et. al stated that *'emotional expression of a music performance can be accurately predicted using a limited set of acoustic tone parameters'* [5].

Studies have found that variation of timing and amplitude are essential features of auditory communication and music, as they carry information about expressive intentions and provide higher resolution of auditory information [6]. In addition to the variety of existing software for extracting tone parameters (e.g. audio to MIDI converters), Friberg et. al developed the algorithm CUEX [5]to identify expressive variations in music performances, rather than providing accurate musical notations. The CUEX algorithm is able to extract a wide variety of parameters, but places significant emphasis on onset and offset detection, as it contributes to many temporal feature extractions, such as tone rate, articulation and onset velocity [5]. Hence, we chose tone rate to provide a rough tempo estimation, and fluctuations in the tone rate are used to identify temporal expressive intentions for our A-DAFx.

Timbral deviations also play an important role in enhancing the expressive value of a performance, and have a significant effect on the interpretation of music. The brightness attribute for example, has a direct correlation to the input force or energy of the sound. As a result, musicians are able to utilise this parameter for control and expression, which demonstrates a relation between acoustical timbre and expressive intention [7].

Loudness, another expressive perceptual attribute, which is related to the energy of a signal. Although amplitude variations do not influence expressivity to the same extent as temporal variations, they still have a significant effect, and as the energy of a signal can be intentionally controlled, it serves as a powerful tool for musical expression. Contrary to loudness, energy is not influenced by frequency, but this low level feature is able to provide sufficient information about the dynamics of the performance [6].

A guitar, as a string instrument produces nearly harmonic sounds, and its pitch can be determined as a result [1]. Harmonies and melodies carry important information on musical structure [3]. Yet musical structure is often established by a composer, therefore variations in pitch are not necessarily related to expressive performance. A focus on improvisational performance however, would allow the investigation of temporal changes in pitch as an expressive detail. The physical counterpart of pitch is the fundamental frequency, which can serve as an adequate indicator.

3. DESIGN

Prior work provided us with an overall set of requirements on developing an A-DAFx, and algorithms for extracting expressive acoustic features.

We found the field of A-DAFx to have an emphasis on implementation of algorithms and mapping strategies, while comparably few user-oriented studies have been done in the field, which highlights a potential area of focus. DAFx are commonly used by electric guitar players, making them an ideal main target group for the purpose of this project. This led to an investigation on how to extract information and expressive features from the electric guitar and how to map these features to the A-DAFx parameters. A set of design criteria were formed to frame and guide the project with a special emphasis on implementing A-DAFx that would be beneficial for electric guitar players.

Design Requirements:

- It should function in most Digital Audio Workstations (DAWs) as a VST format.
- The plugin should work in real-time without notable latency.
- The extracted features should be generalised and mapped to work with several different effects.

For the purpose of creating an A-DAFx, a digital implementation of the spatial 'echo' effect was chosen. To create a more interesting echo effect, a variety of modulation and timbral effects were implemented into the echo engine. These included highpass and lowpass filters, vibrato, reverse delay, and saturation (tube). The choice of effects was influenced by commercial delay plugins that utilise similar effects, for example: Strymon's *'Timeline'*, ChaseBliss Audio's *'Tonal recall'*, and Red Panda's *'Particle Granular Delay/Pitch Shifter'* [17].

During the design phase, the DAFx was tested and tweaked using MATLAB by MathWorks and its Audio System Toolbox Library. The white paper by Mathsworks' DeVane and Bunkhelia served as inspiration on how to work with the library [18].

4. IMPLEMENTATION

4.1 Real-Time DAFx

Since the adaptive effect should work in real-time, some modifications had to be made in order for the application to work. Real-time audio processing follows a frame-based pipeline where the signal is segmented into buffers for processing and output [19]. The frame size (also known as buffer size) and sampling rate is unknown, since these parameters can be changed by the user. To prevent crashing due to using an input size or sampling rate that is too large or small, the program needs to adjust itself and work with variable sizes. Furthermore, efficiency needs to be addressed since the user should not experience any notable latency during playing. If the effect and feature extraction algorithms require too much computational power, the input will output to the DAC too late, causing overrun and undesired output. Although it introduces latency, this problem can be fixed by using a larger frame size. Efficiency however was not a primary concern during the project implementation.

