Estimating Performance Parameters from Electric Guitar Recordings

Zulfadhli Mohamad

Submitted in partial fulfillment
of the requirements of the
Degree of Doctor of Philosophy

School of Electronic Engineering and Computer Science
Queen Mary University of London

March 31, 2018

I, Zulfadhli Mohamad, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third partys copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature: Zulfadhli Mohamad

Date: March 31, 2018

ABSTRACT

The main motivation of this thesis is to explore several techniques for estimating electric guitar synthesis parameters to replicate the sound of popular guitarists. Many famous guitar players are recognisable by their distinctive electric guitar tone, and guitar enthusiasts would like to play or obtain their favourite guitarist's sound on their own guitars.

This thesis starts by exploring the possibilities of replicating a target guitar sound, given an input guitar signal, using a digital filter. A preliminary step is taken where a technique is proposed to transform the sound of a pickup into another on the same electric guitar. A least squares estimator is used to obtain the coefficients of a finite impulse response (FIR) filter to transform the sound. The technique yields good results which are supported by a listening test and a spectral distance measure showing that up to 99% of the difference between input and target signals is reduced. The robustness of the filters towards changes in repetitions, plucking positions, dynamics and fret positions are also discussed. A small increase in error was observed for different repetitions; moderate errors arose when the plucking position and dynamic were varied; and there were large errors when the training and test data comprised different notes (fret positions).

Secondly, this thesis explored another possible way to replicate the sound of popular guitarists in order to overcome the limitations provided by the first approach. Instead of directly morphing one sound into another, replicating the sound with electric guitar synthesis provides flexibility that requires some parameters. Three approaches to estimate the pickup and plucking positions of an electric guitar are discussed in this thesis which are the Spectral Peaks (SP), Autocorrelation of Spectral Peaks (AC-SP) and Log-correlation of Spectral Peaks (LC-SP) methods. LC-SP produces the best results with faster computation, where the median absolute errors for pickup and plucking position estimates are 1.97 mm and 2.73 mm respectively using single pickup data and the errors increased slightly for mixed pickup data. LC-SP is also shown to be robust towards changes in plucking dynamics and fret positions, where the median absolute errors for pickup and plucking position estimates are less than 4 mm. The Polynomial Regression Spectral Flattening (PRSF) method is introduced to compensate the effects of guitar effects, amplifiers, loudspeakers and microphones. The accuracy of the estimates is then tested on several guitar signal chains, where the median absolute errors for pickup and plucking position estimates range from 2.04 mm to 7.83 mm and 2.98 mm to 27.81 mm respectively.

Acknowledgements

I would like to take this opportunity to thank my PhD supervisor, Simon Dixon for the constant guidance and giving meaningful advice over the years, without which this work would not have been possible. I would also like to thank Christopher Harte for supervising my work since the beginning of my PhD and for providing me with his Squier Stratocaster to do some experiments on it. I want to thank Josh Reiss and Andrew McPherson for all the valuable constructive feedbacks given during my PhD's Stage 0, 1 and 2, which helped me steer in the right direction. I enjoyed my time working with such amazing people at the Centre for Digital Music at Queen Mary University of London.

To my soul-mate, my lovely wife, Anisah Nur: I want to thank you for the encouragement and endless support. You are a constant reminder of how lucky I am. I would also like to thank my family, especially my parents who taught me the value of hard work and an education. Many thanks to all of my dear friends for being such fantastic people in my life.

I cannot thank all of you enough, those who have contributed to this thesis and supported me during this awesome journey. I am extremely grateful to each and everyone of you. *Ribuan terima kasih.*

Contents

1 Introduction **9**

1.1 Motivation 10

1.2 Research Goal 11

1.3 Thesis Structure 11

1.4 Contributions 13

1.5 Related Publications by the Author 13

2 Background **15**

2.1 Fundamentals of Electric Guitar Sounds 16

2.1.1 The Electric Guitar 17

2.1.2 Guitar Effects, Amplifiers, Cabinet Loudspeakers and Microphones . . 28

2.2 Retrieving Information from Guitar Recordings 34

2.2.1 Pitch Detection 34

2.2.2 Onset Detection 35

2.2.3 Plucking Point and Pickup Position Estimation 37

2.2.4 Inharmonicity, String Detection, Playing Techniques and Decay Rate . 39

2.3 Overview of Electric Guitar Synthesis 41

2.3.1 Physical Modelling of Electric Guitar 41

2.3.2 Modelling Guitar Amplifier and Effects 43

2.4 Summary 44

3 Digitally Moving an Electric Guitar Pickup **46**

3.1 Dataset I 49

3.1.1 Electric Guitar Under Test 49

3.1.2 Audio Samples 49

3.2 Estimating the FIR Filter Coefficients 50
3.3 Analysing the Filter with an Example 53
3.3.1 Timbral Similarity Measurement 53
3.3.2 Morphing a Pickup Sound 54
3.4 Listening Test Results 57
3.4.1 Suitable Filter Lengths 58
3.4.2 Morphing Three Pickup Positions 62
3.4.3 Filter Robustness 65
3.5 Numerical Results 75
3.5.1 Filter Robustness 75
3.5.2 Comparisons Between Variables 76
3.6 Summary 78

4 Electric Guitar Parameter Estimation: Pickup and Plucking Positions 80
4.1 Dataset II 81
4.1.1 Audio Samples 81
4.1.2 Variations in Plucking Events 82
4.2 Methods 83
4.2.1 Spectral Peaks Method (SP) 83
4.2.2 Autocorrelation of Spectral Peaks Method (AC-SP) 88
4.2.3 Log-correlation of Spectral Peaks Method (LC-SP) 95
4.2.4 Setting the Total Number of Harmonics 104
4.3 Just-Noticeable Difference 104
4.4 Results: Pickup and Plucking Position Estimates 106
4.4.1 Single Pickup Data: Comparison of SP, AC-SP and LC-SP 106
4.4.2 Single Pickup Data: Comparison of Existing Methods 110
4.4.3 Single Pickup Data: Effects of Onset Detection 111
4.4.4 Mixed Pickup Data: Comparison of AC-SP and LC-SP 112
4.4.5 Effects of Plucking Dynamics 114
4.4.6 Effects of Fret Positions 116
4.4.7 Test on Chords 118
4.5 Real-world Applications 120
4.5.1 Identification of Electric Guitar Model 122

4.5.2 Identification of Pickup Selection 123

5 Pickup and Plucking Position Estimation in Real World Settings 125

5.1 Polynomial Regression Spectral Flattening (PRSF) 126

5.2 Test on Various Guitar Signal Chains 127

5.2.1 Audio samples 127

5.2.2 Results 129

5.2.3 Comparing Spectral Flattening Methods 133

5.3 Test on Real World Recordings 133

5.3.1 Distinguishing Between Pickup and Pluck Estimates 135

5.3.2 Identification of Guitar and its Pickup Selection 140

5.4 Summary 144

6 Conclusions 146

6.1 Summary 147

6.1.1 Estimating Filter Coefficients from Electric Guitar Recordings 147

6.1.2 Estimating Playing Parameters from Electric Guitar Recordings . . . 148

6.2 Future Work 150

6.2.1 Replicating Electric Guitar Sounds from Audio Recordings 150

List of Figures

2.1 A typical electric guitar chain 18
2.2 Plucking point effect 19
2.3 Plucking width effect 20
2.4 Plucking dynamic effect 21
2.5 String's inharmonicity 22
2.6 Single and mixed pickup effects 23
2.7 Pickup width effect 27
2.8 Soft and hard clipping effects 30
2.9 Symmetrical and asymmetrical soft clipping effects 31
2.10 Pitch detection: autocorrelation and YIN methods 35
2.11 Onset detections 38

3.1 A diagram of the method that transforms a sound from a pickup into another 48
3.2 The modified Squier Stratocaster diagram 48
3.3 Magnitude spectra of neck and bridge pickup signals and their estimates . . . 55
3.4 Zoom in version of Fig. 3.3 56
3.5 Magnitude spectra of learnt filters and their ground truths 57
3.6 Listening test 58
3.7 Filter orders versus normalised distances: 6th string 60
3.8 Similarity ratings: filter lengths and 6th string 61
3.9 Filter orders versus normalised distances: 1st string 62
3.10 Similarity ratings: filter lengths and 1st string 63
3.11 Similarity ratings: morphing three pickups 66
3.12 Similarity ratings: filter robustness (repetitions) 69

3.13 Similarity ratings: filter robustness (plucking dynamics) 70
3.14 Similarity ratings: filter robustness (plucking positions) 72
3.15 Similarity ratings: filter robustness (fret positions) 73
3.16 Similarity ratings: fret positions improvements 74

4.1 Excerpts from slow-motion video of plucking events: single plucks 84
4.2 Excerpts from slow-motion video of plucking events: chords 85
4.3 Diagram for the Spectral Peaks (SP) method 86
4.4 An electric guitar tone 86
4.5 Fourier series coefficients of the first period of the tone 87
4.6 Diagram for the Autocorrelation of Spectral Peaks (AC-SP) method 88
4.7 Spectral peaks of the tone and its spectral slope found using linear regression 88
4.8 Spectral slope adjustments 90
4.9 Autocorrelations of the electric guitar model and observed data 91
4.10 Autocorrelations of in-phase and out-of-phase mixed pickup models 92
4.11 Autocorrelations of spectral peaks with various slope adjustments 94
4.12 Diagram for the Log-correlation of Spectral Peaks (LC-SP) method 95
4.13 Log-correlation of an electric guitar tone 96
4.14 Log-correlation of an electric guitar tone that is plucked near the pickup . . . 97
4.15 Mains hum from 5 pickup configuration 100
4.16 Energies of mains hum from Dataset II 101
4.17 Autocorrelation and log-correlation of a mixed pickup signal 102
4.18 Log-correlations of a mixed pickup signal with two plucking positions 103
4.19 Just-noticeable difference of plucking points 106
4.20 Pickup and plucking position errors: single pickup data 107
4.21 Pickup position estimates using SP, AC-SP and LC-SP 108
4.22 Comparison of plucking point errors with existing methods 111
4.23 Pickup and plucking position errors: mixed pickup data 113
4.24 Pickup and plucking position errors: effects of plucking dynamics 115
4.25 Pickup and plucking position errors: effects of fret positions 117
4.26 Pickup and plucking position errors: chords 119
4.27 Complete system of estimating the pickup & plucking positions 121
4.28 Target locations of a Fender Stratocaster and a Gibson Les Paul 122

4.29 Target locations of a Fender Stratocaster, Telecaster and Jaguar 123

5.1 Spectral peaks of a tone and its envelope found using polynomial regression . 127
5.2 Various electric guitar signal chains 128
5.3 Pickup and plucking position errors: effects of various guitar signal chains . . 130
5.4 Effects of signal clipping on the estimates 132
5.5 Comparing spectral flattening methods with various guitar signal chains . . . 134
5.6 Excerpt from a guitar signal played by Robby Krieger of The Doors 135
5.7 Love Me Two Times: pickup and plucking position estimates 136
5.8 Excerpt from a guitar signal played by George Harrison of The Beatles 137
5.9 Day Tripper: pickup and plucking position estimates 138
5.10 Diagram for guitar model and pickup selection identification 140
5.11 Love Me Two Times: median pickup position estimates 140
5.12 Day Tripper: median pickup position estimates 141
5.13 Day Tripper: median pickup position estimates and fitted line 141

List of Tables

3.1 Squier Stratocaster measurements 50
3.2 Errors for filters applied to an input/target pair with different variables . . . 77
3.3 Summary table for comparisons between each variable. 78

4.1 Pickup identification results 124

5.1 Love Me Two Times: distance from pickup position estimates to their targets 143
5.2 Day Tripper song: distance from pickup position estimates to their targets . . 143

Chapter 1

Introduction

1.1 Motivation

The electric guitar revolutionised Western popular music and was one of the most prominent instruments in most blues, jazz, rock and pop music for several decades. It provides a vast amount of tonal diversity which allows musicians to explore a wide range of sound for their creative and artistic expression.

A musician can alter the sound of the electric guitar just by switching the pickup selection or adjusting its tone and volume controls while guitar effects, amplifiers and loudspeakers also make significant contributions to the tone. Well-known guitarists often have their own playing style and a particular combination of guitar, amplifier and effects, both of which are key ingredients of their unique sound. This is what makes some of them instantly recognisable and captures the attention of listeners.

The traditional way of replicating the sound of their favourite guitarists is by purchasing the same set of guitar, amplifier and effects and finding the right settings for each of them. This could be highly expensive, as some popular guitarists use vintage electric guitars and amplifiers. Furthermore, some vintage guitars and amplifiers might be discontinued which could also make them very difficult to find in guitar shops nearby and in online stores. Of course, a user also needs to play like his or her favourite guitarist in order to sound exactly like them.

Digital replication of electric guitar, guitar amplifiers and guitar effects has recently grown rapidly in the research community and the music industry. Today's technology allows electric guitars, amplifiers and effects to be digitally emulated using signal processing techniques achieving close resemblance in sonic quality to their analogue counterparts. This allows guitar enthusiasts to copy the sound of their favourite musicians more easily than the traditional method of having to find and purchase expensive equipment.

This motivation spurs the idea of developing methods to replicate the sound of popular guitarists from their published recordings through various signal processing techniques. Principle research questions arise which are how to replicate the sound, what type of information can be extracted from the recording in order to replicate it and how to extract the information. In this thesis, the questions are addressed, where two concepts of replicating the sound are explored and methods to extract relevant information from electric guitar signals are presented.

1.2 Research Goal

This research is mainly driven by developing methods to replicate the sound of popular guitarists from their published recordings. Relevant information could exist in their recordings which can help achieve that ambition. The aim of this research is to develop methods to extract meaningful information from guitar recordings. This information could then be used to help users replicate the sound of their favourite guitarists.

One type of information that can be looked into is the differences between a user's guitar signal and a desired guitar signal. So, estimating filter coefficients that could transform the user's guitar into the desired guitar sound could hold the key to achieve the aim of the main motivation. In order to weigh in the advantages and disadvantages of using this concept, it is first explored by taking the sound of an arbitrarily selected pickup as input and another pickup on the same guitar as the target signal.

Other information that could be extracted from the guitar signal is its performance parameters such as the pickup and plucking positions. In this thesis, a method to estimate the pickup and plucking positions on an electric guitar is proposed and evaluated for various cases.

1.3 Thesis Structure

Chapter 1

The motivations and main goal of research in performance parameter estimation on electric guitar recordings are discussed. The contributions of this research are also highlighted in this chapter.

Chapter 2

This chapter presents existing literature which is relevant to this work. The chapter starts with discussing about the fundamentals of electric guitar sound such as how the sound is produced and the factors that alter the sound. Methods to retrieve information from guitar recordings based on past literature are also discussed in this chapter. Previous literature on synthesising electric guitar and digitally emulating guitar effects, amplifiers, loudspeakers and microphones are then discussed.

Chapter 3 (based on Mohamad et al. (2015))

This chapter presents a technique to transform the sound of an arbitrarily selected magnetic pickup into another on the same electric guitar. The method is evaluated using listening tests, spectral difference measurements, and the effects of fret, dynamics, plucking position and repetition on the accuracy are tested.

Chapter 4 (based on Mohamad et al. (2017a,b,c))

This chapter proposes three frequency domain approaches to estimate the pickup and plucking positions on an electric guitar. The three approaches that are discussed are the Spectral Peaks (SP), Autocorrelation of Spectral Peaks (AC-SP) and Log-correlation of Spectral Peaks (LC-SP) methods which are tested on single pickup data. The AC-SP and LC-SP methods are then tested on mixed pickup data, and used to test the effects of plucking dynamics and fret positions on the accuracy of the estimates.

Chapter 5 (part of it is based on Mohamad et al. (2017b))

This chapter discusses the effects of various guitar signal chains on the accuracy of the pickup and plucking position estimates in order to simulate real-world settings. A guitar signal chain consists of emulated guitar effects, an amplifier, a loudspeaker and a microphone which is applied to the direct input electric guitar tone. The modified LC-SP method is introduced to improve the accuracy of the estimates for electric guitar tones with audio effects. This chapter also presents examples of estimating the pickup and plucking positions on an electric guitar in two commercial recordings.

Chapter 6

The last chapter provides a conclusion of the thesis, and outlines the prospects of further research.

1.4 Contributions

The main contributions of this thesis are:

- Chapter 3: a novel technique is introduced to transform the sound of a selected magnetic pickup into another on the same guitar.
- Chapter 4: a new frequency domain approach is introduced to estimate the pickup and plucking positions on an electric guitar by minimising the difference between the spectral peaks of the tone and the electric guitar model.
- Chapter 4: an improved technique is proposed to estimate the pickup and plucking positions on an electric guitar by minimising the difference between the autocorrelations of the spectral peaks of the tone and the electric guitar model.
- Chapter 4: another novel technique is proposed to estimate the pickup and plucking locations based on the log-correlation of the electric guitar tone.
- Chapter 4 & 5: the Linear Regression Spectral Flattening (LRSF) and the Polynomial Regression Spectral Flattening (PRSF) methods are introduced to improve the pickup and plucking position estimation.

1.5 Related Publications by the Author

In all publications listed below, the author was the main contributor by collecting and analysing the data and developing and implementing the models. The second and third authors of the publications, Simon Dixon and Christopher Harte supervised and edited the papers.

Peer-reviewed conference papers

- Z. Mohamad, S. Dixon, and C. Harte. Digitally moving an electric guitar pickup. In *Proceedings of the International Conference on Digital Audio Effects (DAFx-15)*, pages 284–291, 2015.[1]

[1] basis for Chapter 3

- Z. Mohamad, S. Dixon, and C. Harte. Pickup position and plucking point estimation on an electric guitar. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP17)*, pages 651–655, 2017.[2]

- Z. Mohamad, S. Dixon, and C. Harte. Estimating pickup and plucking positions of guitar tones and chords with audio effects. In *Proceedings of the International Conference on Digital Audio Effects (DAFx-17)*, pages 420–426, 2017.[3]

Journal paper

- Z. Mohamad, S. Dixon, and C. Harte. Pickup position and plucking point estimation on an electric guitar via autocorrelation, in the *Journal of the Acoustical Society of America*, pp. 3530 – 3540, 2017.[2]

[2]basis for Chapter 4
[3]basis for Chapter 4 & 5

Chapter 2

Background

In the previous chapter, the main goal of this research is explained which is to extract meaningful information from electric guitar recordings in order to replicate the guitar sound. This chapter establishes the basic understanding of electric guitar sounds required in later discussions throughout the thesis.

In order to know the type of information that needs to be extracted, the inner workings of the electric guitar sound should be studied. Section 2.1 discusses how the electric guitar produces sound, from the physics of string instruments to the output of a loudspeaker, and the attributes that affect the timbre of an electric guitar.

Section 2.2 explains existing techniques that have been used to extract meaningful information from recordings of electric guitars and other related musical instruments. Techniques to extract relevant information such as the pitch, onset time, plucking point, pickup position, inharmonicity coefficient, string and fret played, playing technique and decay rate are discussed.

Information such as the electric guitar's performance parameters can be used as parameters for electric guitar synthesis. In Section 2.3, previous literatures about synthesising electric guitar sound are discussed.

2.1 Fundamentals of Electric Guitar Sounds

An electric guitar is a plucked string instrument that uses magnetic pickups to convert the vibrations of the strings to electrical signals. Much like an acoustic guitar, the vibrations of its strings produce sound but what makes it different from acoustic guitars is how the sound of the string is amplified. The vibrations of the string in the magnetic field of the pickup induce a weak electrical signal in the pickup, which is normally amplified by a guitar amplifier and sent to the loudspeaker to produce sound. Some musicians use effects such as distortion, overdrive, reverberation and chorus to alter the sound according to their taste for artistic purposes. Case et al. (2013) describe an electric guitar as a "sound synthesiser" capable of a vast range of sounds where different choices of electric guitar model, guitar amplifier, guitar effects, loudspeaker cabinet and microphone change the sound significantly. Not only do different types of equipment render a different sound, but also the settings of each component have an impact on the tonal colouration of the electric guitar timbre.

Fig. 2.1 shows a typical set of equipment used to record an electric guitar, which includes an electric guitar, effects, amplifier, loudspeaker cabinet and microphone. This example shows an illustration of a Fender Stratocaster – which is one of the most popular electric

guitars of all time – that have three magnetic pickups called the neck, middle and bridge pickups. It has a pickup selector which allows the player to choose one of the 3 single pickups or one of 2 mixed pickup configurations. The mixed pickup configuration allows two adjacent pickups to be selected at the same time. Simply changing the switch of the pickup selector can drastically alter the sound. There are also three knobs on the guitar which are two tone controls that basically acts as low pass filters and a volume control. Other electric guitar models such as a Gibson Les Paul have different physical properties that result in a different sound than that of a Fender Stratocaster. The effects of these properties will be discussed later in Section 2.1.1.

The electric guitar may be connected to a chain of guitar effects, and the output of the guitar effects chain is plugged into a guitar amplifier, which produces sound through the loudspeaker. Fig. 2.1 shows an illustration of a Marshall JCM800 which has an amplifier head and a loudspeaker cabinet connected together. A typical guitar amplifier has tone and volume controls which will further vary the sound of the electric guitar. Microphones are used to capture the sound for recordings, where the microphone selection and placement also affects the overall electric guitar tone (Case, 2010).

2.1.1 The Electric Guitar

Plucking Point

Suppose that a string of length L with density μ is fixed with tension ζ between two rigid supports. Taking the x-axis along the string and concerning with transverse vibrations in the $\tilde{y}$ and $\tilde{z}$ directions, the equation describing the transverse waves is given by (Fletcher, 1976):

$$\frac{\partial^2 \tilde{y}}{\partial t^2} = \frac{\zeta}{\mu}\frac{\partial^2 \tilde{y}}{\partial \tilde{x}^2} = c^2 \frac{\partial^2 \tilde{y}}{\partial \tilde{x}^2}, \tag{2.1}$$

with a similar equation for the $\tilde{z}$ displacement and where c is the velocity of the transverse waves.