4.2 Delay Based Effects

4.2.1 Saturation

The saturation algorithm is based on Zölzer's implementation of tube distortion [4, pp. 122-123]. This algorithm replicates the nonlinearities introduced in valve amplifiers by asymmetrical clipping of the input audio signal x.

$$f(x) = \begin{cases} \frac{x-Q}{1-e^{-dist(x-Q))}} + \frac{Q}{1-e^{dist \cdot Q'}}, & Q \neq 0, x \neq Q, \\ \frac{1}{dist} + \frac{Q}{1-e^{dist \cdot Q'}} & x = Q \end{cases} \tag{1}$$

,

At low input levels there should be no distortion in the signal. Clipping and limiting should occur at large negative input values and be approximately linear for positive values [3]. The work point Q, controls how much nonlinear clipping is introduced in the signal. Large negative values for Q will provide a more linear output, and dist increases the amount of distortion in the signal.

4.2.2 Vibrato

The implementation for the vibrato was inspired by Zölzer's et al. vibrato algorithm [3]. This algorithm produces vibrato in an audio signal by creating pitch variations in a way that is similar to the Doppler effect. Varying the distance between the sound source and the listener is the same as varying the delay in the signal. If the variation in the delay is periodic, then the pitch variation is also periodic.

To implement this, a circular buffer is used with a read and write index. The write index is used for writing the audio input in the buffer. The read index reads the delayed audio input from the buffer in an oscillating manner creating periodic variations in the outputted delay.

4.2.3 Reverse Delay

The inspiration for this type of effect stemmed mainly from commercial pedals stated previously. The effect was achieved by saving the input in a delay line, reversing it and then reading from it. A variable was stored to keep track of the delay time in samples with the purpose of using and reversing the necessary buffer length. Pointer variables, a write index and a read index, were used to keep track of where the input was stored in the delay line, and what was required to be sent to the audio output.

4.2.4 Filters

One very common design for a low-pass filter, which can be easily implemented, is the biquad filter. This filter is a second-order linear IIR filter with two poles and two zeroes. One big advantage of using a second-order filter over a high-order IIR filter is greater stability [20]. The filter can be controlled with three user-defined parameters: the sampling frequency f_s, the corner frequency f_0, and the quality factor Q [21].

4.3 Expressive Feature Extraction

4.3.1 Pitch Detection:

The implemented pitch detection algorithm was inspired by the 'Harmonic Product Spectrum (HPS)' technique and is used for extracting the fundamental frequency for any given audio input signal in real-time. The main reason for choosing this method was due to its ability to run in real-time and perform well under a wide range of conditions [22]. An audio signal spectrum will usually consist of a series of peaks, corresponding to the fundamental frequency with harmonic components at integer multiples of the fundamental [23]. The HPS method utilises this by down-sampling the spectrum a number of times, multiplying the compressed versions together, and comparing it to the original spectrum. The resulting spectrum will display an alignment of the strongest harmonic peaks and consequently, the fundamental frequency. The HPS technique measures the maximum coincidence for harmonics for each spectral frame according to:

$$Y(\omega) = \prod_{r=1}^{R} |X(\omega r)| \tag{2}$$

Where R is the number of harmonics being considered, The resulting periodic correlation array is then searched for a maximum value of a range of possible fundamental frequencies to obtain the fundamental frequency estimate [22].

$$\widehat{Y} = arg\, max Y(\omega_i) \tag{3}$$

4.3.2 *Spectral Centroid:*

The spectral centroid provides a strong impression of the timbral feature; brightness. It is calculated by taking a spectrum's centre of gravity according to the following formula [24, 25].

$$SC_{hz} = \frac{\sum_{k=1}^{N-1} k * X^d[k]}{\sum_{k=1}^{N-1} X^d[k]} \tag{4}$$

The calculation was implemented by applying a hamming window for each buffer, to attenuate any unwanted frequencies that result from taking the Fourier Transform of a segment of the audio signal. For music signals, the brightness generally increases with the presence of higher frequency components in the signal, which is often introduced through inharmonic sounds [25, 26].