Assuming that the end supports are rigid, where boundary conditions $\tilde{y}(0,t) = 0$ and $\tilde{y}(L,t) = 0$, the general form of string motion is written as:

$$\tilde{y}(\tilde{x},t) = \sum_{k=1}^{\infty} C_k \sin\left(\frac{k\pi\tilde{x}}{L}\right)\cos(2\pi f_k t + \psi_k), \tag{2.2}$$

where C_k is the modal amplitude, f_k is the frequency and ψ_k is the phase of harmonic number k.

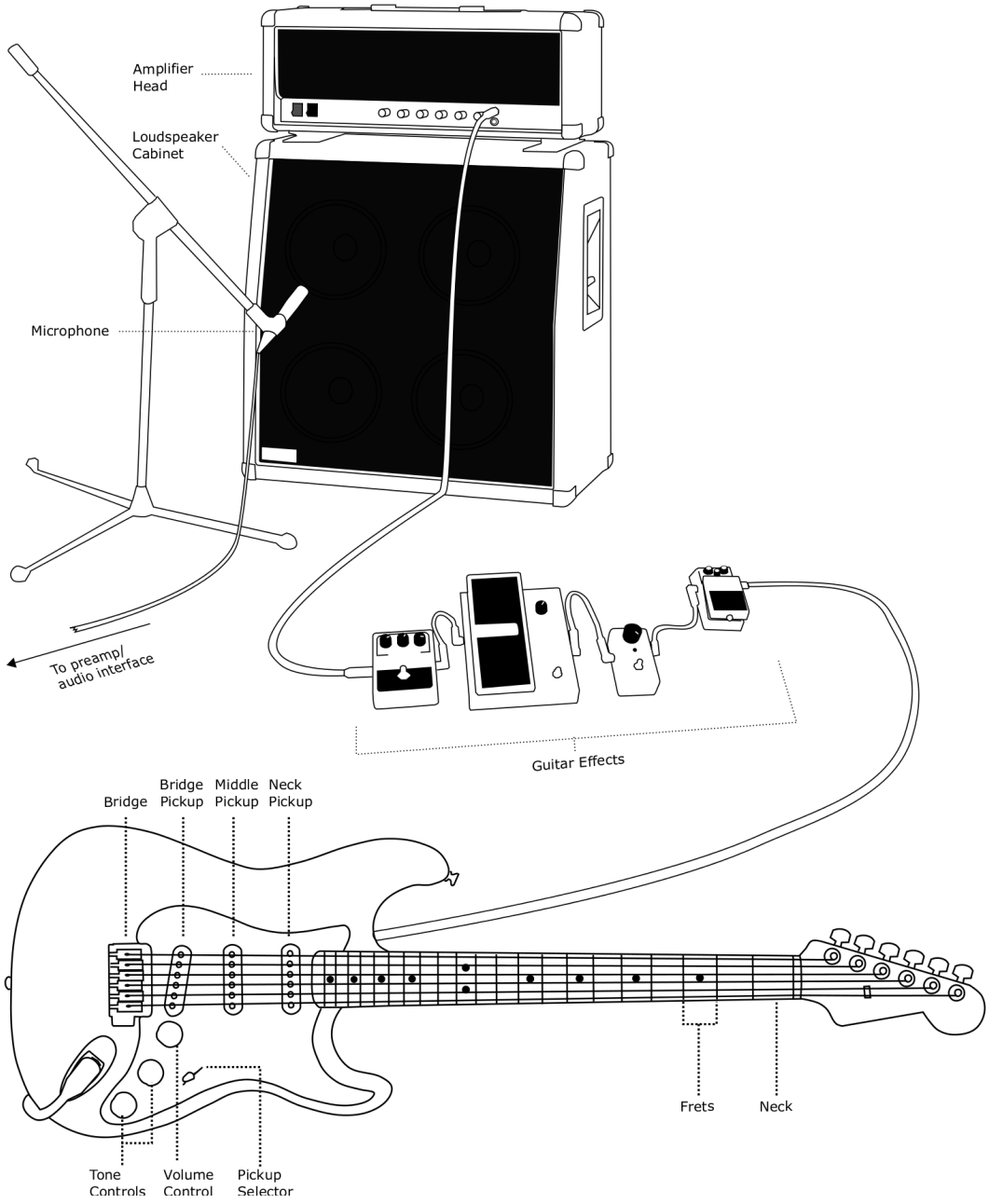


Figure 2.1: A typical electric guitar signal chain consists of: an electric guitar, guitar effects (can be excluded), a guitar amplifier, a loudspeaker cabinet and a microphone.

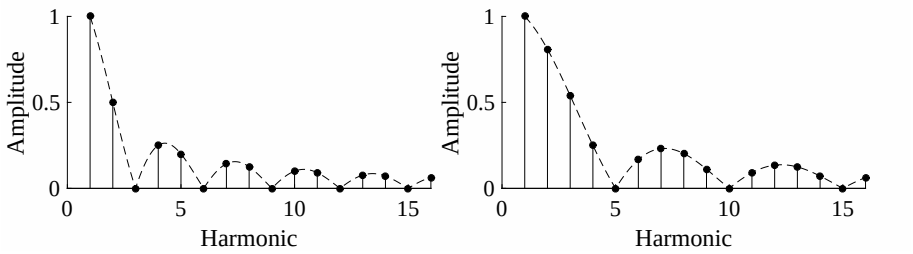


Figure 2.2: Spectra of the string plucked at (a) one-third of the string length and (b) one-fifth of the string length. The maximum of each spectral envelope is normalised to 1.

Suppose that a string is plucked at a distance ρ from the bridge with a vertical displacement a, its Fourier analysis which considers only the real part ($\psi_k = 0$) are given by (Fletcher, 1976):

$$C_k = \frac{2a}{\pi^2 R_\rho (1 - R_\rho)} \frac{\sin(k\pi R_\rho)}{k^2} \tag{2.3}$$

where R_ρ is the ratio between the distance of plucking point from the bridge, ρ.

Fig. 2.2 illustrates two examples of the effects of plucking positions on the spectrum. Fig. 2.2a and Fig. 2.2b show the spectra of a string plucked at one-third and one-fifth of the string length respectively. It is shown that every third and fifth harmonic is suppressed respectively. This means that a player can control the timbre of an electric guitar by varying the plucking point, where plucking near the bridge gives a brighter sound and plucking away from the bridge produces a warmer sound.

Plucking Width

An electric guitar string is usually plucked with a plectrum or a finger of finite width δ. The previous section assumes that the string is plucked with a plectrum of infinitesimally small width. The ideal string equation in Eq. 2.3 can be further extended to include the plucking width as:

$$C_k = \frac{2a}{\pi^2 R_\rho (1 - R_\rho) k^2} \int_{\rho - \frac{\delta}{2}}^{\rho + \frac{\delta}{2}} \sin(k\pi R_{\rho'})\, \mathrm{d}\rho' \tag{2.4}$$

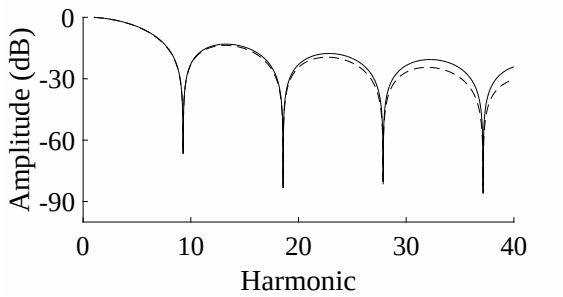


Figure 2.3: Spectra of the ideal string model with plucking width effect (dashed line) and without (solid line). The plectrum width is 20 mm for the string model with plucking width effect. Both strings are plucked at 70 mm from the bridge with a string length of 650 mm.

Deriving the integral:

$$\frac{1}{\delta}\int_{\rho-\frac{\delta}{2}}^{\rho+\frac{\delta}{2}} \sin\left(\frac{k\pi\rho'}{L}\right)\mathrm{d}\rho' = \frac{1}{\delta}\left[-\frac{L}{k\pi}\cos\left(k\pi\frac{\rho'}{L}\right)\right]_{\rho-\frac{\delta}{2}}^{\rho+\frac{\delta}{2}}$$
$$= \frac{2L}{k\pi\delta}\sin\left(\frac{k\pi\rho}{L}\right)\sin\left(\frac{k\pi\frac{\delta}{2}}{L}\right) \tag{2.5}$$

substituting Eq. (2.5) into Eq. (2.4)

$$C_k = A_C A_\delta \frac{S_\rho S_\delta}{k^3}, \tag{2.6}$$

where $A_C = \frac{2a}{\pi^2 R_\rho(1-R_\rho)}$, $A_w = 2/(\pi R_w)$, $S_\rho = \sin(k\pi R_\rho)$ and $S_\delta = \sin(k\pi\frac{R_\delta}{2})$.

The finite plucking width reduces the level of high harmonics with a factor of S_δ, where R_δ is the ratio between the plucking width δ and string length L. This effect introduces a 6 dB/octave rolloff above mode number $k = 2L/(\pi\delta)$, where harmonics above mode number $k_\delta = 2L/\delta$ are very little excited (Fletcher, 1976; Hall and Askenfelt, 1988; Chadefaux et al., 2012). Hence, this will approximately limit the spectrum to $k < k_\delta$.

Fig. 2.3 shows two spectra of the ideal string model in Eq. (2.3) plucked at 70 mm from the bridge with a string length of 650 mm which suppresses around every 9th harmonic with (dashed line) and without (solid line) plucking width effect. The low-pass filtering effect can start to be seen at higher harmonics.

Increasing the plucking width δ will result in a decrease in the amplitudes of higher harmonics. This is one of the reasons why the string plucked using a plectrum sounds brighter than using a finger.

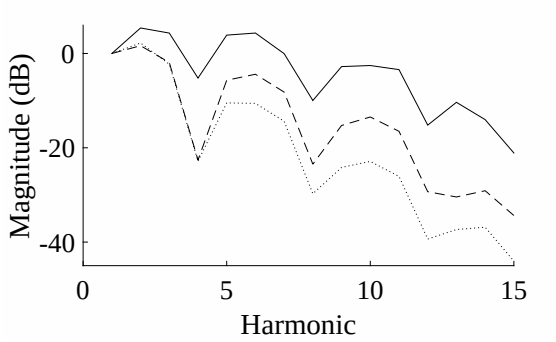


Figure 2.4: Magnitude spectral envelopes of an electric guitar played loudly (solid line), moderately-loud (dashed line) and softly (dotted line). The electric guitar is played on the open 1st string and plucked at 30 mm from the bridge, where the neck pickup configuration is selected. Each magnitude response is normalised to 0db for the fundamental.

Plucking Dynamic

A musician also has the capability of changing the tone of an electric guitar by varying the strength of the pluck. Varying the plucking dynamics of a plucked string instrument affects the level of the tone and the energy of higher partials (Jaffe and Smith, 1983; Laurson et al., 2001; Lindroos et al., 2011; Askenfelt and Jansson, 1993). The relative energy of higher partials reduces for softer plucks. Fig. 2.4 shows three magnitude spectral envelopes of an electric guitar played *forte* (loud), *mezzo-forte* (moderately-loud) and *piano* (soft) on the open 2nd string. The relative levels of the 6th partial resulting from mezzo-forte and piano plucks are 8.7 dB and 14.9 dB lower, respectively, compared to a forte pluck.

In addition to the attenuation of higher partials, Lindroos et al. (2011) also reported that the decay rate of partials changed rapidly for fortissimo plucks, whereas the effect is less dramatic for mezzo-forte and pianissimo plucks. Plucking dynamics also cause a pitch glide at the beginning of the tone, where stronger plucks create a larger pitch glide than softer plucks due to increased tension in the string (Lindroos et al., 2011; Järveläinen and Välimäki, 2001; Lee et al., 2009).

String Material

Thus far, the ideal string is assumed to be flexible. Real strings, particularly electric guitar strings, are made of steel and have non-negligible diameters. As a result, the harmonics are no longer exact integer multiples of the fundamental due to the stiffness of the string. For

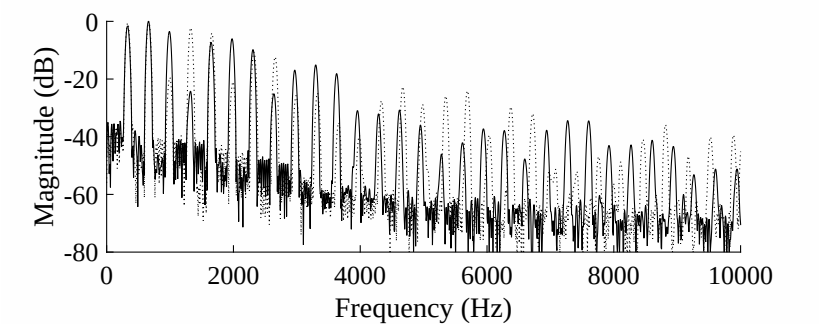


Figure 2.5: Spectra of two electric guitar tones played on the open 1st string (solid line) and fifth fret of the 2nd string (dotted line) ($f_0 = 330$ Hz).

a plucked string with a fundamental frequency f_0 and inharmonicity factor B, the partial frequencies f_k are given by Fletcher (1964):

$$f_k = k f_0 \sqrt{1 + Bk^2} \tag{2.7}$$

where the inharmonicity coefficient B is calculated as:

$$B = \frac{\pi^3 Q \xi^4}{64 L^2 \zeta} \tag{2.8}$$

where the term Q is Young's modulus, ξ is the diameter of the string and ζ is its tension. Thus, the size and material of the string affect the timbre of the electric guitar. Furthermore, the lowest pitched string has a higher inharmonicity factor compared to the highest pitched string mostly due to the larger string diameter. Also, the inharmonicity coefficient increases when the string played is fretted. By ignoring any minor changes in the tension ζ due to fretting the strings from Eq. (2.8), the inharmonicity coefficient when fretted can be expressed as (Lindroos et al., 2011; Barbancho et al., 2012):

$$B_{\mathcal{F}} = B_0 2^{\frac{\mathcal{F}}{6}} \tag{2.9}$$

where $\mathcal{F}$ is the fret number, $B_{\mathcal{F}}$ is the inharmonicity coefficient of the string at fret $\mathcal{F}$ and B_0 is the inharmonicity coefficient of the open string.

Fig. 2.5 shows spectra of two tones of the same pitch played on different strings of an electric guitar. The E4 tones ($f_0 = 330$Hz) are played on the open 1st string and the fifth fret of the 2nd string. It is shown that the higher partials of the electric guitar played on the 2nd string are stretched further towards high frequencies than those on the 1st string. The inharmonicity coefficient for the 2nd string is higher than that of the 1st string due to its larger diameter and the fret number played.

Järveläinen and Karjalainen (2006) found that the inharmonicity of strings is perceived to have a larger effect on lower pitched strings, thus, it is recommended to incorporate this effect for accurate synthesis of electric guitar tones.

Electric guitar string manufacturers such as Ernie Ball, D'Addario, Elixir, Dunlop, GHS, Fender and Gibson produce different sets of strings with different gauges and materials. Musicians have the capability of choosing amongst these products, where their differences will affect the timbre of an electric guitar.

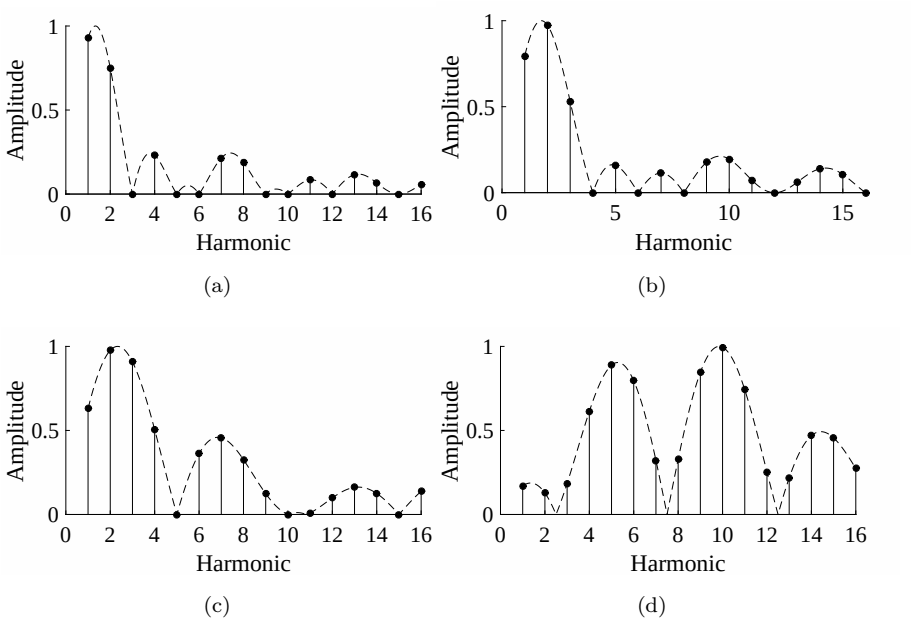


Figure 2.6: (a) Spectral envelope for a string plucked at one-third of the string length with a pickup placed at one-fifth of the string length. (b) Spectral envelope of string plucked at a quarter of the string length with pickup placed at one-sixth of the string length. (c, d) Spectral envelope of a string of length 650 mm plucked at 30 mm from the bridge with (c) in-phase and (d) out-of-phase mixed pickups located at 100 mm and 160 mm from the bridge. The maximum of each spectral envelope is normalised to 1.

String-Fret and String-Fretboard Interactions

The way a musician plays a guitar can significantly change the sound, therefore, each factor must be known in order to realistically replicate its expressive sound. There are other factors that can affect the sound of a guitar such as string-fret and string-fingerboard

interactions. These interactions are nonlinear indeed deforming the shape of the ideally vibrating strings and there are several recent papers discussing about these phenomena (Evangelista and Eckerholm, 2010; Evangelista, 2011; Bilbao and Torin, 2014, 2015).

For an electric guitar which typically has a curved fret, the string is free to move in the horizontal direction (direction that is parallel to the fret) but subject to friction on the fret surface. Evangelista (2011) discusses that this results in the two polarisation modes showing different decay times and their fundamental frequencies differ and vary with time due to unequal elongation of the string.

Bilbao and Torin (2014) provided a visualisation of the time evolution of the string profile under different plucking forces. There is one case that the string vibration is free from colliding with the frets, while other cases show some collisions distorting the profile of the string.

Bilbao and Torin (2014, 2015) also provided a simulation of the time-varying finger position, where the finger slides over a single fret during a playing gesture. Of course, this will effectively change the pitch of the tone. This will also cause a strong variation of the slope of the string at the fret location.

Pickup Position

The magnetic pickup senses the velocity of the string (Horton and Moore, 2009), which therefore requires a time derivative of the ideal string equation (see Eq. (2.2)). The velocity of the ideal string that is sensed at a single point d is given by:

$$v(t) = A_v \sum_{k=1}^{\infty} \frac{S_\rho S_d \sin(2\pi f_k t)}{k} \tag{2.10}$$

where $A_v = \frac{-2ac}{\pi L R_\rho (1-R_\rho)}$, c is the velocity of transverse waves, $S_\rho = \sin(k\pi R_\rho)$, $S_d = \sin(k\pi R_d)$, R_d is the ratio between the distance of pickup position from the bridge, d and the string length L and $f_k = ck/(2L)$ are the modal frequencies.

The effect of pickup placement and plucking point on the timbre can be shown by considering its spectrum. The spectrum of the velocity of the ideal string sensed at a single point d can be computed as:

$$\hat{V}_k = A_v \frac{S_\rho S_d}{k} \tag{2.11}$$

Note that the pickup position produces a comb-filtering effect similar to the plucking point effect, where harmonics whose harmonic number is a multiple of $\frac{L}{d}$ are suppressed. More

harmonics are suppressed if the pickup position is further away from the bridge. This is the reason why a neck pickup sounds warmer than a bridge pickup. Fig. 2.6a shows the magnitude spectrum where the string is plucked at one-third of the string length with a pickup placed at one-fifth of the string length. It is shown that every third and the fifth harmonic is suppressed. Another example is shown in Fig. 2.6b where every fourth and sixth harmonic is suppressed for an electric guitar plucked at a quarter of the string length with a pickup sensed at one-sixth of the string length.

An electric guitar commonly has an option of mixing two pickups together. Tillman (2002) and Paiva et al. (2012) studied the effect of mixed pickups. The electric guitar model in Eq. (2.11) can be extended to include mixing two pickups of distance d_1 and d_2 along the string of length L:

$$\hat{V}_k = A_v \frac{S_\rho S_\mu^+}{k} \tag{2.12}$$

where $S_\mu^+ = S_{d_1} + S_{d_2}$ is the summation of the two sine functions for the single pickups, which can be rewritten using trigonometric equation:

$$S_\mu^+ = 2\sin(k\pi R_\alpha)\cos(k\pi R_\beta) \tag{2.13}$$

where $\alpha = \frac{d_1+d_2}{2}$, $\beta = \frac{d_1-d_2}{2}$, $R_\alpha = \alpha/L$ and $R_\beta = \beta/L$. Thus, a mixed pickup signal produces a sine function that relates to the average of the two pickup locations α and a cosine function that relates to half of the distance between the two pickup locations β. Some electric guitar pickups have out-of-phase connections which can be modelled as:

$$\hat{V}_k = A_v \frac{S_\rho S_\mu^-}{k} \tag{2.14}$$

$$S_\mu^- = 2\sin(k\pi R_\beta)\cos(k\pi R_\alpha) \tag{2.15}$$

where $S_\mu^- = S_{d_1} - S_{d_2}$ represents two mixed out-of-phase pickups. The in-phase connection of the two pickups is more typically used than the out-of-phase connection (Paiva et al., 2012).