4.3.3 *Energy:*

Energy is a basic but powerful acoustic feature as it describes the signal energy [27]. Energy is calculated by squaring and summing up amplitudes in a windowed signal, after which the sum is normalised by the length of the window.

$$E_n = \frac{1}{N} \sum_{n=0}^{N-1} x^2(n) \tag{5}$$

The frame size of the input signal was used to decide the window length. The value generated by the energy was used to estimate the amount of force in the guitar strumming.

4.3.4 *Tone Rate:*

Tone rate was implemented to gain information on the performers speed of playing. It is roughly proportional to tempo, as it refers to the number of successive tones over a given period, but this range is not large enough to serve as a detailed indication of tempo. In order to gain a better resolution, we derive the average inter onset interval (IOI) to indicate speed at which the musician is playing. The onset marks a single point in time within the signal where the transient region or attack starts, highlighting a sudden burst of energy [28].

Onset detection based on spectral features include various robust methods, which generally require less preprocessing and potentially allow the bandwise analyses of polyphonic signals. The reduction function transforms the set of spectral windows into a so called novelty curve, which indicates the location of prominent changes, and therefore the possible onsets. The novelty curve is acquired from the rectified spectral difference at each window, treating the onset as a broadband event, capturing only positive changes in the magnitude spectrum [28, 29].

$$SD = \sum_{k=0}^{K} |X[k]| - |X[k-1]| \tag{6}$$

The localisation of an onset relies on two threshold parameters. First, an absolute threshold on the amount of positive change in the magnitude spectrum, to filter out the spurious peaks. The second being temporal threshold, which makes sure that only one onset is detected in a given range. In order to calculate the average inter-onset interval, the onset intervals are accumulated over time in a vector and divided by the total number of onsets in the same period.

4.4 Parameter Mapping

In order to demonstrate the capabilities of the effect we created a specific preset. The parameter mappings for the pre-

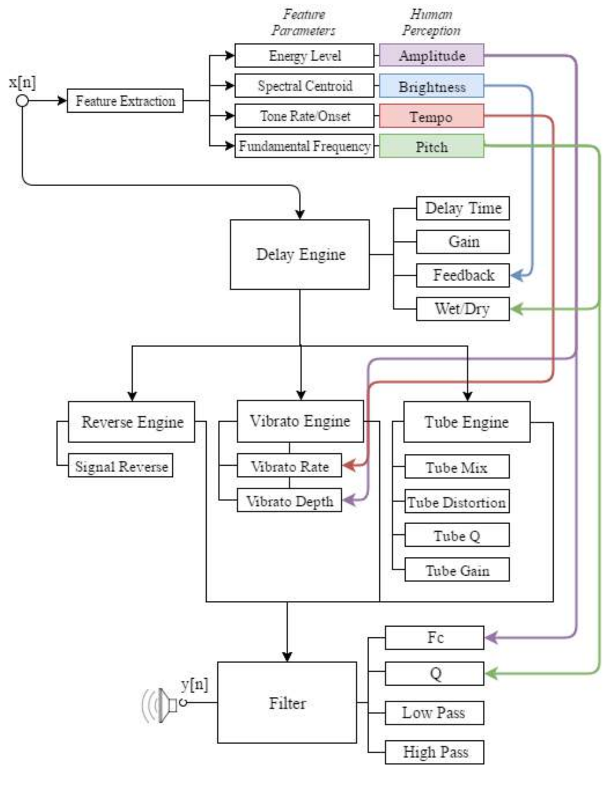

Figure 2. Effect engine and parameter mapping structure

set called 'Dreamy', were determined by an attempt to construct meaningful relationships between feature extractions and effect controls, and to also create an accurate representation of the preset name. The notion of dreamy music is often associated with rich ambient and reverberant sounds, which can be digitally replicated by utilising effects such as delay, reverb and vibrato. An attempt to achieve this type of 'sound identity' was made through the parameter relationships described. We found adaptive mappings such as 'Pitch Height → Delay Mix' to make sense musically, as overly delayed signals in low frequencies can create an undesired muddy low end. The mappings were decided from what we found aesthetically pleasing and fitting to the chosen notion of dreamy. Further, to inform the choice of adaptive mapping, the dreamy preset was subject to initial tests. The non-adaptive versions static parameters were centred in relation to the range of the adaptive mapping. Figure 2 displays the signal path of the whole effect, with the coloured lines outlining the parameter mappings for 'Dreamy'. The adaptive mappings for the 'Dreamy' preset are as follows:

- Energy Level → Vibrato Depth + Filter Fc
- Spectral Centroid → Delay Feedback
- Tone/Rate → Vibrato Rate
- Fundamental Frequency → Delay Mix + Filter Q

5. EVALUATION

Having designed and implemented an adaptive digital audio effect that uses expressive audio features as input, we now turn to evaluating its performance and use for a performer. In the following we will describe the evaluation which was done to aswer the research question: Are users able to distinguish between the adaptive and non-adaptive version, and do they understand what the difference is?

5.1 Participants

In order to take part in the evaluation, we required that the participant be able to perform basic guitar techniques. Additionally, we deemed knowledge about effects and various musical terms necessary for the participant to understand the questions asked during the experiment. Hence the target group was defined as intermediate guitar players well acquainted with audio effects.

The participants were recruited through purposive sampling [30]. The researchers chose associative participants, who fit the target group. A total of seven people participated in the evaluation. The participants were between 22 and 27 years old and the average guitar playing experience was 9.8 years.

5.2 Procedure

We chose a within-participant design where players compared two versions of our effect, one adaptive and one non-adaptive. In order not to affect the behaviour of the participants in a systematic way [31], the presentation order of the two versions was counterbalanced so that Participant 1 tested with A = non-adaptive, B = adaptive, while Participant 2 tested with A = adaptive, B = non-adaptive, etc.).

Initially a pilot test was performed, and the experimental procedure was adjusted accordingly. The final procedure was deducted as follows: Each participant was given a short scripted introduction. They were told they would be trying two versions of an effect, and were introduced to the plugin interface. Thereafter the participant would be asked to sign a terms of agreement, allowing for the experiment to be recorded. A small demographic questionnaire was also conducted to obtain information on age, experience, and preferred genre. This was done to make sure that participants fit the chosen target group.

The remaining test was structured as follows:

1. Basic tasks; 2. Short interview; 3. Expressive improvisational tasks; and 4. Follow-up interview.

In the first set of tasks, the participant was asked to perform a variety of basic playing styles on the guitar. The goal of these tasks was to make sure that the participant explored every aspect of the effect, and was able to gather some thoughts about the difference between the two versions. After performing the basic tasks, a short interview was conducted, where the participant was asked to explain the difference between the two versions. There were questions specifically asking about the effect of certain features (such as strum power or pitch), which helped clarify if the participant understood changes in the sound.

The next set of tasks focused on the participants own interpretation of expressiveness. They were asked to try version A and B when playing deadpan and expressively. The prediction was that the difference between A and B would be more noticeable when playing expressively. Deadpan was described as mechanical playing, adding no expressive variation to the playing. Expressively was defined as performing as though you were trying to convey some musical expression and emotion to an audience.

Finally, another short interview was conducted. The participant was asked to explain if the difference between version A and B was more noticeable when playing expressively. Afterwards the difference was revealed by the experimenter, and the participant was asked to guess which of the versions was the adaptive one. The complete test lasted between 25-35 minutes per participant.

5.3 Data Collection

1) Qualitative Interview: The main method for gathering data was through qualitative structured interviews: All of the questions required the participants to provide a specific answer (yes, no, A or B). This allowed for a structured comparison between all their answers. Additionally, the participants were encouraged to elaborate on each question, with the test experimenter taking notes of their answers throughout the interviews.

2) Observation: The tests were all video recorded for post-observation. The guitar was recorded with the effects applied. The recordings served several purposes:

1. Review the interviews, in case the test experimenter missed any important notes. 2. Check for user errors did they understand the tasks correctly? 3. Check for systematic errors glitches, and other types of unpredictable behaviour. 4. Triangulate with the data from the interviews.