Fig. 2.6c and 2.6d show examples of the spectral envelope of electric guitar tones with in-phase and out-of-phase mixed pickups respectively. Notice that every fifth harmonic ($\frac{L}{\alpha}$) is weakly sensed for the in-phase mixed pickup and strongly sensed for the out-of-phase mixed pickup.

Note that a humbucker pickup (or a double coil pickup) can be considered as a mixed pickup. It usually combines two single pickup signals together with opposite coil winding

directions and polarities enabling the guitar signals to be in-phase while the hum (or noise) at 50 Hz is out of phase. This eliminates the hum.

Pickup Width

A magnetic pickup senses the velocity of a string around an area of width w rather than at a single point. Hence, the electric guitar model in Eq. (2.11) can be further extended to include the effect of pickup width as:

$$\hat{V}_k = A_v \frac{S_\rho}{k} \frac{1}{w} \int_{d-\frac{w}{2}}^{d+\frac{w}{2}} \sin(k\pi R_{d'})\, \mathrm{d}d' \tag{2.16}$$

Expanding the integral:

$$\begin{aligned} \frac{1}{w} \int_{d-\frac{w}{2}}^{d+\frac{w}{2}} \sin\left(\frac{k\pi d'}{L}\right) \mathrm{d}d' &= \frac{1}{w} \left[-\frac{L}{k\pi} \cos\left(k\pi \frac{d'}{L}\right) \right]_{d-\frac{w}{2}}^{d+\frac{w}{2}} \\ &= \frac{2L}{k\pi w} \sin\left(\frac{k\pi d}{L}\right) \sin\left(\frac{k\pi \frac{w}{2}}{L}\right) \end{aligned} \tag{2.17}$$

substituting Eq. (2.17) into Eq. (2.16)

$$\hat{V}_k = A_v A_w \frac{S_\rho S_d S_w}{k^2} \tag{2.18}$$

where $A_w = 2/(\pi R_w)$ and $S_w = \sin(k\pi \frac{R_w}{2})$. The effect adds a 6 dB/octave rolloff above mode number $k = 2L/(\pi w)$, where harmonics above mode number $k_w = 2L/w$ are very little excited. Notice that the pickup width effect is similar to the plucking width effect where a wider pickup sensitivity lowers the level of high harmonics. The limit of the spectrum is now reduced to $k < \min(k_\delta, k_w)$.

In this model, the area sensed is assumed to have a rectangular shape. In practice, the string is more strongly sensed around the middle of the pickup than at the ends. Paiva et al. (2012) proposed using a Hamming window to model the pickup width effect.

The final electric guitar model can be computed by introducing the pickup width effect and plucking width effect into the in-phase mixed pickup model by substituting Eq. (2.17) into Eq. (2.12) and adding the plucking width factor, where w_1 and w_2 are the pickup widths of the two pickups:

$$\hat{V}_k = A_v A_\delta \frac{S_\rho S_\delta}{k^3} \left(\frac{2S_{d_1} S_{w_1}}{\pi R_{w_1}} + \frac{2S_{d_2} S_{w_2}}{\pi R_{w_2}} \right) \tag{2.19}$$

where $A_\delta = 2/(\pi R_\delta)$. A mixed pickup usually has two pickups with the same width such that $w_1 = w_2$, so the model can be written as:

$$\hat{V}_k = A_v A_\delta A_w \frac{S_\rho S_\delta S_\mu^+ S_w}{k^3} \tag{2.20}$$

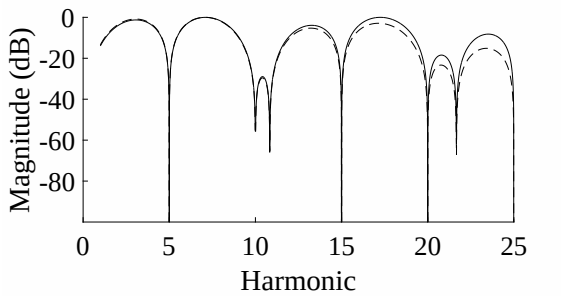


Figure 2.7: Spectra of the electric guitar model plucked with a plectrum of width 2 mm with pickups of width 20 mm (solid line) and 40 mm (dashed line). A mixed pickup is selected, where the two pickups are situated at 100 mm and 160 mm from the bridge. The string of length 650 mm is plucked at 30 mm from the bridge.

Fig. 2.7 shows two spectra of the electric guitar model from Eq. (2.20) with different pickup widths. Similar to the plucking width effect, a greater pickup width lowers the amplitude of higher harmonics.

Pickup Height

Typical electric guitar pickups have six individual magnetic pole-pieces to optimally capture the vibrations of each string. The height of the pickup can be adjusted; moving the pole-pieces closer to the strings will increase the overall output level of the pickup (Jungmann, 1994; Lindroos et al., 2011). The pickup height should ideally be adjusted to obtain even volumes from each string. Also, distortion occurs due to the magnetic pickup's interaction with the strings if the distance between the pickup and the string is too close (Jungmann, 1994; Celi et al., 2004).

Pickup Nonlinearity

The mapping between the string displacement and the resulting magnetic pickup signal is reported to be nonlinear (Horton and Moore, 2009; Paiva et al., 2012). The nonlinear mapping between the resulting magnetic pickup signal and the string's transverse vibration is different for the string vibrating in vertical and horizontal directions. Odd and even harmonics are generated with a high emphasis up to the third harmonic for vertical directions, whereas only even harmonics are generated for horizontal directions (Paiva et al., 2012). Mustonen et al. (2014) added that the magnetic pickup distorts vertical string vibrations

more significantly compared to horizontal vibrations.

2.1.2 Guitar Effects, Amplifiers, Cabinet Loudspeakers and Microphones

The electric guitar sound is not complete without including guitar amplification. Apart from the playing techniques of the musician and the electric guitar's components and settings, the guitar amplifier also plays a crucial role in defining the tone of an electric guitar. Some external devices may also be added such as guitar effects to modify the sound. In this section, a brief overview of guitar effects, amplifiers, cabinet loudspeakers and microphones is provided, and the factors that affect the sound of an electric guitar.

Guitar Amplifier and Loudspeaker

The output sound of a solid body electric guitar without plugging into a guitar amplifier is barely audible. Therefore, the signal needs some type of amplification to be played in a live performance. There are three types of guitar amplification known today which are tube, solid-state and digital amplifiers. Analogue guitar amplifiers such as tube and solid-state amplifiers use vacuum tubes and transistors respectively to increase the amplitude of the guitar signal. Digital amplifiers mostly replicate existing analogue guitar amplifiers in digital forms that can be used in computers, mobile phones and hardware devices. The anatomy of a basic guitar amplifier consists of pre-amplifier, tone control, power amplifier and loudspeaker. The pre-amplifier consists of a triode circuit that provides input matching and amplifies the guitar signal to a level that can drive the power stage. The tone control acts as an equaliser allowing a user to boost or cut certain frequency bandwidths. The power amplifier increases the power of the signal so that it is strong enough to drive the loudspeaker. The output transformer couples the power amplifier to the loudspeaker for impedance matching.

Early guitar amplifiers use vacuum tubes (before transistors were developed and widely used in electronics) that are easily driven into nonlinearity producing distortion. In Western popular music during the 1950s, blues, jazz and rock and roll guitarists such as Willie Johnson, Chuck Berry and Pat Hare utilised this effect by intentionally increasing the volume beyond the linear region. Most of them started experimenting to achieve their signature sound by adjusting the controls of the amplifier or deliberately modifying the components in

the amplifier. Since then, tube distortion has been increasingly popular amongst guitarists and music listeners. Meanwhile, solid-state guitar amplifiers are produced to compensate the disadvantages of vacuum tube amplifiers: large size, greater expense and high power consumption. The distortion produced by vacuum tubes and solid state amplifiers are different, and most guitarists tend to prefer the sound of tube distortion.

Some papers have studied the perceptual differences between vacuum tube and solid state distortion (Hamm, 1973; Bussey and Haigler, 1981; Santo, 1994). Subjective criteria such as "thin" and "metallic" are often used to describe solid-state amplifiers, and "warmth" and "punchier" are used to describe tube amplifiers (Bussey and Haigler, 1981). They sound similar when both of their frequency responses and group delay characteristics are matched and both are at a low distortion level (Santo, 1994). The audible differences between tube and solid-state amplifiers can only be heard when they are clipped. This is mainly because typical tube amplifiers produce a soft-clipping distortion, while solid-state amplifiers produce a hard-clipping distortion resulting in different harmonic content. Fig. 2.8a shows an example of soft-clipping and hard-clipping distortions, where hard-clipping distortion clips the signal abruptly due to the behaviour of transistors, while soft-clipping distortion produces a smoother clipping. Figs. 2.8b and 2.8c show both distortions produce odd harmonics with predominant third harmonic. Hard-clipping distortion produces more odd harmonics at high frequencies compared to soft-clipping distortion. The tube amplifier often clips the signal asymmetrically which means that the positive or negative values of the signal are clipped more or less than the other due to the offset of the bias point from the centre of the load line. Fig. 2.9a shows the symmetrically and asymmetrically clipped signal, where the asymmetrically soft-clipping distortion cuts the positive values more than the negatives. This produces even and odd harmonics as shown in Fig. 2.9b. Moreover, the high output impedance of the tube power amplifiers cause a peak in their frequency response at the resonant frequency of the speaker they are driving, which is reported to sound "warmer" than solid-state power amps (Santo, 1994).

The behaviour of the output transformer is another reason that makes a tube amplifier sound differently from a solid-state amplifier. The output transformer saturates along with the output of the vacuum tubes which introduces additional distortion and hysteresis, whereas solid-state amplifiers cannot model this effect (Barbour, 1998).

A typical guitar loudspeaker is designed as an electroacoustic transducer which converts the amplified guitar signal into vibration. It is designed differently from hi-fi speakers; its

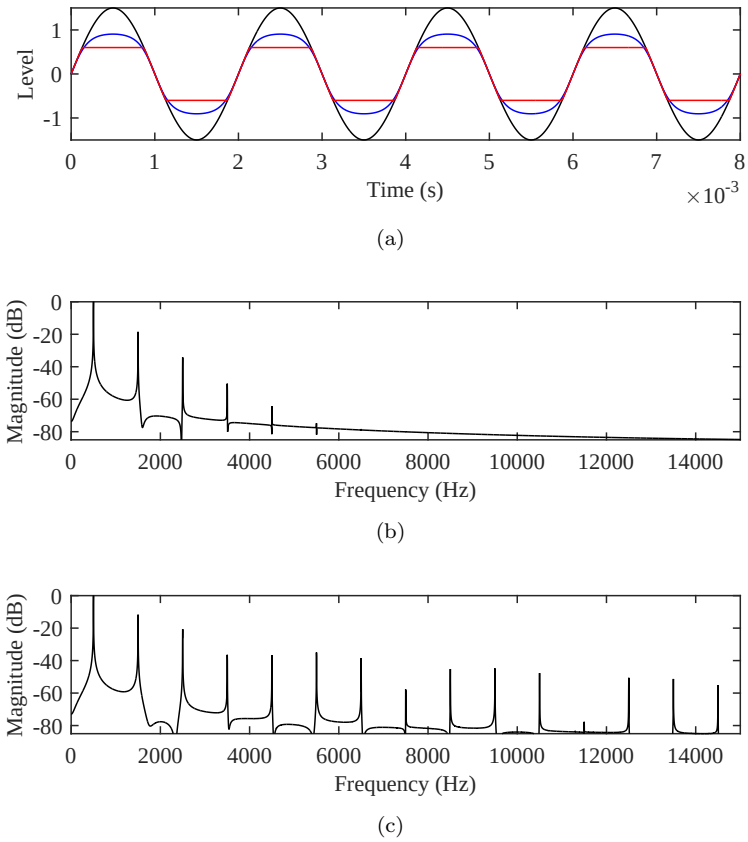


Figure 2.8: (a) Sinusoidal waveform (black) with soft (blue) and hard (red) clipping distortions. The spectra of (b) the soft and (c) hard clipped waveforms.

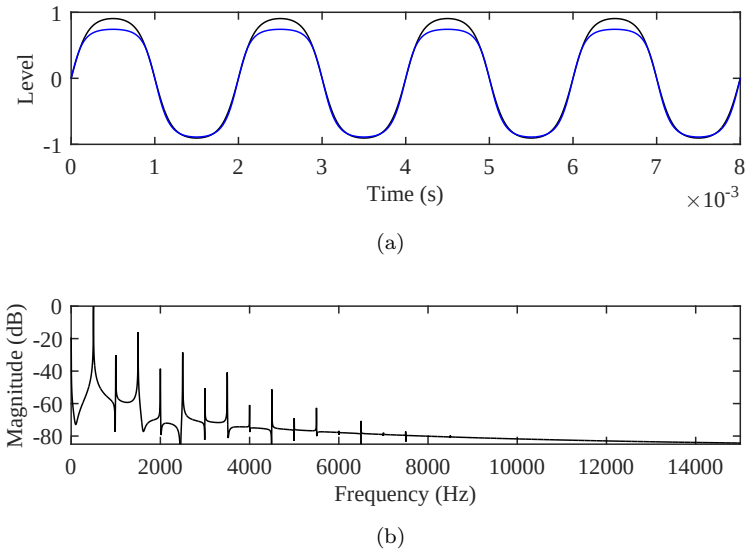


Figure 2.9: (a) Symmetrically (black) and asymmetrically (blue) soft clipped waveforms. (b) The spectrum of the asymmetrically soft clipped waveform.

frequency response is not flat and it produces some distortion artefacts, as these effects are desirable to guitarists. A loudspeaker may be built in with an amplifier in a wooden cabinet which is usually described as a "combo" amplifier, or as an external loudspeaker cabinet ready to be connected with an amplifier. There are a wide range of speaker configurations made available in guitar cabinets on the market, ranging from cabinets containing only a single speaker (e.g. 1 × 10" or 1 × 12") or multiple speakers (e.g. 2, 4 or 8 × 10"). There are also a wide selection of speakers, where each modifies the electric guitar tone differently. For an example, Fig. 2.1 shows a Marshall JCM800, where the amplifier (equipped with 100 W EL34 vacuum tubes) connects to an external loudspeaker cabinet (which contains 300 W 4 × 12" Celestion loudspeakers) via a cable. Zölzer (2008, p. 121) reports that the frequency responses of typical loudspeaker cabinets show an uneven bandpass characteristic with many resonances at mid frequencies. A selection of several frequency responses estimated from well-known loudspeakers are presented by Goetze (2017). Furthermore, the nonlinear characteristics of a loudspeaker cabinet were analysed by Yeh et al. (2008), where a loudspeaker contributes additional harmonic and inharmonic distortion.

The tone stack of a guitar amplifier is based on passive filter networks which typically produces a V-shaped equalisation, and provides users additional tonal control. A typical

tone stack has three knobs: Bass, Middle and Treble, that enable users to alter the gain of the respective frequency bands. Tone stacks that have different circuitry or different resistive and capacitive values will result in different frequency responses.

In conclusion, the electric guitar sound can be altered immensely by a guitar amplifier. Amongst popular guitar amplifiers such as Fender, Marshall, Vox and Mesa Boogie, each of them produces a unique tone where different components and electrical circuits have significant influence on the sound.

Guitar Effects

It is well understood that distortion produced by guitar amplifiers is a desirable effect. Other than distortion, effects such as artificial reverberation in guitar amplifiers are used by many musicians. These effects can be replicated externally. Guitar effects are available in a variety of forms, and the most popular form is in small boxes which are often described as "stompboxes" or "guitar pedals" because they are usually placed on the floor and guitarists step on their switches to turn the effects on and off, where examples can be seen in Fig. 2.1.

There are a plethora of guitar effects made available in the market that can further change a guitar tone. One of the most popular guitar effects is the distortion effect, which distorts the shape of the signal and produces harmonics. There are three main types that are often categorised by manufacturers which are *overdrive*, *distortion* and *fuzz*. These categorisations are often associated with the degree of signal clipping; Zölzer (2008, p. 124) states that an *overdrive* effect operates in the linear and nonlinear regions with a smooth transition, a *distortion* effect mainly operates in the nonlinear region and a *fuzz* effect operates completely in the nonlinear region. Examples of popular *overdrive* pedals are the Ibanez TS-808 Tube Screamer, Boss SD-1 Super Overdrive and Fulltone Full-Drive 2 Mosfet. For *distortion* pedals, typical examples are the Boss DS-1 Distortion and ProCo Rat, and examples for *fuzz* pedals are the Dunlop Fuzz Face and Electro-Harmonix Big Muff.

Time-based effects are also available such as *delay*, *chorus*, *reverb* and *flanger* effects. The *delay* or *echo* effect adds a copy of the original signal at a delayed time. Well-known delay guitar effects are the Boss DD-3 Digital Delay, MXR Carbon Copy and Electro-Harmonix Deluxe Memory Man. The *chorus* effect pedal mixes the signal with slightly different timbre and pitch, which attempts to mimic the effect of a choir or string orchestra. One notable use of this effect can be heard in Nirvana's 'Come as you are' recording. The *reverb* pedals often simulate the spatial effects produced by small and larger rooms. The *flanging* effect

has a short time delay and continuously varies the delay time with a low frequency (Zölzer, 2008, p. 76).

Microphones

Although we could stop at loudspeakers, this research involves analysing recorded electric guitar tones, thus, studying the effects of recording equipment such as the microphone are important. In order to record the sound of the electric guitar, a transducer is required, which is usually placed in front of the loudspeaker. There are a wide range of different techniques, microphone placements and microphone types for the audio engineer to choose from that can further alter the sound of an electric guitar.

Bartlett (1981) studies the perceptual tonal differences of microphone placements. A flat-response omni-directional microphone is used, and it is suggested that there are significant tonal differences due to different microphone placements compared to the reference sound. Furthermore, Case (2010) shows various frequency responses produced by different angles, distances from the center of the loudspeaker grille cloth and distances from the guitar amplifier. The plots shown suggest that there are significant differences in the responses for each variable. Generally, pop and rock genres rely on close microphone techniques for recording electric guitar. One of the motivations behind close microphone strategies (placing the microphone a few inches away from the amplifier) is to increase isolation from other musical instruments that are played simultaneously (Case et al., 2013).

There are a few types of microphones that are usually used in recording electric guitar such as dynamic, condenser and ribbon microphones. The sound captured by these microphones is perceived to differ with microphone choice, and audio engineers usually select the best type of microphone according to their preference. As an example, an audio engineer might choose a condenser microphone over a dynamic microphone because the microphone captures the low frequencies better (Senior, 2017). Furthermore, not only different types of microphones alter the sound of the guitar, but different models in each type have slight differences in their frequency response. Amongst dynamic microphones, the Shure SM57 is the most popular choice in relation to electric guitar recording, and it is suggested that its frequency response complements the electric guitar sound produced by the loudspeaker (Senior, 2017).

Techniques such as using multiple microphones can also alter the captured sound, which enables audio engineers to mix the sounds from each microphone and adjust their balance.

Finally, further effects are usually applied after recording such as compression, reverberation and equalisation.

2.2 Retrieving Information from Guitar Recordings

Extracting meaningful information from guitar recordings can find application in resynthesising the sound, automatic music transcription to produce a score or a guitar tablature, instrument recognition and many more. In this section, we discuss some existing techniques of extracting relevant information from guitar recordings.

2.2.1 Pitch Detection

One of the most important concepts in music is the pitch, and detecting the pitch of a guitar signal is essential for this research. In the time domain approach, finding the pitch period T_0 of the signal would yield the estimated fundamental frequency f_0, where:

$$f_0 = \frac{1}{T_0} \tag{2.21}$$

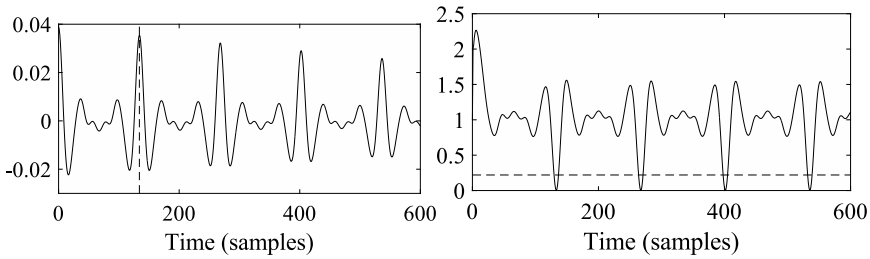
The autocorrelation function (ACF) of a signal $x(n)$ of duration N can be used to detect the pitch period, which is given by:

$$\Gamma(\tau) = \frac{1}{N} \sum_{n=0}^{N-\tau-1} x(n)x(n+\tau) \tag{2.22}$$

Generally, the highest peak of the ACF for positive lags τ corresponds to the pitch period. An example is given in Fig. 2.10a showing the ACF of an electric guitar plucked on the open 1st string ($f_0 = 329$ Hz and a sampling rate of 44100 Hz). The highest peak of the ACF for positive lags is at 134 samples (as shown by the dashed line), which means that the estimated fundamental frequency is 329.1 Hz. However other peaks may appear such as peaks of similar amplitude at integer multiples of the pitch period.

An alternative approach to determine the periodicity of a signal is called the YIN algorithm proposed by de Cheveigné and Kawahara (2002). The sum of differences of a time frame of duration N from its shifted version is calculated as:

$$\gamma(\tau) = \sum_{n=0}^{N-\tau-1} (x(n) - x(n+\tau))^2 \tag{2.23}$$



(a) (b)

Figure 2.10: (a) Autocorrelation function and (b) normalised mean difference function (YIN algorithm) of an electric guitar plucked on the open 1st string ($f_0 = 329$Hz).

The function shows troughs at lags of high correlation instead of peaks like the ACF. The difference function is then normalised:

$$\gamma'(\tau) = \frac{\gamma(\tau)}{\frac{1}{\tau}\sum_{n=1}^{\tau}\gamma(n)} \tag{2.24}$$

The first minimum of the normalised difference function γ' below a fixed threshold corresponds to the pitch period. In Fig. 2.10b, the threshold is set to 0.22 and the first minimum is at 134 samples, which means that the estimated fundamental frequency is 329.1 Hz. Quadratic interpolation can be used to further refine the location of the trough (Smith, 2011).