Question asked	Choice Adapt.	Choice Non-Adap.	Unsure
Q1: Perceived noticeable difference?	Yes: **7**	No: 0	0
Q2: In what version did the energy level change the effect?	**6**	1	0
Q3: In what version did the tempo change the effect?	1	1	**5**
Q4: In what version did the pitch change the effect	**4**	1	2
Q5: Were you at any point confused?	Yes:3	No:**4**	0
Q6: In which version did you have more control?	**5**	2	0
Q7: More noticeable difference when playing expressively?	**6**	1	0
Q8: Which version made you motivated to play expressively?	3	3	1
Q9: Preferred version	**4**	3	0
Q10: Which was adaptive?	**5**	1	1

Table 1. Summary of the participant's choices to the simple questions during the evaluation. The questions have been condensed here, but can be found their actual format in the appendix. The data in 'bold' outlines the successful choices in the interview. Ad = the adaptive version. Non-Ad = the non-adaptive version. U = Unclear, Dont know or No perceived difference.

5.4 Analysis Methods

As a first step, participants' choices to the simple questions were summarised (See table 1), after which a more thorough analysis of the interview was performed using Meaning Condensation and Content Analysis. The full interview answers were reduced by performing meaning condensation, after which the content was analysed by searching for keywords, patterns, and outliers to gain a broad perspective of the answers. To strengthen the validity of the study, the data from the interviews was compared with the observation data. If the two sources are not congruent we considered the data to be invalid.

6. RESULTS

6.1 Difference Noticed Between Versions

1) General: All participants perceived a difference between the adaptive and the non-adaptive version. The non-adaptive version was described as a flat, dry and more simple compared to the adaptive. The adaptive version was described to have a thicker/fuller effect where there was more happening with the effect.
2) Effect Change According to Energy Level: A vast majority said that the power of their strum changed some parameter of the effect in adaptive version. Generally, some parameter of the effect changed according to the power of the strum. There were different interpretations about what happened, some of which were: sensitivity of an envelope filter, added harmonics, and presence.
3) Effect Change According to Tempo: There was a lot of confusion regarding the mapping of the tempo. Generally, the participants were unable to perceive a difference from their tempo fluctuations. The participants could not tell the difference between the adaptive version and the non-adaptive version.
4) Effect Change According to Pitch: There was some confusion regarding the mapping of the pitch. Most were able to hear a change of the effect in the adaptive version, however some could not perceive a difference between the two versions. A change in the adaptive version was identified, but the description of the effect was inconsistent between participants. They could hear that something happened but could not define exactly what it was.

6.2 Comparing 'Deadpan' to 'Expressive' Playing

6.2.1 Expressive Motivation:

The version that made the participants more motivated to play expressively appeared to be random, and the version that participants preferred in general was fairly random. There was no agreed preference. The non-adaptive version was deemed simpler and predictable, while the adaptive version was deemed playful, stronger, and interesting for experimenting with sounds.

6.2.2 More Noticeable Differences:

All, except for one participant, thought that the difference between the two versions was more noticeable when playing expressively compared to deadpan. The majority of participants guessed correctly when asked which the adaptive version was. Participants thought there was something adaptive happening in both versions when playing expressively, but found it more clear in the adaptive version. Comments include that the effect seemed more static in both versions when playing in a deadpan manner.

6.3 Usability

Some participants were confused about certain aspects of the effect, but most found it intuitive. Most participants felt that they had more control in the adaptive version, but it was difficult to determine exactly what happened with the sound. The adaptive version was deemed complex in the sense that there were more things going on, while the non-adaptive was deemed simpler in terms of control and understanding.

7. DISCUSSION

The results showed that the participants were able to perceive a notable difference between the two versions of the effect, and they were able to tell which one was adaptive. However, they had trouble understanding exactly what features actually affected the parameters of the effect. To gain an understanding of why some of these issues occurred, triangulation was performed.

7.1 General Comparison Between Versions

The participants described the non-adaptive version as flat, dry and more simple compared to the adaptive. The adaptive version was described to have a thicker/fuller effect where there was more happening with the effect. The influence of the energy was the most noticeable and a possible reason for this could be the direct relation between the power of the strum and the perceived changes in the vibrato depth and the cut off frequency. The influence of playing tempo and the pitch on the effect was less clear to the participants. The players could detect the influence of pitch, but were not able to identify what parameter it was controlling. The least successful mapping was the tempo and the vibrato rate, which the participants were perceiving the same in both cases. The noticed difference can also be biased by the imbalanced influence of different mappings, therefore the masking of one effect of another.

In the more free task involving expressive playing, players thought both the non-adaptive and the adaptive versions were adaptive. But the players described the non-adaptive version to be simpler and predictable, while the adaptive version was deemed playful, stronger, and interesting for experimenting with sounds.

7.2 Further Comments

The responses to the questions regarding general usability that participants answered at the end of the test can be helpful to set design requirements for further development. All the participants thought that an adaptive effect would be useful, but mostly for experimental purposes. They generally wished that they had more control, or at least knew what kind of parameters were active in the effect, which highlights a clear need for transparency in the further design of such tool. The participants especially liked the idea of being able to control the intensity of 'adaptive expression' applied to the effect. They also had some ideas for mapping the effect differently.

8. CONCLUSIONS

An adaptive digital audio effect was implemented to run real-time and was tested on intermediate guitar players. The effect was a multi-dimensional delay engine with added vibrato, reverse delay, saturation, and filters. Several sound features, pitch, brightness, tone rate and energy were extracted from the input signal and transformed the output signal by varying the effect parameters over time. The design aimed to exploit the musical intention of the players in order to manipulate effect parameters. Through testing it was found that the participant generally did not find that the adaptive version made them play more expressively than the non-adaptive, and that users did not prefer the adaptive effect. However, when using the adaptive version, the participants felt they had more control and noticed a bigger difference in the sound produced by the effect when playing expressively compared to deadpan. It was also found that the mapping of tempo and pitch were not apparent to the users, whereas the energy mapping was easily recognised. To investigate A-DAFx further, it would be worthwhile to test the effect in various environments (such as in a recording or live setting). In conjunction with this, it may be beneficial to apply a longitudinal testing approach so that participants could integrate the effect in their own creative and compositional work process.

Acknowledgments

The authors greatfully thank the guitarists involved in the evaluation.
Authors Holfelt, Csapo, Andersson, Zabetian, and Castanieto were responsible for the main work which was performed as a student project during the first semester of Sound and Music Computing, Aalborg University Copenhagen. Authors Erkut and Dahl supervised the project and participated in the writing together with author Overholt.

9. REFERENCES

[1] D. M. Campbell, C. A. Greated, A. Myers, and M. Campbell, *Musical Instruments: History, Technology, and Performance of Instruments of Western Music*. Oxford: Oxford University Press, 2004.

[2] J. Bruck, A. Grundy, and I. Joel, "An Audio Timeline," 2016, accessed: Dec. 15, 2016. [Online]. Available: http://www.aes.org/aeshc/docs/audio.history.timeline.html

[3] U. Zölzer, *DAFX: Digital Audio Effects*, 2nd ed. Oxford, England: Wiley, John & Sons, 2011.

[4] V. Verfaille and D. Arfib, "A-DAFX: Adaptive Digital Audio Effects," *Digital Audio Effects (DAFX-01)*, 2001.

[5] A. Friberg, E. Schoonderwaldt, and P. Juslin, "Cuex: An Algorithm for Automatic Extraction of Expressive Tone Parameters in Music Performance from Acoustic Signals," *Acta Acustica United With Acustica*, vol. 93, no. 3, 2007.

[6] A. Bhatara, A. K. Tirovolas, L. M. Duan, B. Levy, and D. J. Levitin, "Perception of Emotional Expression in Musical Performance," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 37, no. 3, 2011.

[7] M. Barthet, R. Kronland-Martinet, and S. Ystad, "Improving Musical Expressiveness by Time-Varying Brightness Shaping," *Computer Music Modelling and Retrieval*, vol. 4969, 2008.

[8] R. A. Kendall and E. C. Carterette, "The Communication of Musical Expression," *Music Perception: An Interdisciplinary Journal*, vol. 8, no. 2, 1990.