A comparison of several pitch detection methods using time or frequency domain approaches for monophonic guitar signals is presented by von dem Knesebeck and Zölzer (2010). It is suggested that the YIN algorithm is the most suitable pitch detection method for real-time single note guitar tracking. For polyphonic guitar signals, Klapuri (2003) proposed a frequency domain approach to estimate the fundamental frequencies along with the inharmonicity coefficient and the spectral envelope.

2.2.2 Onset Detection

Accurate detection of the starting time of each musical note is important in many music signal analysis applications. There are several onset detection methods reviewed and compared by (Bello et al., 2005; Dixon, 2006), such as using spectral flux, high frequency content, phase deviation and complex domain methods.

The High Frequency Content (HFC) of a signal is computed by applying a linear weighting

to the local energy as follows:

$$HFC(n) = \frac{1}{N} \sum_{f=-\frac{N}{2}}^{\frac{N}{2}} |f| \cdot |X(n,f)| \tag{2.25}$$

where $X(n,f)$ represents the amplitude of the fth frequency bin of the nth frame and N is the total number of frames. Penttinen and Välimäki (2004) extend this method by applying a high pass filter with 6kHz cutoff frequency before calculating the HFC of each frame. The energy at high frequencies is prominent only during onsets of a plucking event of a guitar and quickly decays over time. Applying the high pass filter takes advantage of this and highlights the energy of high frequencies, resulting in a more accurate onset estimation.

The Spectral Flux (SF) method takes a time-frequency domain approach, where a short time Fourier Transform of the signal with a Hamming window is calculated for every frame. The power spectrum for a frame is compared against the previous frame, and positive changes in magnitude in each frequency bin are summed, which gives the onset function SF (Masri, 1996, p. 137–141):

$$SF(n) = \sum_{f=-\frac{N}{2}}^{\frac{N}{2}} H(|X(n,f)| - |X(n-1,f)|) \tag{2.26}$$

where $H(x) = \frac{x+|x|}{2}$ is the half-wave rectifier function.

So far, SF and HFC are based on the magnitude frequency content of the signal. The complex domain approach also considers the phase information of the signal by calculating the expected amplitude and phase of the current bin $X(n,f)$, based on the previous two bins $X(n-1,f)$ and $X(n-2,f)$. The target value $X_T(n,f)$ is calculated by assuming constant amplitude and rate of phase change (Dixon, 2006):

$$X_T(n,f) = |X(n-1,f)|e^{\psi(n-1,f)+\psi'(n-1,f)}, \tag{2.27}$$

where $\psi(n,f)$ is the phase of $X(n,f)$ and $\psi'(n,f) = \psi(n,f) - \psi(n-1,f)$. Thus, the Complex Domain detection function (CD) is defined as the sum of absolute deviations from the target values:

$$CD(n) = \sum_{f=-\frac{N}{2}}^{\frac{N}{2}} |X(n,f) - X_T(n,f)|. \tag{2.28}$$

Dixon (2006) extends the CD method using a similar idea used in the SF method, where only increases in energy in spectral bins are considered to distinguish onsets from offsets. This technique is called the Rectified Complex Domain (RCD) method.

Fig. 2.11 shows examples of the onset detection methods. The methods chosen for this demonstration are the Spectral Flux (SF), High Frequency Content as proposed by Penttinen and Välimäki (2004) (P-HFC) and Rectified Complex Domain (RCD) methods. Fig. 2.11a shows the signal under test. All of the onset detection functions are normalised to have a zero mean and standard deviation of 1. The window size for each method is set to 20 ms with a 50% overlap. Since the tone is a single pluck, the maximum of each detection function yields the estimated onset time.

For this example, the expected onset time is 1.824 seconds and the onset times estimated from using the SF, P-HFC and RCD methods are 1.81, 1.82 and 1.82 seconds respectively. These three methods correctly identify the onset of the tone, but further refinements are needed to get closer to the expected onset time.

2.2.3 Plucking Point and Pickup Position Estimation

Estimating the plucking point and pickup position of an electric guitar from an audio recording can be used as parameters for electric guitar synthesisers. Furthermore, the position of the pickup can help distinguish which pickup on the electric guitar is being used (for a known guitar) or which guitar model is being used (for an unknown guitar).

Several research papers propose methods to estimate the plucking point of an acoustic guitar, using either a frequency domain approach (Bradley et al., 1995; Traube and Smith, 2000; Traube and Depalle, 2003) or a time domain approach (Penttinen and Välimäki, 2004). In the frequency domain approach proposed by Bradley et al. (1995), the partials of the analysed guitar signal are identified, and the plucking point is estimated by finding the minimum error between the ideal string magnitude spectrum and the observed data spectrum. Traube and Smith (2000) used a similar method but corrected a flaw in the ideal string equation used by Bradley et al. (1995).

The plucking point can also be estimated by locating the minimum of the autocorrelation function of the signal. This is because after the plucking event, the pulses and their inverted reflections arrive at the under-saddle pickup (or microphone) at different times causing a strong negative correlation at time lag τ_ρ. Traube and Depalle (2003) introduced a technique that calculates the autocorrelation of the signal from the log amplitude spectrum which is called "log-correlation". The minimum in the log-correlation yields a good first approximation of the plucking point. Then, a weighted least squares is performed to refine the ideal string magnitude spectrum to better fit the observed data spectrum yielding

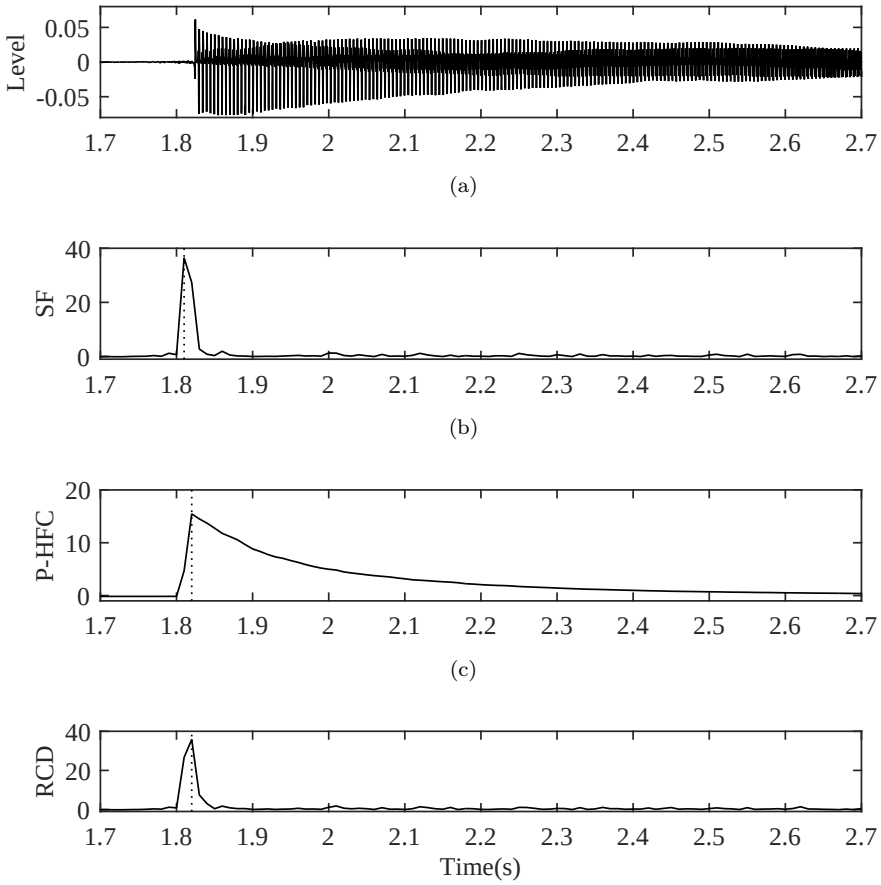


Figure 2.11: (a) Signal of an electric guitar plucked on the open 3rd string at 70 mm from the bridge with a pickup located at 45 mm from the bridge. (b) Spectral Flux, (c) High Frequency Content as proposed by Penttinen and Välimäki (2004) and (d) Rectified Complex Domain onset detection functions with dotted lines showing their maximum values.

a more accurate plucking point estimate. Penttinen and Välimäki (2004) proposed a much simpler technique to estimate the plucking point. The autocorrelation of a period of the tone is calculated, and the minimum of the autocorrelation is located to yield the plucking point estimate.

The existing methods expect the vibrations of an acoustic or a classical guitar string as input which would not work well electric guitar and electric bass guitar signals. An electric guitar signal has a combination of two strong comb filtering effects produced by the plucking point and pickup position. These two effects must be considered when finding the plucking point, therefore, it is much more feasible to estimate both parameters simultaneously when neither of the two locations is known. The existing methods can be extended to simultaneously estimate the plucking point and pickup position of an electric guitar. In Section 4.2.1 a novel technique that estimates the plucking point and pickup position of an electric guitar at the same time. The methods proposed by Bradley et al. (1995) and Traube and Depalle (2003) are extended by comparing the electric guitar model and the observed data instead.

Novel techniques that provide further improvements are proposed by analysing the autocorrelation of the electric guitar signal in Sections 4.2.2 and 4.2.3. Theoretically, the autocorrelation of an electric guitar signal has two minima that corresponds to the pickup and plucking positions. However, for a real electric guitar signal, the local minimum near zero lag is not apparent which is caused by the low-pass filtering effect made by the pickup width, plectrum width and plucking dynamic. A spectral flattening method is introduced to compensate for this effect which emphasises the local minima near zero lag. Furthermore, novel techniques to find the two minima that corresponds to the pickup and plucking positions are proposed which is either by grid searching (an exhaustive search through a specified subset of the hyperparameter space of a learning algorithm) or trough detection.

Details of the proposed techniques are presented in Chapters 4 and 5. Furthermore, the plucking point estimations proposed by Penttinen and Välimäki (2004) and Traube and Depalle (2003) are evaluated on electric guitar signals in Section 4.4.2.

2.2.4 Inharmonicity, String Detection, Playing Techniques and Decay Rate

There are several previously published papers that propose techniques to extract other relevant information from guitar recordings such as the inharmonicity coefficient of the string,

fretboard position, playing techniques and decay rate.

Rauhala et al. (2007) present an efficient iterative process for estimating the inharmonicity coefficients and the fundamental frequencies of synthetic and recorded piano tones. The inharmonicity factor B is estimated by minimising the differences between the expected and estimated partial frequencies. Dixon et al. (2012) introduce a technique for estimating the inharmonicity factor of synthetic and recorded harpsichord tones. It is estimated by taking the median of all inharmonicity coefficient estimates for each pair of estimated partial frequencies. This method is used by Mohamad et al. (2017b) in the pickup and plucking position estimation to extract the spectral peaks of an electric guitar signal, which will be discussed in detail later in Section 4.2.2.

The estimated inharmonicity coefficient can be used in detecting the fretboard position of a recorded guitar tone (Barbancho et al., 2012; Abeßer, 2012; Kehling et al., 2014), since each string and fret combination have different inharmonicity factor. Kehling et al. (2014) also propose a method to distinguish between 3 plucking styles of an electric guitar such as finger-style, picked, and muted and 5 expression styles such as slide, vibrato, bending, harmonics and dead-note. Meanwhile, Paté et al. (2014) introduce an efficient prediction of the decay time of electric guitar tones, where it could be used as a parameter for electric guitar synthesis.

Two papers attempt to distinguish three classical guitar models using Support Vector Machines (SVM) by extracting relevant features such as the time series of partial tones and the Mel Frequency Cepstral Coefficients from their recordings (Dosenbach et al., 2008; Fohl et al., 2012). The classification performance reduces significantly when the testing player is not included in the training because a player's interaction with the instrument has a large influence on the timbre of the guitar.

There are papers that propose a method to detect and distinguish typical guitar effects in electric guitar and bass recordings. Stein et al. (2010) introduces a technique to distinguish ten commonly known guitar effects using SVM based on extracted features such as spectral centroid, spread, skewness, flux, roll-off, slope and flatness from monophonic and polyphonic recordings. Stein (2010) extends the method to distinguish multiple audio effects since an electric guitar sound may have more than one audio effects applied to it. The method is able to successfully classify arbitrary combinations of up to three audio effects, but only monophonic electric guitar tones are considered.

2.3 Overview of Electric Guitar Synthesis

One recent trend in audio technology is the transition of guitar technology into the digital domain. This allows acoustic guitars, electric guitars, amplifiers, loudspeakers, effects and microphones to be digitally reproduced using state-of-the-art signal processing techniques to achieve close resemblance in sonic quality to existing analogue equipment. In order to replicate the sound of popular guitarists, each component discussed previously (see Sec. 2.1) should be accurately modelled. This section reviews approaches for emulating such components.

2.3.1 Physical Modelling of Electric Guitar

The Karplus-Strong algorithm is an early and simple synthesis algorithm for plucked string instruments, where the parameters available for control are pitch, amplitude and decay time (Karplus and Strong, 1983). Pairs of successive samples in the circular buffer are averaged which produces a slow decay of the waveform, and the resulting sound surprisingly resembles a natural plucked string even though the technique is quite simplistic. The circular buffer is initialised with a short noise burst to model the excitation which resembles a pluck. Further research conducted by Jaffe and Smith (1983) involves extending the existing Karplus-Strong model which improves the decay time and simulates the inharmonicity of the string, plucking dynamic, plucking position, glissandi and slurs. The noise burst is filtered with a feed forward comb-filter to model the effect of plucking point, where the notches suppress the harmonics that are related to the position of the pluck. The plucking dynamic is modelled with a one pole low-pass filter that controls the dynamic level. Moreover, the inharmonicity of the string is emulated using an all-pass filter.

A more specific project which focuses on synthesising the electric guitar sound is proposed by Sullivan (1990), whereby the Karplus-Strong algorithm is extended to include a pickup model with simple distortion and feedback effects. The simple distortion effect is emulated by a static waveshaper. Similar to the plucking point model, the pickup position effect is also modelled using a comb filter to suppress each harmonic in relation to the position of the pickup.

Jungmann (1994) confirms that there are distinctive differences in the electrical behaviour of popular guitar pickups, showing measurements and approximations of the pickup circuit responses and the influence of different cables and amplifier input resistances. The induc-

tance and capacitance in the circuit are shown to have a significant effect on the position of the resonance peak. He also reports that a guitar pickup exhibits nonlinear distortion due to the nonlinear magnetic field.

Commercial manufacturers such as Roland and Line 6 produce guitar synthesisers that use a hexaphonic pickup which outputs six signals, one from each string allowing them to process sound from each string separately (Hoshiai, 1994; Celi et al., 2004). Celi et al. (2004) model the pickups of popular electric guitars including the effects of pickup position, pickup mixing, height and magnetic apertures.

Cuzzucoli and Lombardo (1999) present a simple model for the interaction of the finger or pick with the string which is represented by a damped spring-mass system. They model three temporal phases of the player's action on the string which are excitation, release and damping. The manner in which the string is excited and released will affect the resulting guitar sound. Recent papers introduce techniques to model the interactions of the plucking fingers or guitar picks with the strings, the interactions of the fretting fingers with the strings, and the interactions of the strings with frets (Evangelista and Eckerholm, 2010; Bilbao and Torin, 2015).There are several recent papers discussing techniques to model the interactions or collisions between the string and the fret or fretboard. The contact between a vibrating string and a barrier is simulated by stable numerical formulations based on a modal expansion approach (van Walstijn and Bridges, 2016). Issanchou et al. (2018) discuss the percussive sound produced by the interactions such as "pop" and "slap" and investigated this nonlinear behaviour both numerically and experimentally.

Lindroos et al. (2011) introduce a parametric electric guitar synthesis model where conditions that affect the sound can be changed such as the force of the pluck, plucking point and pickup position. They introduce a novel excitation model that reproduces the sound of the plucking noise (which is audible especially when the guitar is played through a distortion effect). The inharmonicity of the pickup is taken into account and modelled with allpass filters. Paiva et al. (2012) added some details to model the electric guitar pickup. The pickup width effect is modelled with a Hamming window, replicating a pickup's magnetic aperture. The pickup circuit response effect that creates a linear resonant filtering can be modelled using the discrete-time transfer function of the pickup circuit or wave digital filters.

The mapping between the string displacement and the magnetic pickup signal is proven to be nonlinear. This behaviour can be modelled using a simple Hammerstein model consisting of a static nonlinear function followed by a linear filter (Paiva et al., 2012; Remaggi et al.,

2012). There is also an attempt to model this behaviour using a generalised Hammerstein model. Novak et al. (2016, 2017) reports that using a generalised Hammerstein model containing several parallel branches with static nonlinear functions and linear filters is not necessary, where a simple Hammerstein model seems to be sufficient at modelling the pickup's nonlinearity.

2.3.2 Modelling Guitar Amplifier and Effects

Pakarinen and Yeh (2009) provide a review of the techniques to digitally emulate vacuum-tube amplifiers. First, they discuss past literature of modelling the tone stack of vacuum-tube amplifiers. There are two main approaches which are the black-box approach and the white-box approach. In the black-box approach, various techniques presented by Foster (1986); Abel and Berners (2006) are both well-known for extracting the impulse responses for several settings of the parameters (low, mid and high tone knobs of the tone stack). These impulse responses can be used directly as a finite impulse response (FIR) filter to replicate the measured system. The white box approach is based upon deriving the discrete-time filter coefficients that simulate the circuit response of the tone stack. For an example, Yeh et al. (2008) derived the parameter values for the linear third-order transfer function to replicate the '59 Fender Bassman's tone stack. Next, Pakarinen and Yeh (2009) discuss literature on nonlinear modelling with or without memory. The most basic and direct approach is to apply a nonlinear mapping to the input signal to simulate the distortion effect. Examples of the nonlinear functions can be found in (Araya and Suyama, 1996; Doidic et al., 1998). These are described as static waveshapers because the mapping does not change over time, which does not describe the behaviour of a real tube amplifier. There are researchers who propose various dynamic waveshapers (Pritchard, 1991; Kuroki and Ito, 1998; Gustaffsson et al., 2012). Other techniques have also been proposed such as solving the ordinary differential equation simulating the behaviour of classic tube circuits (the white-box approach) (Karjalainen and Pakarinen, 2006; Yeh and Smith, 2008).

FIR filters are often used to approximate the linear behaviour of classic loudspeaker cabinets (Karjalainen et al., 2006; Goetze, 2017). Techniques mentioned earlier that are proposed by Foster (1986) and Abel and Berners (2006) can also be used to extract the impulse response of the loudspeaker cabinets and the microphone placements. Modelling the linear characteristics of loudspeaker cabinets is not sufficient to fully model a loudspeaker. Yeh et al. (2008) proposed a model with linear transfer functions and static nonlinearity, in

an attempt to fully approximate the real behaviour of a loudspeaker.

Yeh et al. (2007) proposed a computationally efficient approach to model distortion and overdrive effects, particularly, the Boss DS-1 Distortion and Ibanez Tube Screamer. The approach aims for fast computation for real-time use, where the model consists of a conditioning filter, followed by a static waveshaper and an equalisation filter. The filter transfer functions and memoryless nonlinearities are derived by analysing the circuits.

2.4 Summary

In this chapter, we have discussed how an electric guitar sound is produced, and the many factors that help shape a guitarist's signature sound. Understanding these factors are important in order to extract relevant information to replicate their sound. Then, published literature on retrieving information from guitar recordings and electric guitar synthesis are discussed.

Each of the individual components from the musician's playing technique to the microphone significantly influences the sound of the electric guitar. As Case et al. (2013) suggest, the electric guitar sound is a complex nonlinear system.

There are a wide range of electric guitar models, each having their own unique tone such as the Fender Stratocaster, Fender Jaguar, Fender Jazzman, Fender Mustang, Gibson Les Paul, Gibson SG, Rickenbacker 360 and Gretsch Country Gentleman. The differences in tone are mostly due to having different pickup positions, widths and heights. There are other factors that influence the sound such as pickup circuit response (Jungmann, 1994; Paiva et al., 2012) and the material of the body and neck of the guitar (Fleischer and Zwicker, 1999). Jungmann (1994) reported that there are significant differences among popular electric guitar pickups by measuring pickup circuit responses. The position of the resonance peak produced by the circuit response is different for each pickup, and is responsible for the unique sound of each individual pickup. The volume control circuit of the electric guitar is also reported to colour the output signal (Paiva et al., 2012). A musician can further alter the tone by varying the plucking position and the strength of the pluck, adjusting the tone knobs on the guitar and changing the strings with different gauges.

The selection of equipment used such as the model of the electric guitar, effects, amplifier, loudspeakers and microphones are chosen based on the musician's taste, where each of them have significant affect on the resulting sound. Thus, musicians such as Jimi Hendrix, Jimmy

Page, Brian May and Kurt Cobain and Johnny Marr, all have their recognisable tone. In this research, some of these factors that influence the sound of popular guitarists are studied, in order to be able to replicate their sound by extracting important parameters from their published recordings.

Chapter 3

Digitally Moving an Electric Guitar Pickup

The previous chapter discusses the complex nonlinear system of the electric guitar sound, where there are numerous of ways to alter the sound. Thus, replicating an arbitrary sound must not be simple. One method of replicating the sound could be implemented by extracting each parameter for electric guitar synthesis from recordings. Therefore, it requires an individual parameter extraction technique or an exhaustive search of a high-dimensional parameter space for each equipment used by the musician.

One of the concepts explored to achieve similar resulting sound as the electric guitar timbre of popular musicians is to morph the user's guitar sound into the target sound without having prior knowledge of the physical parameters of the desired guitar sound. The morphing of the input sound could be performed by applying a linear and a nonlinear filter. While the filters are unknown, the coefficients of these filters can be estimated using optimisation methods. The scope of this chapter is focused on analysing and obtaining the linear filter.

There are limitations of using this concept such as the input signal must be played as similarly as possible to the target signal, in order to only need to account for the differences due to the instrument and effects. This is because playing techniques, type of plectrum (e.g. finger or pick) and fret positions can alter the guitar sound. Preferably, the input signal needs to be played at the same plucking position, plucking dynamic, plucking angle, pitch, string and fret position with the same type of plectrum as the target sound. Other limitations will be discussed later in Section 3.6.

In this chapter, the concept is first explored by testing whether a sound from an arbitrarily selected magnetic pickup can be transformed so that it sounds like it comes from another pickup selection on the same electric guitar using a finite impulse response (FIR) filter. This preliminary experiment is an important step because the pickup has a large impact on the timbre of an electric guitar. The coefficients of the FIR filter are obtained using a least squares estimator. The robustness of the learnt filters are then tested on the following variables: plucking position, dynamics, fret position and random variation in human plucking.

The overview of the system is depicted in Fig. 3.1. The input signal $x(t)$ is an isolated tone captured by a pickup and the target signal $y(t)$ is the same tone captured by another pickup at a different position. The least squares estimator finds the coefficients of the filter $\hat{h}(t)$ which minimises the differences between the output and target signals. The learnt filter is applied to the input signal to produce the estimated signal $\hat{y}(t)$ replicating the target signal.

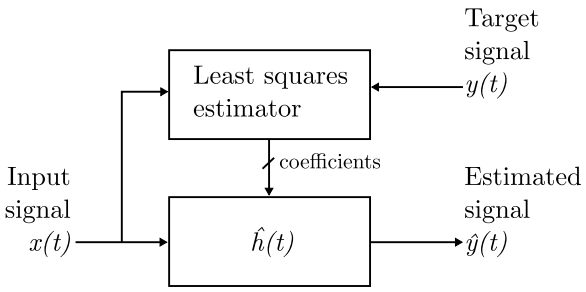


Figure 3.1: A schematic overview of the method that transforms the sound of a pickup into another.

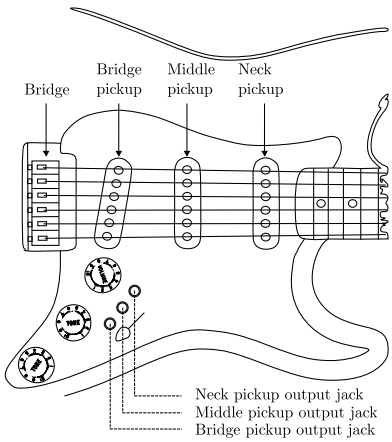


Figure 3.2: The modified Squier Stratocaster diagram (Mohamad et al., 2015). Three 1/8" output jacks allow us to tap separate signals from each magnetic pickup simultaneously. The three plucking positions are directly above each pickup.

Section 3.1 describes the electric guitar and audio samples used in this study. Section 3.2 presents an overview of the mathematical foundation to optimally determine the FIR filter coefficients. The morphing of the sound of a pickup into another is explained by providing a particular example, and the selection of the optimum number of coefficients is discussed in Section 3.3. Also in section 3.3, the calculation of the accuracy of the estimated signal is presented. Section 3.4.3 evaluates the robustness of the learnt filters when applied to an input signal with different repetitions, plucking positions, plucking dynamics and fret

positions. The conclusions from the experiment are presented in Section 3.6.

3.1 Dataset I

3.1.1 Electric Guitar Under Test

This work utilises a modified electric guitar which allows recording from each individual pickup output simultaneously. Therefore, each set of signals will be played at exactly the same plucking point, dynamic, pitch and timing. Fig. 3.2 shows the three 1/8" output jacks that allow signals from each pickup to be recorded simultaneously. The electric guitar that is used in this study is a Stratocaster model manufactured by Squier with three stock magnetic pickups. The pickup positions are situated around 160 mm (neck pickup PU_N), 102 mm (middle pickup PU_M) and 40 mm (centre of bridge pickup PU_B) measured from the bridge. The bridge pickup is slanted to 10°. Note that each bridge saddle position varies, which leads to slightly different measurements for each string. The guitar has an option to mix two pickups together. The mixed pickup configurations are a mix between neck and middle pickups PU_{N+M} and a mix between middle and bridge pickups PU_{M+B}.

Table 3.1 shows the precise measurements of the string length and pickup position for each string. The pickup positions are measured from the bridge saddle to the middle of the pickup, where the string is most strongly sensed. The strings used are nickel wound strings with gauges .010, .013, .017, .026, .036 and .046 inches manufactured by Ernie Ball.

3.1.2 Audio Samples

Isolated tones are recorded from the 3 single pickups which are played on 6 strings at 3 plucking points, 3 plucking dynamics and 3 fret positions with each condition repeated 3 times, leading to a total of 1458 samples. Details of the playing conditions are as follows: the plucking points are situated above each pickup; the plucking dynamics are loud (forte), moderately loud (mezzo-forte) and soft (piano); and the fret positions are played at open string, fifth fret and twelfth fret. The strings are plucked using a 0.88 mm thick plastic plectrum manufactured by Dunlop. All of the audio samples are recorded at 44100 Hz sampling rate, and the duration of the samples ranges from 3 – 28 seconds (depending on the decay rate of each string). This dataset will be referred to as Dataset I.

		Pickup distances from the bridge		
String	L (mm)	$d_{\mathrm{PU_B}}$ (mm)	$d_{\mathrm{PU_M}}$ (mm)	$d_{\mathrm{PU_N}}$ (mm)
1st, E4	649	38	99	157
2nd, B3	650	41	100	158
3rd, G3	652	45	102	160
4th, D3	651	46	101	159
5th, A2	652	49	102	160
6th, E2	650	49	100	158

Table 3.1: Measurements of string length L and pickup distances from the bridge for each string.

3.2 Estimating the FIR Filter Coefficients

The transformation of a pickup sound into another pickup on the same guitar can be achieved by convolving the input signal $x(n)$ with an FIR filter $h(n)$ to estimate the target signal $y(n)$, assuming that the filter is linear time invariant.

The filter $h(n)$ is unknown, thus, requiring an optimisation technique to estimate the coefficients for filter $\hat{h}(n)$. Barchiesi and Reiss (2009, 2010) proposed using a least squares estimator to find the coefficients of a filter to reverse engineer a target mix. The least squares estimator algorithm used in this study is proposed by Manolakis et al. (2000, p. 406).

For a causal discrete-time FIR filter, the filtering process in direct-form is written as:

$$y(n) = \sum_{m=0}^{M} h_m x(n-m) \tag{3.1}$$

where h_m is the unknown coefficient at index m of an FIR filter of length M. This can be expressed in matrix notation. To simplify the illustration, the number of samples N is 6 and

the number of filter taps is 3, so the expression is obtained as:

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ y(3) \\ y(4) \\ y(5) \end{bmatrix} = \begin{bmatrix} x(0) & 0 & 0 \\ x(1) & x(0) & 0 \\ x(2) & x(1) & x(0) \\ x(3) & x(2) & x(1) \\ x(4) & x(3) & x(2) \\ x(5) & x(4) & x(3) \end{bmatrix} \begin{bmatrix} h(0) \\ h(1) \\ h(2) \end{bmatrix}. \tag{3.2}$$

Since samples $x(-1), \cdots, x(-M+1)$ are not available, it is often filled with zeros. Also note that the number of samples of the input and target signal must be equal, otherwise, the signal with shorter length should be zero padded at the end so that the length of both signals are the same. In general, the above equation can also be written as:

$$\mathbf{y} = \mathbf{X}\mathbf{h} \tag{3.3}$$

where $\mathbf{X}$ is composed of shifted versions of the input signal $x(n)$, which is an $N \times M$ matrix, and $\mathbf{y}$ is a vector of the target signal with N number of samples and $\mathbf{h}$ is a vector of the unknown FIR filter coefficients with M number of samples.

The input and target signals are always much longer than the filter length (the filter length is discussed later in Section 3.4.1), where $N \gg M$ which means that the system is overdetermined. For an overdetermined system, the unknown coefficients $\mathbf{h}$ is found by minimising the energy of the error:

$$\begin{aligned} E &= \|\mathbf{y} - \mathbf{X}\mathbf{h}\|_2^2 \\ &= (\mathbf{y} - \mathbf{X}\mathbf{h})'(\mathbf{y} - \mathbf{X}\mathbf{h}) \\ &= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\mathbf{h} - \mathbf{h}'\mathbf{X}'\mathbf{y} + \mathbf{h}'\mathbf{X}'\mathbf{X}\mathbf{h} \\ &= \mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\mathbf{h} + \mathbf{h}'\mathbf{X}'\mathbf{X}\mathbf{h}, \end{aligned} \tag{3.4}$$

where $\mathbf{X}'$ and $\mathbf{y}'$ are the transpose of $\mathbf{X}$ and $\mathbf{y}$ respectively. Taking the derivative gives:

$$\frac{\partial E}{\partial \mathbf{h}} = -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\mathbf{h}, \tag{3.5}$$

and the derivative is then set to zero that leads to:

$$\mathbf{X}'\mathbf{X}\mathbf{h} = \mathbf{X}'\mathbf{y}. \tag{3.6}$$

Therefore, the set of optimal filter coefficients $\hat{\mathbf{h}}$ can then be found by:

$$\hat{\mathbf{h}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}. \tag{3.7}$$

This expression can be illustrated using the example in Eq. (3.2):

$$\begin{bmatrix}\hat{h}(0)\\ \hat{h}(1)\\ \hat{h}(2)\end{bmatrix} = \left(\begin{bmatrix} x(0) & x(1) & x(2) & x(3) & x(4) & x(5)\\ 0 & x(0) & x(1) & x(2) & x(3) & x(4)\\ 0 & 0 & x(0) & x(1) & x(2) & x(3)\end{bmatrix} \begin{bmatrix} x(0) & 0 & 0\\ x(1) & x(0) & 0\\ x(2) & x(1) & x(0)\\ x(3) & x(2) & x(1)\\ x(4) & x(3) & x(2)\\ x(5) & x(4) & x(3)\end{bmatrix} \right)^{-1} \begin{bmatrix} x(0) & x(1) & x(2) & x(3) & x(4) & x(5)\\ 0 & x(0) & x(1) & x(2) & x(3) & x(4)\\ 0 & 0 & x(0) & x(1) & x(2) & x(3)\end{bmatrix} \begin{bmatrix} y(0)\\ y(1)\\ y(2)\\ y(3)\\ y(4)\\ y(5)\end{bmatrix}. \tag{3.8}$$

The size of matrix $\mathbf{X}$ can be large which is computationally expensive and uses a lot of computer memory. The product of the multiplying matrices $\mathbf{X}'$ and $\mathbf{X}$ is the time-average correlation matrix $\hat{\mathbf{R}}$, where its elements $\hat{r}$ can be calculated as:

$$\hat{r}_{i,j} = \sum_{n=0}^{N-1} x(n+1-i)x(n+1-j) \quad 1 \leq i,j \leq M, \tag{3.9}$$

where i and j are the row and column indices of matrix $\hat{\mathbf{R}}$ respectively. So instead of multiplying matrices of size $M \times N$ and $N \times M$, the result of the multiplication (which is the time-average correlation matrix $\hat{\mathbf{R}}$ of size $M \times M$ matrix) can be obtained easily. Further techniques to reduce the number of operations to calculate the time-average correlation matrix $\hat{\mathbf{R}}$ are described by Manolakis et al. (2000, p. 408).

The multiplication between matrix $\mathbf{X}'$ and vector $\mathbf{y}$ is the cross-correlation vector $\mathbf{u}$ between the target signal and the input signal, which can be written as:

$$\mathbf{u}_m = \sum_{m=0}^{M-1} y(n)x(n-m). \tag{3.10}$$

So, substituting the matrix $\hat{\mathbf{R}}$ and the vector $\mathbf{u}$ in Eq. (3.7) gives:

$$\hat{\mathbf{h}} = \hat{\mathbf{R}}^{-1}\mathbf{u}. \tag{3.11}$$

Although the method described estimates the filter coefficients in the time domain, the least squares solution is equivalent in the frequency domain. Barchiesi and Reiss (2010) proved that the least squares equation is identical with orthogonal transforms.

3.3 Analysing the Filter with an Example

3.3.1 Timbral Similarity Measurement

Once the filter and the estimated sound are obtained, a timbral similarity measurement is needed to evaluate how close the resulting sound is to the target sound. One way of measuring the similarity of two signals is by measuring the difference between the sounds in the time-frequency domain, where the short-time Fourier transform (STFT) is typically used for time-frequency analysis. Lai et al. (2006) used this measurement in order to find optimal parameters for frequency modulation matching synthesis with a genetic optimisation approach.

The estimated and target signals are divided into short segments and a discrete Fourier transform is performed for each segment. The length of each frame is set to 1024 samples with an overlap of 512 samples. The distance (or error) D_R between the estimated and target signal is calculated as follows:

$$D_R(\hat{Y}, Y) = \frac{1}{N} \sum_{n=1}^{N} \sum_{f=1}^{F} |\hat{Y}(n, f) - Y(n, f)|^2 \tag{3.12}$$

where $\hat{Y}(n, f)$ is the magnitude spectrum of the estimated signal at frequency bin f at time frame n, $Y(n, f)$ is the magnitude spectrum of the target signal at frequency bin f at time frame n, N is the total number of frames in the STFT and F is the total number of frequency bins in each frame. The estimated sound is considered to be more similar to the target when the distance is near zero. Due to the variations in plucking dynamics in the dataset, the distance in Eq. (3.12) should be normalised to compensate for differences in loudness. Otherwise, the results will be biased toward *piano* (soft pluck) recordings. The raw distance D_R is divided by the average energy of the estimated and target signal, where the normalised distance D is given by:

$$D(\hat{Y}, Y) = \frac{2D_R(\hat{Y}, Y)}{\sum_{n=0}^{N-1} |\hat{y}(n)|^2 + \sum_{n=0}^{N-1} |y(n)|^2} \tag{3.13}$$

where N is the sample length of the estimate and target signals. In this experiment, the sample lengths of both signals are always the same, where the pickups are simultaneously recorded as mentioned in Section 3.1 (which means that the sample lengths of the input, estimated and target signals are the same).

3.3.2 Morphing a Pickup Sound

In this section, two examples are demonstrated: the transformations of the sound from the neck pickup $x(n)$ into the sound of the bridge pickup $y(n)$, and vice versa. In these examples, the electric guitar is plucked moderately loud (*mezzo-forte*) directly above the bridge pickup on the 3rd open string. Due to the effect of the plucking point, multiples of 14.5 harmonics are suppressed. Due to the effect of pickup position, the bridge and neck pickup positions suppress multiples of 14.5 and 4.1 partials respectively. Fig. 3.3a and 3.3c show two spectra X and Y for the neck and bridge pickup signals respectively. The effect of the pickup position can be seen in these figures, where the 4th (neck pickup) and 14th (bridge pickup) partials have low amplitudes.

The coefficients for FIR filters $\hat{h}_{x\to y}(n)$ and $\hat{h}_{y\to x}(n)$ are estimated using the method described in Section 3.2 to transform the sound of the neck into bridge pickup and the sound of the bridge into neck pickup respectively. In this example, the filter length M is set to 1024. Once the filter coefficients are obtained, $\hat{h}_{x\to y}(n)$ is applied to the signal $x(n)$ to produce the estimated bridge pickup sound $\hat{y}(n)$ and $\hat{h}_{y\to x}(n)$ is applied to the signal $y(n)$ to produce the estimated neck pickup sound $\hat{x}(n)$. The spectra of the signals $\hat{x}(n)$ and $\hat{y}(n)$ are shown in Fig. 3.3b and 3.3d respectively. The spectra of the estimated neck $\hat{X}$ and bridge pickup $\hat{Y}$ signals have similar spectral envelope as their target signals. The amplitude of every 4th partial of spectrum X (see Fig. 3.3a) is boosted (see. Fig. 3.3d) to match its target spectrum in Fig. 3.3c. Meanwhile, the amplitude of every 4th partial of spectrum Y (see Fig. 3.3c) is decreased (see. Fig. 3.3b) to match its target spectrum in Fig. 3.3a. The accuracies of the synthetic signals are calculated as the normalised distance D using Eq. (3.13). Distance $D(\hat{X}, X)$ is 0.0047 and distance $D(\hat{Y}, Y)$ is 0.1327, and comparing these distances with the initial distance $D(X, Y) = 0.7148$ yield reductions of 99% and 81% in spectral differences.

In the case where a partial has a very low amplitude and the target partial has a relatively higher amplitude, the estimated filters cannot amplify this partial. An example is shown where there are noticeable spectral peaks at 8970 Hz and 9218 Hz in Fig. 3.4a, and at the same frequencies, there are partials with very low amplitudes (almost completely suppressed) in Fig. 3.4c. In Fig. 3.4b, it is shown that the partial is not amplified to match the amplitude of its target partial. Evidently, a new frequency component cannot be generated by linear filters.

The estimated filters $\hat{h}_{x\to y}(n)$ and $\hat{h}_{y\to x}(n)$ that were obtained are analysed and com-

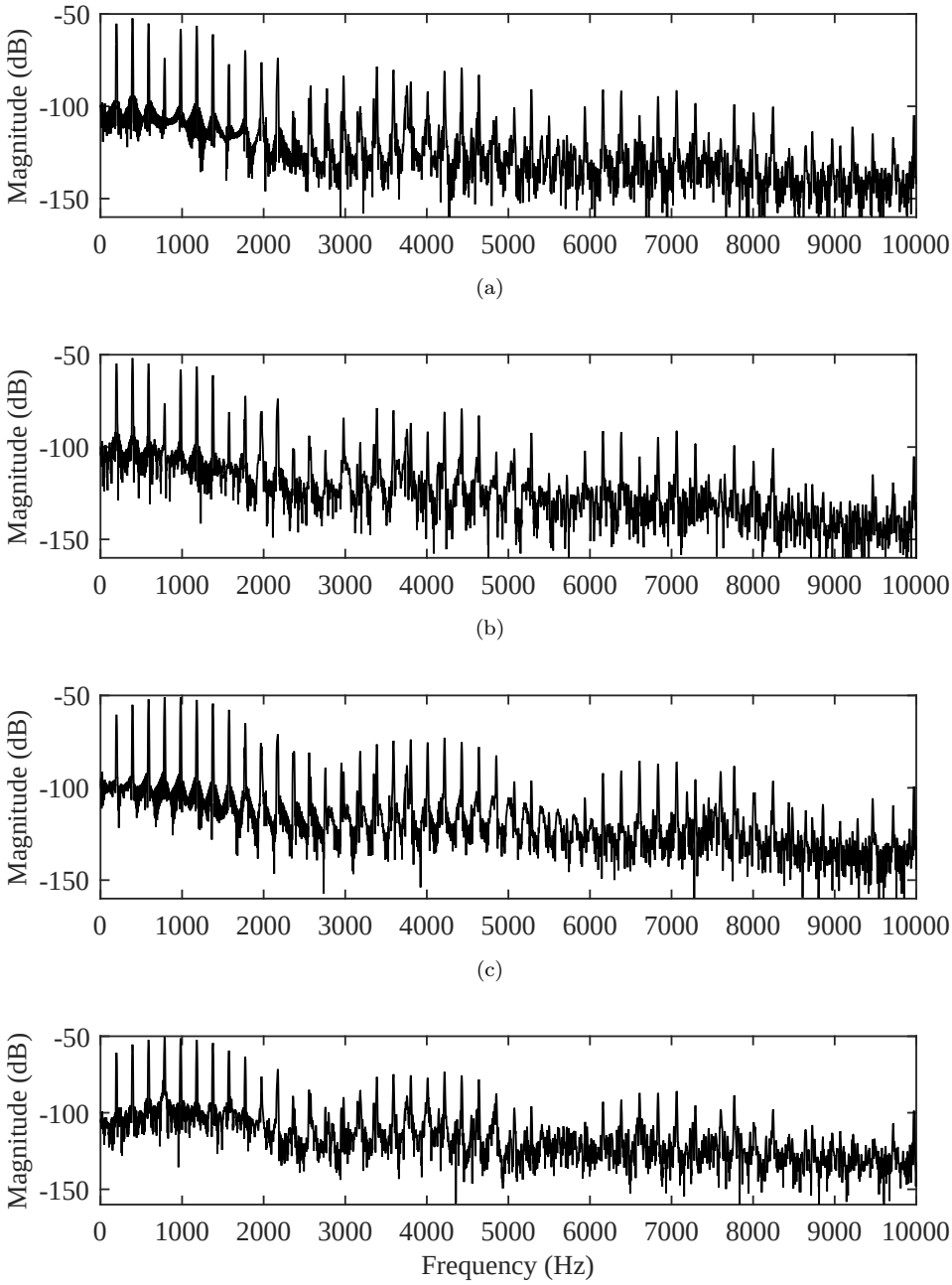


Figure 3.3: Magnitude spectra for the attack part of a guitar tone plucked directly above the bridge pickup on the open 3rd string ($f_0 = 196$ Hz), calculated from (a) the neck pickup signal, (b) estimated neck pickup signal, (c) bridge pickup signal and (d) estimated bridge pickup signal. The spectral analysis is performed on the attack part of the signal of length 8192 samples with a Hamming window and zero padding factor of 4.