[9] L. Mion and G. D'Incà, "Analysis of Expression in Simple Musical Gestures to Enhance Audio in Interfaces," *Virtual Reality*, vol. 10, no. 1, 2006.

[10] C. Oshima, K. Nishimoto, Y. Miyagawa, and T. Shirosaki, "A Concept to Facilitate Musical Expression," *Proceedings of the 4th conference on Creativity & cognition*, 2002.

[11] U. Z. V. Verfaille and D. Arfib, "Adaptive Digital Audio Effects (a-DAFX): A New Class of Sound Transformations," *IEEE Transactions on Audio, Speech and Language Processing*, no. 5, 2006.

[12] V. Verfaille, M. 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, 2006.

[13] B. Corporation, "Te-2 — Tera Echo," 2016, accessed: Dec. 13, 2016. [Online]. Available: https://www.boss.info/us/products/te-2/

[14] ——, "Da-2 — Adaptive Distortion," 2016, accessed: Dec. 13, 2016. [Online]. Available: https://www.boss.info/us/products/da-2/features/

[15] ——, "Mo-2 — Multi Overtone," 2016, accessed: Dec. 13, 2016. [Online]. Available: https://www.boss.info/us/products/mo-2/

[16] M. Müller, *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications: 2015*. Switzerland: Springer International Publishing AG, 2015.

[17] G. Tanaka, "Reviews of the Best Guitar Pedals & Gear," 2013, accessed: Dec. 12, 2016. [Online]. Available: http://www.bestguitareffects.com/top-best-delay-guitar-effects-pedals-buyers-guide/

[18] C. DeVane and G. Bunkheila, "Automatically Generating VST Plugins from MATLAB Code," 2016, accessed: Dec. 12, 2016. [Online]. Available: http://www.aes.org/e-lib/browse.cfm?elib=18142

[19] MathWorks, "Audio I/O: Buffering, Latency, and Throughput," 2016, accessed: Dec. 12, 2016. [Online]. Available: https://se.mathworks.com/help/audio/gs/audio-io-buffering-latency-and-throughput.html

[20] T. Instruments, "Configure the Coefficients for Digital Biquad Filters in TLV320AIC3xxx family," 2010, accessed: Dec. 12, 2016. [Online]. Available: http://www.ti.com/lit/an/slaa447/slaa447.pdf

[21] R. Bristow-Johnson, "Cookbook Formulae for Audio EQ Biquad Filter Coefficients," 2016, accessed: Dec. 12, 2016. [Online]. Available: http://www.musicdsp.org/files/Audio-EQ-Cookbook.txt

[22] T. Smyth, "Music 270a: Signal Analysis," 2015, accessed: Dec. 12, 2016. [Online]. Available: http://musicweb.ucsd.edu/~trsmyth/analysis/analysis.pdf

[23] G. Middleton, "OpenStax CNX," 2003, accessed: Dec. 12, 2016. [Online]. Available: http://cnx.org/contents/i5AAkZCP@2/Pitch-Detection-Algorithms

[24] T. H. Park, *Introduction to Digital Signal Processing*, 1st ed. Singapore: World Scientific Publishing Co., 2010.

[25] E. Schubert, "Does Timbral Brightness Scale with Frequency and Spectral Centroid?" *Acta Acustica United With Acustica*, vol. 92, 2006.

[26] H. Järveläinen and M. Karjalainen, "Importance of Inharmonicity in the Acoustic Guitar," *Forum Acusticum*, vol. 2005, 2005.

[27] F. Eyben, *Real-time Speech and Music Classification by Large Audio Feature Space Extraction*. Springer, 2016.

[28] J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. B. Sandler, "A Tutorial on Onset Detection in Music Signals," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 5, 2005.

[29] M. Müller, D. Ellis, A. Klapuri, and G. Richard, "Signal Processing for Music Analysis," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 6, 2011.

[30] T. Bjørner, *Qualitative Methods for Consumer Research: The Value of the Qualitative Approach in Theory & Practice*. Denmark: Reitzel, Hans, 2015.

[31] A. Field and G. J. Hole, *How to Design and Report Experiments*. London: SAGE Publications, 2003.