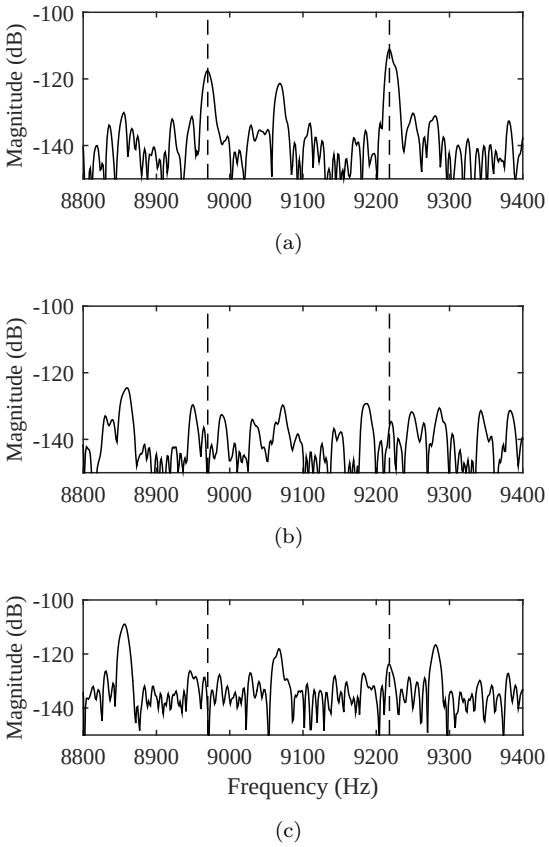


Figure 3.4: Zoom in version of Fig. 3.3. Magnitude spectra of (a) neck pickup, (b) estimated neck pickup and (c) bridge pickup.

pared with their ground truths. It is assumed that the ground truths are the spectral difference between input and target signal. Fig. 3.5 compares the ground truths (black lines) with the frequency responses of the estimated filters (transparent blue lines). The estimated filters tend to be closer to their ground truths at or near the partial frequencies (vertical dotted lines). This is due to most of the energy of the input and target signals being concentrated at the partials, so the filter optimisation is biased to yield low errors at these frequencies rather than the frequencies between partials.

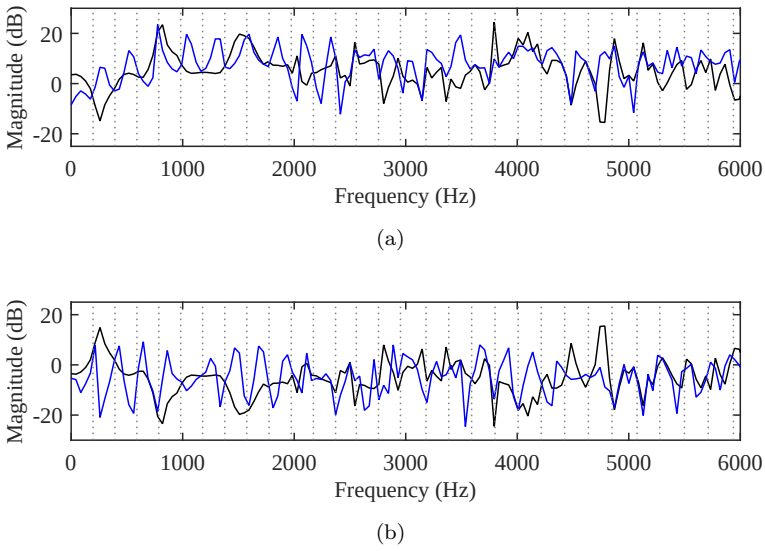


Figure 3.5: The spectral differences (a) $Y - X$ and (b) $X - Y$ drawn as black lines, and the frequency responses of (a) $\hat{h}_{x \to y}(n)$ and (b) $\hat{h}_{y \to x}(n)$ shown as transparent blue lines. The guitar is plucked directly above the bridge pickup on the open 3rd string ($f_0 = 196$ Hz). The partial frequencies are shown as vertical dotted lines.

3.4 Listening Test Results

A subjective evaluation is performed using a Multiple Stimuli with Hidden Reference and Anchor (MUSHRA) (Radiocommunication Sector of ITU (ITU-R), accessed March 6, 2018) style listening test using the Web Audio Evaluation Tool designed by Jillings et al. (2015) in order to evaluate the perceived sound similarity between target sounds and estimated sounds. A total of 10 listeners in the age group of 23 to 39 with various music and audio backgrounds participated in this study, where 7 of them plays at least one musical instrument and 4 of them are electric guitar players with 9 – 35 years of experience. Each participant is given an AKG-K92 headphone to listen to the audio samples.

Each participant is presented with 20 test pages asking to rate the sound similarity between a reference sound and some audio samples under test. Each test page has a reference sound (target sound), a hidden version of the reference, one or two hidden anchors (input sounds) and estimated sounds. A variety of fixed parameters e.g. strings and plucking positions are chosen for each test page so that the results would not be biased toward only one case. The similarity between the reference sound and the test samples is measured on

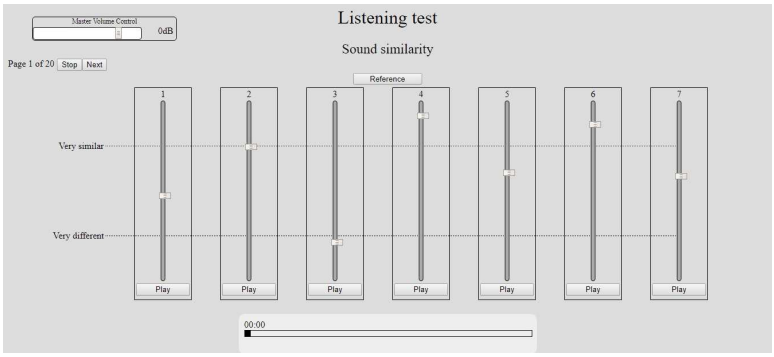


Figure 3.6: Listening test using the Web Audio Evaluation Tool.

a relative scale (0–1) from "very different" to "very similar" (Mehrabi et al., 2016), which can be seen in Fig. 3.6. A higher grade is given when the test sample is similar to the reference sound. Conversely, a lower grade is given when the test sample sounds different from the reference sound. The highest grade is only given when the listener cannot perceive any differences between the reference sound and the test sample.

This listening test consists of four parts: (1) 4 questions are dedicated to evaluate suitable filter lengths M for the emulation, (2) 3 questions are about morphing the three pickup sounds, (3) 12 questions are about the robustness of the filters when applied to different repetition, plucking dynamic, plucking position and fret position and 1 question is about improving the filter for different fret positions.

3.4.1 Suitable Filter Lengths

The filter optimisation finds the spectral differences between the input and target signals, which are biased to give low errors at partial frequencies where there are energies. Thus, setting the filter length M to have sufficient frequency resolution is crucial so that each spectral difference is modelled correctly, where $f_s/f_0 < M$. The sample rate f_s is 44.1 kHz and f_0 is the fundamental frequency of the input and target signals.

Open 6th String

As an example, the bridge pickup signal of the guitar plucked moderately loud on the open 6th string ($f_0 = 82.4$ Hz) directly above the bridge pickup is transformed into the neck pickup sound using filters with various numbers of filter taps. The solid line in Fig. 3.7 shows the normalised distances of the neck pickup signal and the estimated neck pickup

signal for various filter lengths. It can be seen that the error converges above $f_s/f_0 = 535$ samples.

Participants are given a hidden reference sound (neck pickup signal), a hidden anchor (bridge pickup signal) and 5 test samples (estimated neck pickup signal using filters with 5 different filter lengths) to compare with the reference sound (neck pickup signal). The filter lengths are 350, 450, 550, 1100 and 1650, which are shown as vertical dotted lines in Fig. 3.7.

Fig. 3.8a shows the box plots of the sound similarity ratings. The horizontal line in each box is the median, the top line in each box is the upper (75%) quartile values and the bottom line in each box is the lower (25%) quartile values. The whiskers ($|---|$) are the extent of the rest of the data except for outliers which are shown as cross symbols (+). As expected, the listeners are able to rate the hidden anchor as very different from the reference sound, where most of the data lies below 0.5. Also, the hidden reference are given high ratings, where all of the data lies above 0.5. The listeners gave low grades for the estimated sounds with filter lengths 350 and 450, which are expected since their frequency resolutions are not sufficient. High grades are given for the estimated sounds with filter lengths 550, 1100 and 1650, where their data are similar to the hidden reference. The estimated sounds that are very similar to the reference sound can also be confirmed by performing paired-sample t-tests comparing against the hidden reference data. The difference between the hidden reference and the estimated sounds with filter lengths 550, 1100 and 1650 are not statistically significant with significance level of 0.05 (a 5% chance of not being true).

A similar question is asked, where the input and target signals are reversed (morphing the sound of the neck pickup into the bridge pickup). The dotted line in Fig. 3.7 shows the normalised distances of the real bridge pickup signal and the estimated bridge pickup signal for various filter lengths, where the error slowly converges. The converged errors are higher than the previous example because the estimated bridge pickup sound is perceived to have noisier background sound than the real bridge pickup sound. This is because the noise floor in the neck pickup sound is also amplified to match the level of the bridge pickup sound. The amplitude of the bridge pickup sound is higher than the neck pickup due to a higher pickup height. Note that the root mean square (RMS) values for the bridge and neck pickup sounds are 0.0012 and 0.0007 respectively. Participants are given a hidden reference sound (bridge pickup signal), a hidden anchor (neck pickup signal) and 5 test samples (estimated bridge pickup signal using filters with the same filter lengths as the previous example) to compare with the reference sound (bridge pickup signal). Fig. 3.8b shows that the estimated

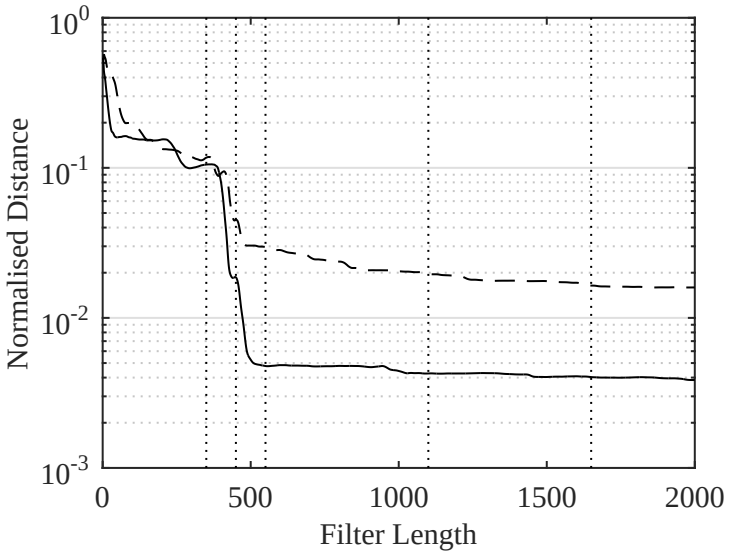


Figure 3.7: Normalised distances D between the estimated signals and their target signals as a function of filter length. The transformation of the bridge pickup sound into the neck pickup sound is shown as a solid line and the transformation of the neck pickup sound into the bridge pickup sound is shown as a dashed line. The electric guitar is plucked directly above the bridge pickup on the open 6th string. The vertical dotted lines at 350, 450, 550, 1100 and 1650 are the filter lengths used in the listening test. Note that the y-axis is in log scale.

sounds are increasingly similar to the reference sound, where the ratings for filter length 1650 is slightly lower than the hidden reference due to the perceivable amplified noise.

Open 1st String

A similar test as the case for the open 6th string is performed using recordings of the electric guitar plucked on the open 1st string ($f_0 = 329$ Hz) directly above the neck pickup. Similarly, two cases are tested which are transforming the sound of the bridge pickup into the neck pickup and transforming the sound of the neck pickup into the bridge pickup.

The errors shown in Fig. 3.9 converge above $f_s/f_0 = 134$ samples. The listening test results in Fig. 3.10 show that high grades are given for estimated sounds with filter lengths above 134 samples. The differences in the ratings for estimated sounds with filter lengths above 134 samples are not statistically significant from each other. This means that the

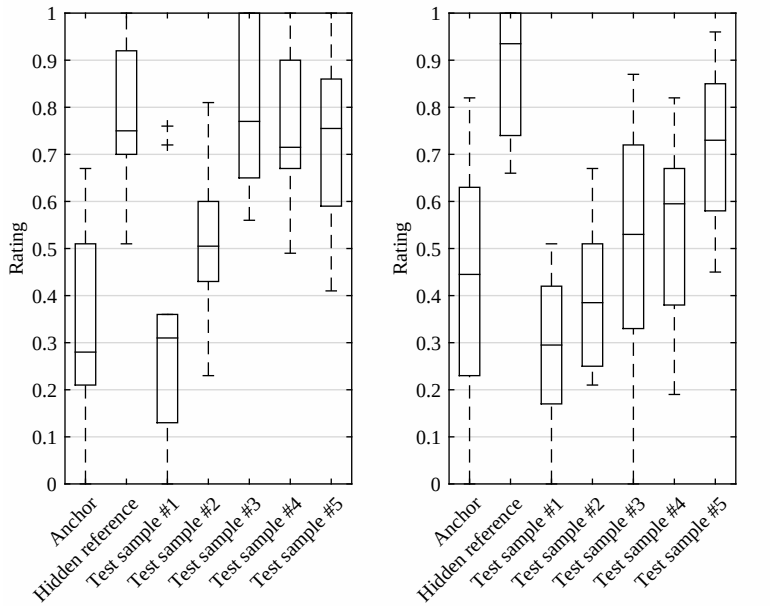


Figure 3.8: Box plots of the sound similarity ratings between the reference sound and test samples testing for finding suitable filter lengths. (a) The reference sound is the neck pickup signal, the anchor is the bridge pickup signal and test samples 1–5 are the estimated neck pickup signal with filter length 350, 450, 550, 1100 and 1650 respectively. (b) The reference sound is the bridge pickup signal, the anchor is the neck pickup signal and test samples 1–5 are the estimated bridge pickup signal with filter length 350, 450, 550, 1100 and 1650 respectively.

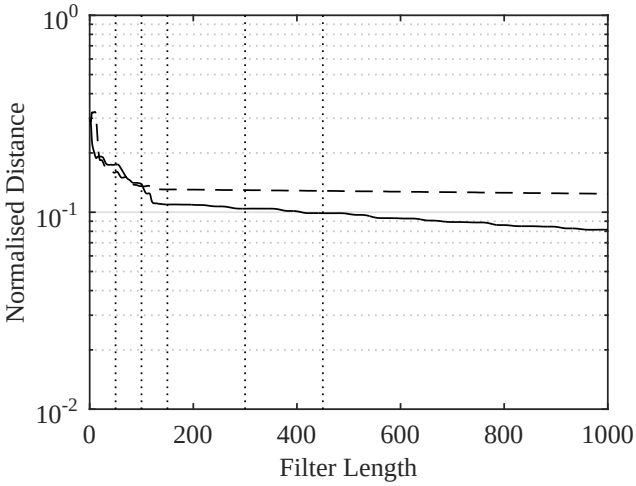


Figure 3.9: Normalised distances D between the estimated signals and their target signals as a function of filter length. The transformation of the bridge pickup sound into the neck pickup sound is shown as a solid line and the transformation of the neck pickup sound into the bridge pickup sound is shown as as dashed line. The electric guitar is plucked directly above the neck pickup on the open 1st string. The vertical dotted lines at 50, 100, 150, 300 and 450 are the filter lengths used in the listening test. Note that the y-axis is in log scale.

listeners cannot perceive significant differences between the estimated sounds, suggesting that filter length of 150 samples (just above f_s/f_0) is sufficient in this case. The ratings for the estimated sounds are quite close to the hidden references, which may suggest that the similarity between estimated sounds and their target sounds are close with some differences.

In conclusion, filter lengths with frequency resolution above f_s/f_0 are suitable for transforming the sound of a pickup into another, giving estimated sounds that are close to the target sounds.

3.4.2 Morphing Three Pickup Positions

In this section, a subjective evaluation is performed on the morphings of the three pickup positions using recordings of the electric guitar plucked moderately loud on the open 3rd string ($f_0 = 196$ Hz). Participants are presented with three questions relating to: (1) transforming the middle and bridge pickup signals into the neck pickup, (2) transforming the neck and bridge pickup signals into the middle pickup and (3) transforming the neck and

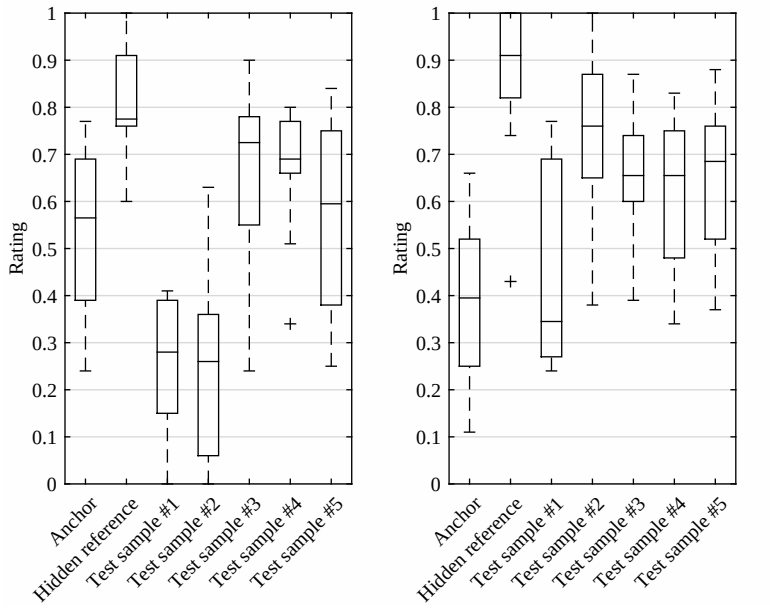


Figure 3.10: Box plots of the sound similarity ratings between the reference sound and test samples testing for finding suitable filter lengths. (a) The reference sound is the neck pickup signal, the anchor is the bridge pickup signal and test samples 1–5 are the estimated neck pickup signal with filter length 50, 100, 150, 300 and 450 respectively. (b) The reference sound is the bridge pickup signal, the anchor is the neck pickup signal and test samples 1–5 are the estimated bridge pickup signal with filter length 50, 100, 150, 300 and 450 respectively.

middle pickup signals into the bridge pickup signals. Since the filter length must be higher than f_s/f_0, the filter length is set to 450, which is around $(2f_s)/f_0$.

Transforming the Middle and Bridge Pickup Signals into the Neck Pickup

In this test, participants are asked to rate the sound similarity between the neck pickup signal (reference sound) and a hidden version of the reference sound, two hidden anchors and two test samples. The hidden anchors are the middle and bridge pickup signals and the two test samples are the estimated neck pickup signals transformed from the middle and bridge pickup signals.

Fig. 3.11a shows the box plots of the ratings. The anchor sounds are rated with low grades, where their median scores are below 0.5. As expected, the middle pickup sound is rated to be closer to the neck pickup sound than the bridge pickup, where the middle pickup suppresses more partials than the bridge pickup producing a warmer sound. Note that difference between the ratings for the middle and bridge pickups are statistically significant.

Fig. 3.11a also shows that the ratings for the estimated sounds are high, where most of the ratings are above 0.7, suggesting that the sounds are very close to the target sound. By performing the paired-sample t-test, the difference between the ratings for the estimated sounds and the hidden reference are not statistically significant, confirming that the estimated signals are as good as the hidden reference with no significant differences.

Transforming the Neck and Bridge Pickup Signals into the Middle Pickup

Participants are asked to rate the sound similarity between the middle pickup signal (reference sound) and a hidden version of the reference sound, two hidden anchors and two test samples. The hidden anchors are the neck and bridge pickup signals and the two test samples are the estimated middle pickup signals transformed from the neck and bridge pickup signals.

The ratings are shown as box plots in Fig. 3.11b. As expected, the anchor sounds are rated with low grades. Although the neck pickup is perceived to sound closer to the middle pickup than the bridge, the difference between the ratings for the neck and bridge pickup is not statistically significant.

The ratings for the estimated sounds are mixed as shown in Fig. 3.11b. The ratings for the estimated middle pickup signal transformed from the bridge pickup are high, where the median score is above 0.8. By comparing between the ratings for the estimated middle

pickup signal that is morphed from the bridge pickup signal and the ratings for the hidden reference, the differences are not statistically significant. This strongly suggests that the emulation is quite close to the target sound with no significant differences. Meanwhile, the ratings for the estimated middle pickup signal transformed from the neck pickup are collectively lower. Although the timbre (e.g. brightness) is close to the target sound, there is a slight perceivable difference in the attack part of the emulated signal and a slightly noisier background.

Transforming the Neck and Middle Pickup Signals into the Bridge Pickup

Participants are then asked to rate the sound similarity between the bridge pickup signal (reference sound) and a hidden version of the reference sound, two hidden anchors and two test samples. The hidden anchors are the neck and middle pickup signals and the two test samples are the estimated bridge pickup signals transformed from the neck and middle pickup signals.

The box plots of the ratings are shown in Fig 3.11c. The two anchors are rated with low grades, where their median scores are below 0.5. The middle pickup sound is perceived to be closer to the bridge pickup than the neck pickup. This is expected, where the middle pickup suppresses less partials than the neck pickup producing a brighter sound. The difference between the ratings for the middle and neck pickup is statistically significant.

The ratings for the estimated bridge pickup sounds are mostly rated with high grades, where their median scores are above 0.7. The ratings are not as high as the hidden reference ratings suggesting that the estimated bridge pickup sounds are very similar to the target sound but with perceivable differences.

3.4.3 Filter Robustness

In this section, the generality of learnt filters for each string is tested to assess the effect of different playing techniques such as variations in plucking dynamics and plucking positions. To measure the robustness of the filter, a filter for a particular input/target pair (the training pair) is extracted to test how well it performs given a different input/target pair (the testing pair). The training and testing pairs are different instances (repetitions) of the same parameters (string, fret, plucking dynamic, plucking position), but variations in the other parameters are tested one at a time (except for string). Therefore, the variables under test

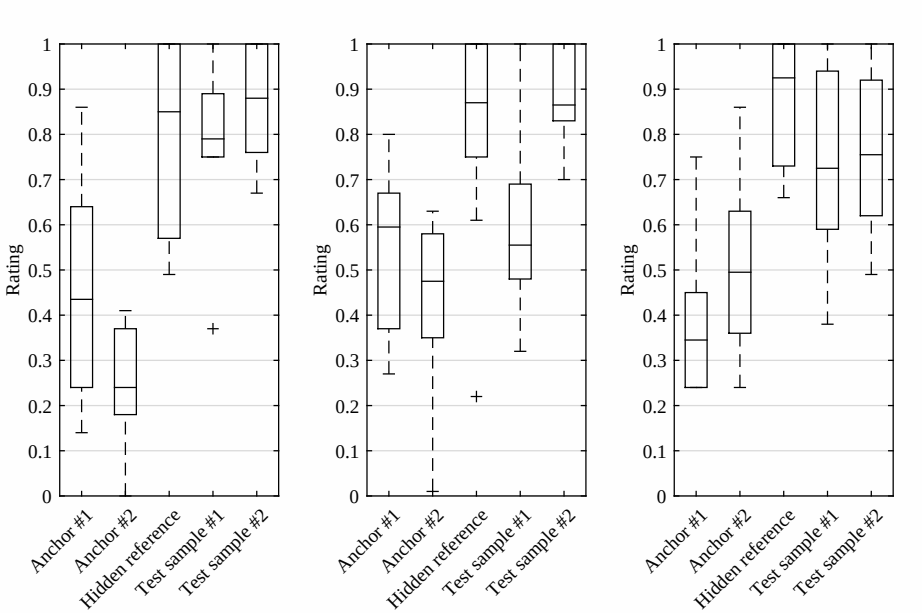


Figure 3.11: Box plots of the sound similarity ratings between the reference sound and test samples testing for morphing three pickup positions. (a) The reference sound is the neck pickup signal, the anchor 1 and 2 are the middle and bridge pickup signal respectively and test samples 1 and 2 are the estimated neck pickup signal transformed from the middle and bridge pickup respectively. (b) The reference sound is the middle pickup signal, the anchor 1 and 2 are the neck and bridge pickup signal respectively and test samples 1 and 2 are the estimated middle pickup signal transformed from the neck and bridge pickup respectively. (c) The reference sound is the bridge pickup signal, the anchor 1 and 2 are the neck and middle pickup signal respectively and test samples 1 and 2 are the estimated bridge pickup signal transformed from the neck and middle pickup respectively.

are *repetition*, *plucking dynamic*, *plucking position* and *fret position*.

The variable *repetition* is taken as an example to explain the method of analysing each variable. The process for analysing the filter robustness to different *repetitions* is as follows:

1. Three input/target pairs $x_i(n)$ and $y_i(n)$ are taken, where $i \in \{1, 2, 3\}$ is the index of the *repetition* and other variables remained constant.
2. The estimated filters $\hat{h}_i(n)$ are obtained as described in Section 3.2 for each *repetition*.
3. Each estimated filter $\hat{h}_j(n)$ is applied to the input signals $x_i(n)$ separately to produce the estimated signals $\hat{y}_{i,j}(n)$, where $j \in \{1, 2, 3\}$
4. Steps 1, 2 and 3 are repeated for each case (i.e. each combination of plucking dynamic, plucking position, fret position and string), which leads to a total of 162 cases.

The same process is used for analysing the filter robustness to variables *plucking position*, *plucking dynamic* and *fret position* by replacing the variable *repetition* with the variable under test.

Repetitions

Three questions are presented to the listeners to evaluate the accuracy of the estimated filters when applied to different repetitions. In these questions, recordings of the electric guitar plucked moderately loud on the open 4th string ($f_0 = 147$ Hz) directly above the middle pickup with 3 repetitions are used. For Question i, where $i \in \{1, 2, 3\}$, the hidden reference sound is the neck pickup signal of *repetition* i, the anchor sound is the bridge pickup signal of *repetition* i and test samples 1–3 are the estimated neck pickup signals of *repetition* 1–3 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ extracted from the three input/target pairs respectively. The filter length is set to 600, which is around $(2f_s)/f_0$. Listeners are asked to rate the sound similarity between those sound samples and the reference sound.

Fig. 3.12 shows the ratings given by the listeners. It can be seen that ratings for the estimated sounds are as high as the hidden reference, suggesting that the estimated sounds closely resemble their target sounds. By performing the paired-sample t-test, the differences between the estimated sounds and their hidden references are not statistically significant, except for Test Sample # 2 in Fig. 3.12c but its ratings are still high with median score above 0.7. Almost all of the estimated sounds are as good as the hidden references, where

no significant difference is perceived, meaning that the filter is very robust to changes in repetitions.

Plucking Dynamics

The previous test is repeated, where the variable *repetition* is changed to *plucking dynamic*. Recordings of the electric guitar plucked on the open 5th string ($f_0 = 110$ Hz) directly above the neck pickup with a forte (*plucking dynamic* 1), mezzo-forte (*plucking dynamic* 2) and piano (*plucking dynamic* 3) pluck are used in this test. The filter length is set to 802, which is around $(2f_s)/f_0$.

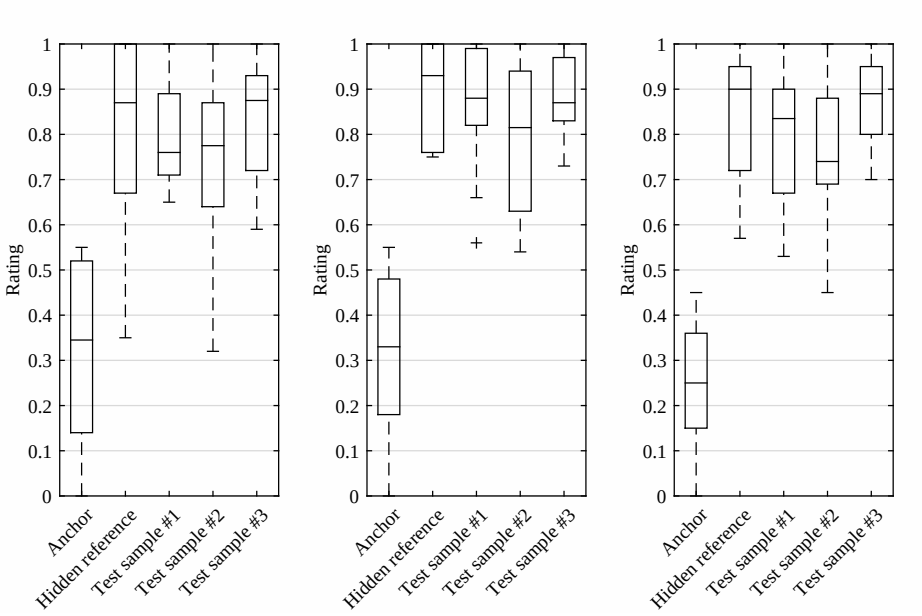
Fig. 3.13 shows the ratings given by the participants. The estimated sounds of a forte pluck using filters extracted from mezzo-forte and piano plucks are perceived to be less similar than the hidden reference (see Fig. 3.13a). This might be due to the lack of information for the filter optimisation to learn at high frequencies, where a softer pluck produces a low-pass filtering effect. Nevertheless, their median scores are above 0.5 and higher than the anchor sound. This suggests that the filters are less robust to forte plucks.

On the other hand, there are very small perceivable differences between the estimated sounds of mezzo-forte and piano plucks and their hidden references (see Figs. 3.13b and 3.13c). By performing the paired-sample t-test, the difference between the ratings for the estimated sounds and their hidden references are not statistically significant, meaning that the estimated sounds are as good as the hidden references. This suggests that the filters are robust to mezzo-forte and piano plucks.

Plucking Positions

The previous test is repeated, where the variable *plucking dynamic* is changed to *plucking position*. Recordings of the electric guitar plucked moderately loud on the open 2nd string ($f_0 = 247$ Hz) directly above the bridge pickup (*plucking position* 1), middle pickup (*plucking position* 2) and neck pickup (*plucking position* 3) are used in this test. The filter length is set to 358, which is around $(2f_s)/f_0$.

The ratings given by the participants are shown in Fig. 3.14. It can be seen that the ratings for all of the estimated sounds are high, where the median score is above 0.7. By performing the paired-sample t-test, the differences between the estimated sounds and the hidden references are not statistically significant, except for Test sample # 3 and Test sample # 1 in Figs. 3.14b and 3.14c respectively. Although the ratings for those test samples are



(a) (b) (c)

Figure 3.12: Box plots of the sound similarity ratings between the reference sound and test samples testing for the effects of repetition. (a) The reference sound is the neck pickup signal of *repetition* 1, the anchor is the bridge pickup signal of *repetition* 1 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *repetition* 1 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively. (b) The reference sound is the neck pickup signal of *repetition* 2, the anchor is the bridge pickup signal of *repetition* 2 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *repetition* 2 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively. (c) The reference sound is the neck pickup signal of *repetition* 3, the anchor is the bridge pickup signal of *repetition* 3 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *repetition* 3 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively.

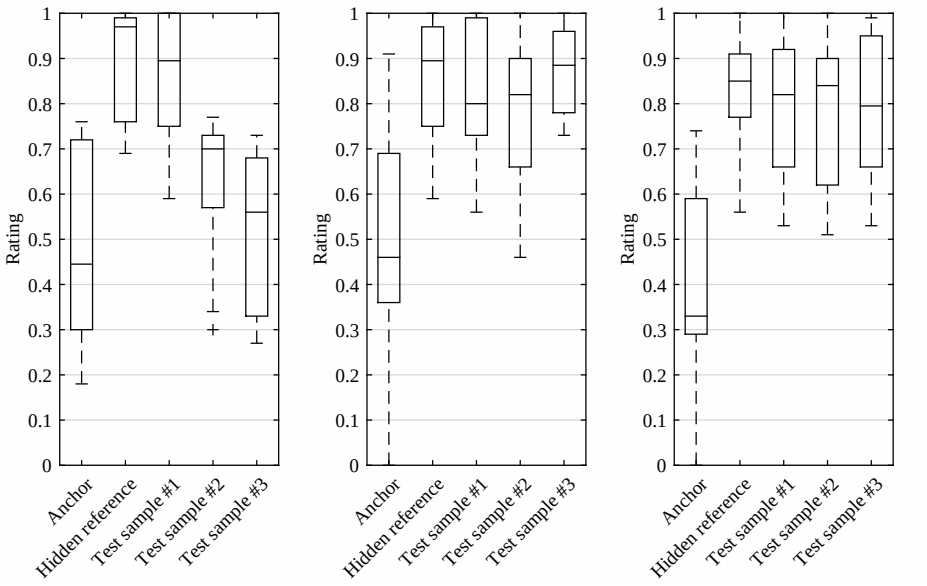


Figure 3.13: Box plots of the sound similarity ratings between the reference sound and test samples testing for the effects of plucking dynamic. (a) The reference sound is the neck pickup signal of *plucking dynamic* 1, the anchor is the bridge pickup signal of *plucking dynamic* 1 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *plucking dynamic* 1 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively. (b) The reference sound is the neck pickup signal of *plucking dynamic* 2, the anchor is the bridge pickup signal of *plucking dynamic* 2 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *plucking dynamic* 2 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively. (c) The reference sound is the neck pickup signal of *plucking dynamic* 3, the anchor is the bridge pickup signal of *plucking dynamic* 3 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *plucking dynamic* 3 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively.

not similar to the hidden references, their median scores are quite high, suggesting that the listeners perceive some small differences. This means that the filters are quite robust to the changes in plucking positions, where slight perceivable differences are audible in some cases.

Fret Positions

The previous test is repeated, where the variable *fret position* is now under test. Recordings of the electric guitar plucked moderately loud on the open 3rd string directly above the neck pickup with 3 fret positions: open fret (*fret position* 1), fifth fret (*fret position* 2) and twelfth fret (*fret position* 3) are used in this test. The filter length is set to 450, which has a sufficient frequency resolution for the three pitches.

The ratings given by the listeners are shown in Fig. 3.15. As expected, the estimated sounds of *fret position* i only works well with filters learnt from input/target pairs of *fret position* i, where their ratings are higher than other estimated sounds and very close to the hidden references. This means that the filter is not robust to changes in fret positions.

Fret Positions: Improvements

From the previous test, the filters can be improved to be more robust towards changes in fret positions. The input (respectively target) signals played on the open string, at the fifth fret and at the twelfth fret are concatenated, in order to learn a composite filter.

Participants are given the neck pickup signal played on the open string as the reference sound and the bridge pickup signal played on the open string as the anchor sound. The test samples are the estimated neck pickup signals using 5 different filters. The filter $\hat{h}_1(n)$ is learnt from the open string only, $\hat{h}_2(n)$ is learnt from the open string and fifth fret signal pairs, $\hat{h}_3(n)$ is learnt from the fifth and twelfth fret pairs, $\hat{h}_4(n)$ is learnt from the open string and twelfth fret pairs and filter $\hat{h}_5(n)$ is learnt from all three fret positions.

Fig. 3.16 shows the ratings given by the participants. It can be seen that the filters learnt from the open string are given high scores, where their median scores are above 0.6, and the filter that does not learn from the open string ($\hat{h}_3(n)$) is rated with low scores, where its median score is below 0.4.

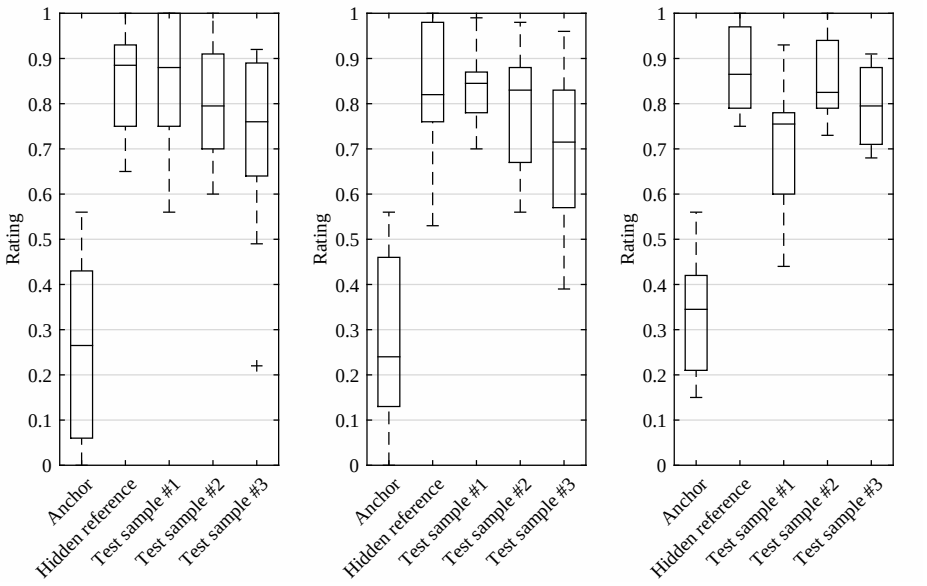


Figure 3.14: Box plots of the sound similarity ratings between the reference sound and test samples testing for the effects of pickup position. (a) The reference sound is the neck pickup signal of *plucking position* 1, the anchor is the bridge pickup signal of *plucking position* 1 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *plucking position* 1 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively. (b) The reference sound is the neck pickup signal of *plucking position* 2, the anchor is the bridge pickup signal of *plucking position* 2 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *plucking position* 2 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively. (c) The reference sound is the neck pickup signal of *plucking position* 3, the anchor is the bridge pickup signal of *plucking position* 3 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *plucking position* 3 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively.

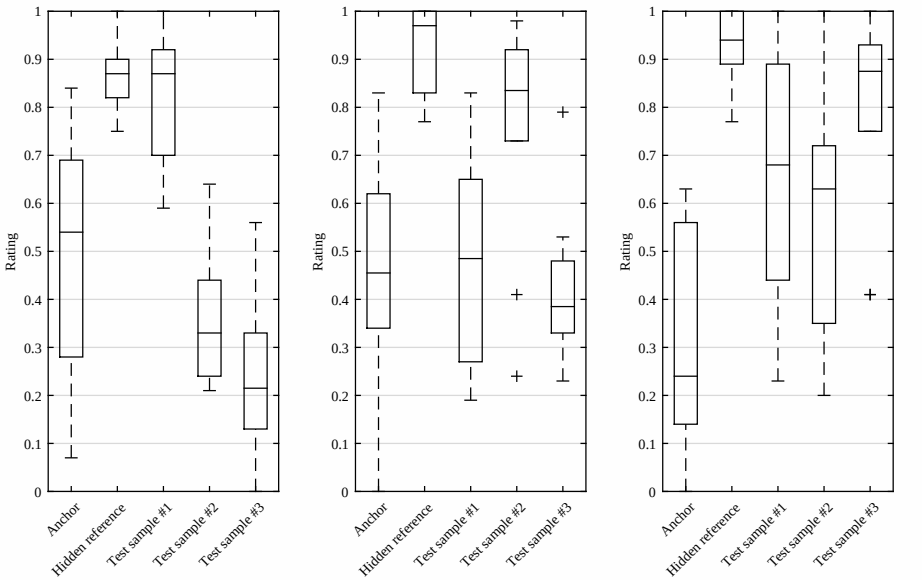


Figure 3.15: Box plots of the sound similarity ratings between the reference sound and test samples testing for the effects of fret position. (a) The reference sound is the neck pickup signal of *fret position* 1, the anchor is the bridge pickup signal of *fret position* 1 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *fret position* 1 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively. (b) The reference sound is the neck pickup signal of *fret position* 2, the anchor is the bridge pickup signal of *fret position* 2 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *fret position* 2 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively. (c) The reference sound is the neck pickup signal of *fret position* 3, the anchor is the bridge pickup signal of *fret position* 3 respectively and test samples 1, 2 and 3 are the estimated neck pickup signal of *fret position* 3 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ respectively.

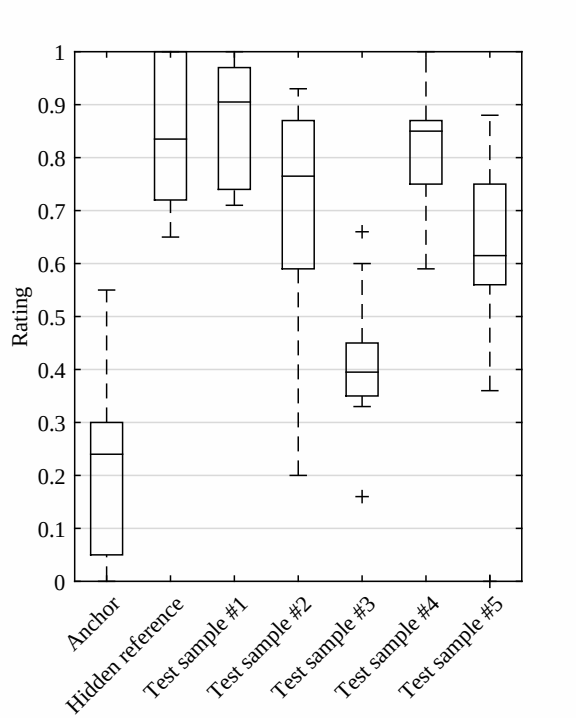


Figure 3.16: Box plots of the sound similarity ratings between the reference sound and test samples improving the effects of fret position. The reference sound is the neck pickup signal of *fret position* 1, the anchor is the bridge pickup signal of *fret position* 1 respectively and test samples 1, 2, 3, 4 and 5 are the estimated neck pickup signal of *fret position* 1 using filters $\hat{h}_1(n)$, $\hat{h}_2(n)$, $\hat{h}_3(n)$, $\hat{h}_4(n)$ and $\hat{h}_5(n)$ respectively.

3.5 Numerical Results

3.5.1 Filter Robustness

Nine distances (or errors) for each of 162 cases are obtained to analyse each variable (*repetition*, *plucking dynamic*, *plucking position* and *fret position*). Each of the nine distances is then averaged over all 162 cases. The average errors for analysing the filter robustness when applied to an input/target pair with different *repetition*, *plucking position*, *plucking dynamic* and *fret position* are shown in Tables 3.2a, 3.2b, 3.2c and 3.2d respectively. The values of the variable under test are indexed by i for the learnt filters $\hat{h}_i(n)$. As an example, in Table 3.2d, $\hat{h}_1(n)$, $\hat{h}_2(n)$ and $\hat{h}_3(n)$ are extracted from input/target pairs that are played on the open fret $\mathcal{F}_0$, fifth fret $\mathcal{F}_5$ and twelfth fret $\mathcal{F}_{12}$. The filters are then evaluated on input/target pairs from each of the three fret positions. The diagonal values in bold emphasise the cases where the training and testing pairs coincide. Therefore, these can be used as reference values, in order to compare the loss in accuracy due to changes in the variable under test.

In Table 3.2a, the error increases when the filters are applied to different *repetitions*, which ranges from 16% to 39%. The filters are expected to be reasonably robust towards other *repetitions* because the notes are plucked at similar positions, dynamics, frets and strings. It is worth noting that a mechanical plucking device was not used in this experiment, thus, there are slight random variations in the parameters of the pluck (i.e. strength, position and angle) between repetitions, but there should be no systematic variation. Thus, all non-bold values (off-diagonal values) are shown to be quite consistent because different *repetitions* have similar timbre.

In Table 3.2b, the error increases by a factor of 2 to 5 when a filter is applied to a different *plucking position*. The learnt filter $\hat{h}_2(n)$ has lower off-diagonal errors compared to filters $\hat{h}_1(n)$ and $\hat{h}_3(n)$. Similarly, input/target pair 2 has lower off-diagonal errors than the other pairs. The effect of the middle plucking point, like the position itself, appears to be closer to the other plucking points than they are to each other.

Table 3.2c shows that the error increases when the learnt filter is applied to a signal with different *plucking dynamics*. It appears that changing the *plucking dynamics* increases the error by approximately 2 folds, although for filter $\hat{h}_3(n)$ learnt from a softer pluck *p'*, a much larger error is observed. This could be because there could be a lack of information for the filter optimisation to learn at high frequencies, which is due to the low-pass filtering effect of softer plucks.

The largest errors are recorded in Table 3.2d when the filters are applied to signals with different *fret positions*. There are two reasons for the large errors in the result: first, the filter learns accurately only at or near partial frequencies of the training tone, therefore, for an input signal with different fret positions, the partial frequencies are different and the filter cannot provide an accurate frequency response at those partials. The second reason is that different fret positions produce different comb filtering effects of the pickup position. For example, the neck pickup is located approximately $\frac{1}{4}$ of the way along an open string, but $\frac{1}{3}$ of the way between the fifth fret and the bridge, and $\frac{1}{2}$ way between the twelfth fret and the bridge. Thus, the comb filters attenuate every 4th, 3rd or 2nd partial in the respective cases.

The results in Table 3.2d can be improved. In order to learn a composite filter, the input (respectively target) signals played on the open string, at the fifth fret and at the twelfth fret are concatenated. The filter $\hat{h}_1(n)$ is learnt from the open string and fifth fret signal pairs, $\hat{h}_2(n)$ is learnt from the open string and twelfth fret pairs, $\hat{h}_3(n)$ is learnt from the fifth and twelfth fret pairs and filter $\hat{h}_4(n)$ is learnt from all three fret positions. The accuracy of the filters is then evaluated on input/target pairs for the three fret positions which are shown in Table 3.2e. A considerable improvement is shown compared to the errors measured in Table 3.2d. As expected, the filter with the least error is $\hat{h}_4(n)$ which uses information from all input/target pairs.

3.5.2 Comparisons Between Variables

Table 3.3 provides a summary of the results in Table 3.2, where the second and third columns show the averages of the diagonal and off-diagonal values respectively for Tables 3.2a–3.2d. The third column also shows the average error for the improved composite filter (averages for the fifth column in Table 3.2(e)). The results in Table 3.3 give insight into the relative contributions of the variables, and thus the robustness of the filters towards changes in *repetition*, *plucking position*, *plucking dynamic* and *fret position*. It is shown that the filters are most robust to changes in *repetition*, which should have no systematic difference. The changes in *plucking positions* result in a 3 fold increase in error, while the changes in *plucking dynamics* result in a 2 fold increase in error. The filters are least robust to changes in *fret position*, where the error is increased 5 fold. By using the learnt composite filters, the errors increase by only about 30%.

Input signal	$\hat{h}_1(n)$ Rep. 1	$\hat{h}_2(n)$ Rep. 2	$\hat{h}_3(n)$ Rep. 3
Rep. 1	**0.186**	0.240	0.260
Rep. 2	0.216	**0.184**	0.238
Rep. 3	0.217	0.225	**0.187**

(a)

Input signal	$\hat{h}_1(n)$ ρ_1	$\hat{h}_2(n)$ ρ_2	$\hat{h}_3(n)$ ρ_3
ρ_1	**0.154**	0.391	0.705
ρ_2	0.597	**0.222**	0.366
ρ_3	0.836	0.453	**0.181**

(b)

Input signal	$\hat{h}_1(n)$ *f'*	$\hat{h}_2(n)$ *mf*	$\hat{h}_3(n)$ *p'*
f'	**0.195**	0.364	0.535
mf	0.358	**0.211**	0.287
p'	0.352	0.258	**0.152**

(c)

Input signal	$\hat{h}_1(n)$ $\mathcal{F}_0$	$\hat{h}_2(n)$ $\mathcal{F}_5$	$\hat{h}_3(n)$ $\mathcal{F}_{12}$
$\mathcal{F}_0$	**0.135**	0.705	0.591
$\mathcal{F}_5$	1.080	**0.111**	0.505
$\mathcal{F}_{12}$	0.773	1.949	**0.310**

(d)

Input Signal	$\hat{h}_1(n)$ $\mathcal{F}_0 + \mathcal{F}_5$	$\hat{h}_2(n)$ $\mathcal{F}_0 + \mathcal{F}_{12}$	$\hat{h}_3(n)$ $\mathcal{F}_5 + \mathcal{F}_{12}$	$\hat{h}_4(n)$ $\mathcal{F}_0 + \mathcal{F}_5 + \mathcal{F}_{12}$
$\mathcal{F}_0$	**0.151**	**0.241**	0.432	**0.242**
$\mathcal{F}_5$	**0.121**	0.495	**0.152**	**0.152**
$\mathcal{F}_{12}$	0.644	**0.311**	**0.333**	**0.325**

(e)

Table 3.2: Errors measured for filters applied to an input/target pair with different (a) repetition, (b) plucking position, (c) plucking dynamic, (d) fret position and (e) fret position (improved filter). The plucking positions ρ_1, ρ_2 and ρ_3 are located directly above the neck, middle and bridge pickup respectively. The plucking dynamics are *forte* (*f'*), *mezzo-forte* (*mf*) and *piano* (*p'*). The fret positions are the open string $\mathcal{F}_0$, fifth fret $\mathcal{F}_5$ and twelfth fret $\mathcal{F}_{12}$.

Variables	Average Diagonal	Average Off-diagonal	Improved
Repetitions	0.186	0.233	
Plucking positions	0.186	0.558	
Plucking dynamics	0.186	0.359	
Fret positions	0.186	0.934	0.240

Table 3.3: Summary table for comparisons between each variable.

3.6 Summary

In this chapter, a method is proposed for transforming the sound of an arbitrary pickup into that of another pickup on the same guitar using an FIR filter. The filter coefficients are obtained using a least square estimator. This technique is introduced as a preliminary step towards morphing the sound of a user's guitar into a desired electric guitar sound in an audio recording.

Formal listening tests suggest that the technique yields good results. By performing paired-sample t-test, most estimated sounds are as good as the hidden reference sounds, meaning that the participants perceived no significant differences between the estimated sounds and the hidden references. Moreover, these estimated sounds are given high grades, suggesting that the estimated sounds are very similar to the reference sound. The effectiveness of this technique is also supported by a spectral distance measure which shows that up 99% of the spectral difference between the input and target signal is reduced. The listening test and numerical results conclude that filter lengths above f_s/f_0 are suitable for morphing the sound of a pickup into another.

However, by using this approach, it is suggested that there are some limitations to replicate the target signal in the practical use case, such as a guitar synthesiser with hexaphonic pickups. First of all, the filter optimisation is biased to give low errors at or near partial frequencies. This means that the input signal must be played at the same pitch (string and fret position) and playing technique (i.e. plucking position, dynamic and angle) as the target signal, so that both the input and target partial frequencies align. This also means that the inharmonicity of the string must be the same, otherwise, the accuracy of the emulation is

affected.

In guitar synthesisers, once the filter is obtained, each string is processed by an individual filter to replicate the target sound. The robustness of the filters is evaluated via listening test and measuring the spectral differences between estimated sounds and target sounds to determine whether a player has the flexibility to play with a different playing technique (i.e. plucking position or dynamic) and fret position. A small degradation is observed due to random differences between repetitions. The filter is somewhat less robust when applied to an input signal with different plucking position or dynamic. The filter is not at all robust when applied to an input signal with different fret positions. The results for different fret positions are improved by learning the filter using multiple tones played on different frets along the string, which allows the filter to "fill the gaps" of unknown values in the frequency response between partials of a single training tone. This method could also be applied to improve performance across different values of other variables.

Chapter 4

Electric Guitar Parameter Estimation: Pickup and Plucking Positions

In the previous chapter, a technique for transforming an arbitrarily selected pickup sound into another on the same guitar using an FIR filter is proposed. This is a preliminary step to investigate if it is possible to morph the sound of an arbitrary guitar into a desired guitar sound in an audio recording using digital filters. This approach has some limitations, whereby the user does not have much flexibility to change plucking position and dynamic, where the filter only works best on the same plucking position and dynamic as the target sound. Also, the filter works best on the notes (string and fret position) that are played in the target signal, which limits the user to play only those notes.

Another approach of replicating a desired guitar sound is to extract the many parameters for electric guitar synthesis, such as the pickup and plucking positions of the electric guitar, just to name a few. The electric guitar synthesis could provide the user the ability to change playing techniques and notes. This chapter discusses three frequency domain approaches to estimate the locations of the pickup and pluck of an electric guitar tone, which are based on published literature (Mohamad et al., 2017a,b,c). The Spectral Peaks (SP) method finds the parameters (pickup and plucking positions) that best fit the electric guitar model by minimising the difference between the spectral peaks of the first period of the tone and the model. The second and third methods find the pickup and plucking locations based on the autocorrelation of the spectral peaks. The Autocorrelation of Spectral Peaks (AC-SP) method searches for the minimum mean squared error between the autocorrelation of the observed data and the model to yield the model parameters. The Log-correlation of Spectral Peaks (LC-SP) method uses trough detection to find the two minima of the log-correlation of the observed data to yield the pickup and plucking position estimates.

This chapter presents a comparison between these three methods. The effects of changing plucking dynamics and fret positions on the accuracy of the estimates are discussed in this chapter. Furthermore, some real-world applications of estimating the pickup and plucking position are discussed in this chapter, which show that there are several applications other than using the estimates as parameters for electric guitar synthesis.

4.1 Dataset II

4.1.1 Audio Samples

The first dataset contained recordings from only three different plucking positions. In order to test the accuracy of the system, another dataset was recorded containing guitar tones

from the Squier Stratocaster described in Section 3.1.1 with more plucking points. The same string gauges and material as mentioned in 3.1.1 were used. Moreover, this dataset includes recordings of the guitar's mixed pickup configuration signals. This will later be referred to as Dataset II.

This dataset consists of recordings of electric guitar tones played moderately loudly at 8 plucking points on each of the 6 open strings, using 5 different pickup selections (3 single and 2 mixed). The locations of the targeted plucking points are marked on the string with a felt-tip permanent marker pen as described by Penttinen and Välimäki (2004). Some measuring error of about ±2 mm is expected due to inaccuracies in string length measurements and the actual locations of the pickup and plucking events (Penttinen and Välimäki, 2004). Note that the mixed pickup selections are recorded on a separate occasion, as opposed to the single pickup selections which can be simultaneously recorded. All samples were recorded at 44100 Hz sampling rate.

The plucking positions range from 30 mm to 170 mm from the bridge with 20 mm intervals and the strings are plucked using a 0.88 mm thick plastic plectrum (same brand/model as the plectrum mentioned in Section 3.1.2). The two in-phase mixed pickups are a mix between neck and middle pickups, $\mathrm{PU_{N+M}}$ and a mix between middle and bridge pickups, $\mathrm{PU_{M+B}}$.

4.1.2 Variations in Plucking Events

While the pickup positions remain fixed at a point, the plucking positions may vary due to natural plucking. It is important to measure these variations since the accuracy of the pickup position and plucking point estimation is later evaluated.

A high speed camera is used to capture plucking events at 240 frames per second. The camera is placed parallel to the strings reducing parallax error when measuring the distance between the plucking positions and their target positions. The electric guitar is plucked at 3 different locations (on top of the bridge, middle and neck pickups) and 3 strings (1st, 3rd and 6th strings) with 10 repetitions.

Fig. 4.1 shows excerpts from the slow-motion video of the plucking events captured using the high speed camera. It can be seen that the plucking positions vary for each repetition. By measuring the distance between each plucking position and its target, the distance can lead up to ±1 mm, which is still small.

Furthermore, the accuracy of the plucking events are also tested with chords. A G major chord is played and all six strings are plucked at the three target plucking positions with

10 repetitions. Excerpts of the plucking events from the 6th to the 1st string are shown in Fig 4.2, where the distances between the plucking positions and their targets are very small. By analysing each pluck, the deviations are higher than single plucks, but still small, which can lead up to ±2.5 mm error.

4.2 Methods

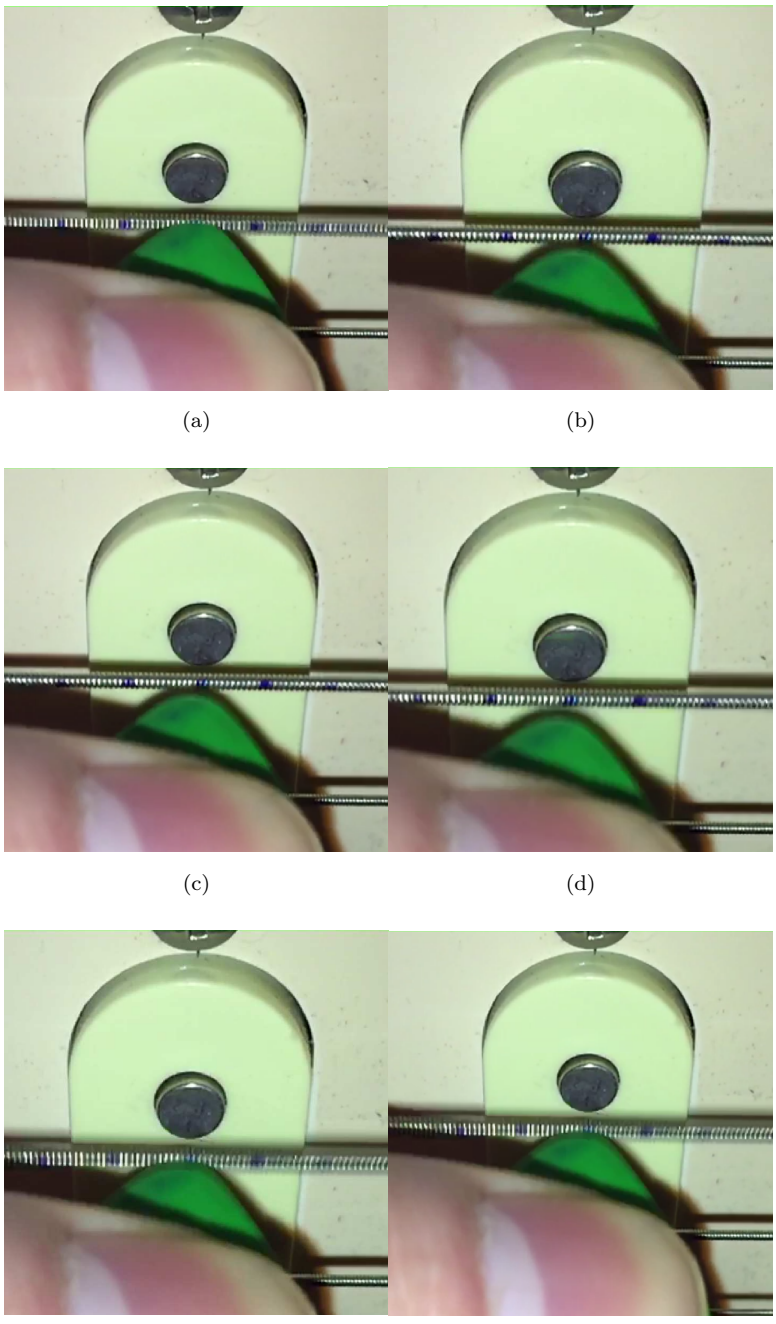
4.2.1 Spectral Peaks Method (SP)

The overview of the first method is shown in Fig. 4.3. It consists of three main steps, described below. Firstly, the first period of the electric guitar tone is extracted. Secondly, the Fourier series coefficients are then computed. Finally, a grid search is performed to find the pickup position $\hat{d}$ and the plucking point $\hat{\rho}$ estimates that minimise the error between the Fourier coefficients and the electric guitar model.

Onset Detection and Retrieving the First Period of the Tone

The first period of the electric guitar tone needs to be extracted to analyse the tone in the frequency domain using Fourier series. The spectral flux of the signal is calculated using Eq. (2.26) to estimate the onset time (see Section 2.2.2 for details). A window size of length 23 ms with overlapping windows of 50% is used to calculate the spectral flux, where in this case, the highest peak in the spectral flux gives an initial estimate of the onset time. This initial estimated onset time usually comes before the plucking noise due to the window overlap, therefore, it is necessary to refine the estimate to be closer to the plucking event. A window of size 92 ms is taken to refine the onset time. The first peak of the signal is detected such that $x(n-1) < x(n) > x(n+1)$, where peaks which are less than 20% of the maximum value in the window are discarded in order to avoid unwanted small peaks at the beginning of the tone due to the plucking noise. The last zero crossing before the first peak of the signal determines the start of the plucking event. An example is shown in Fig. 4.4 where the electric guitar is plucked on the open 4th string at 90 mm from the bridge and the pickup is 159 mm from the bridge. The signal shown in Fig. 4.4 starts from the initial estimate of the onset time at 1.570 s and the first dashed vertical line represents the refined onset time at 1.588 s.

After refining the onset time, the fundamental frequency f_0 is estimated using the YIN

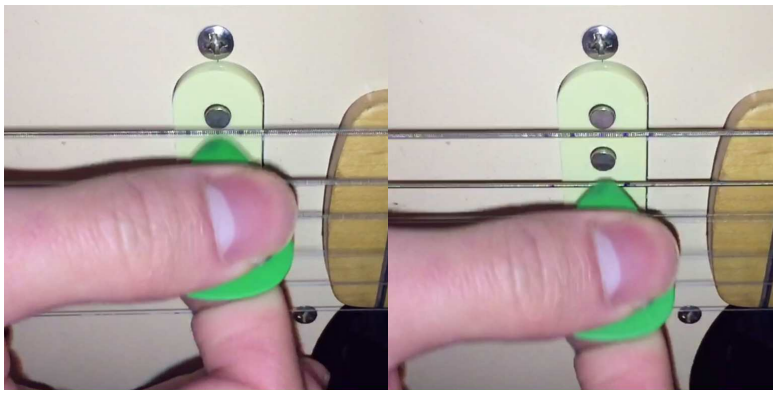


(a) (b)

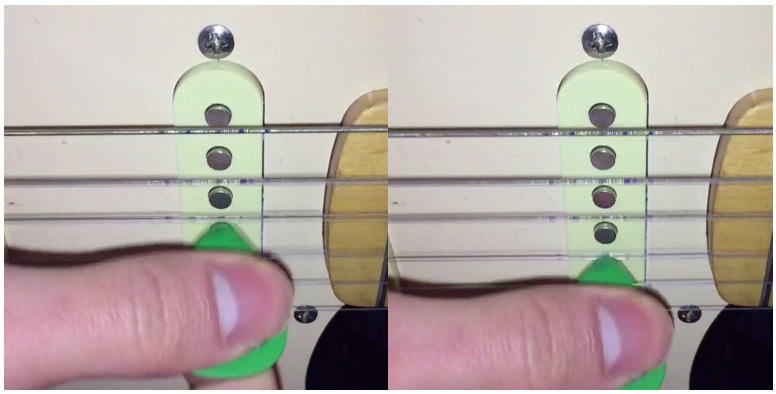
(c) (d)

(e) (f)

Figure 4.1: Excerpts from the slow-motion video of the plucking events on the 6th string captured using a high speed camera. The target plucking position is on top of the neck pickup, and it is marked with a blue marker pen. Locations around the target plucking positions (±5 mm and ±10 mm) are also marked with a blue marker pen.



(a) (b)



(c) (d)

(e) (f)

Figure 4.2: Excerpts from the slow-motion video of the plucking events from the (a) 6th to (f) 1st string when playing a G major chord captured using a high speed camera. The target plucking position is on top of the neck pickup, and it is marked with a blue marker pen. Locations around the target plucking positions ($\pm$5 mm and $\pm$10 mm) are also marked with a blue marker pen.

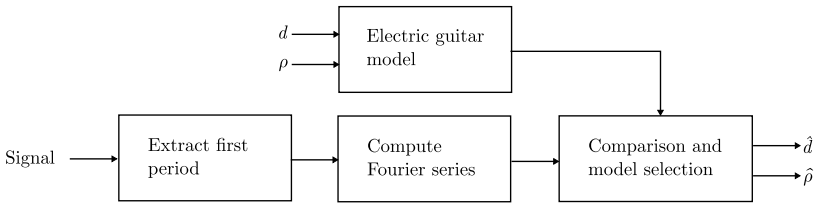


Figure 4.3: Block diagram for estimating pickup position and plucking point of an electric guitar using the Spectral Peaks method (SP).

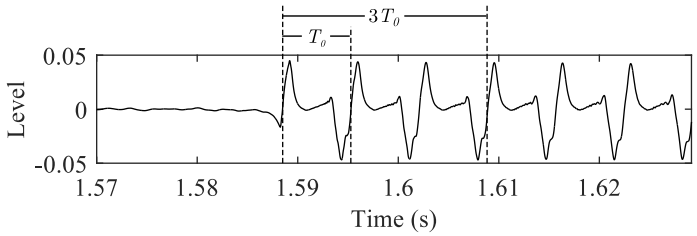


Figure 4.4: An excerpt of an electric guitar tone which starts from the initial estimate of the onset time (1.57s). The electric guitar is plucked on the open 4th string 90 mm from the bridge and the neck pickup $\mathrm{PU_N}$ is 159 mm from the bridge. For the SP method, the analysed window of length T_0 starts from the refined onset time (first dashed line from the left) to the end of the first period (second dashed line). The analysed window for the AC-SP and LC-SP methods is shown as an example for a window size of $3T_0$.

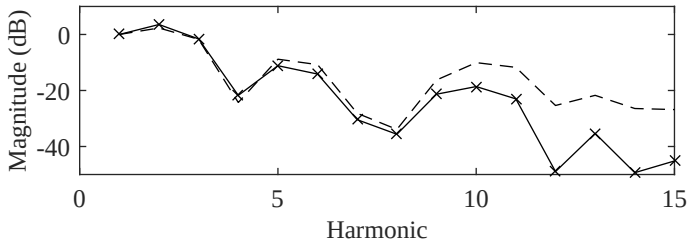


Figure 4.5: The Fourier series coefficients of the first period of the tone in Fig. 4.4 (crosses) and the model (dashed line).

algorithm in Eq. (2.24), where the window size is set to 46 ms. Finally, the first period of the tone can be extracted, where the signal starts at the refined onset time and its window size is T_0 seconds ($T_0 = \frac{1}{f_0}$). An example is shown in Fig. 4.4 where the region between the first and second dashed lines represents the first period.

Fourier Series Coefficients of the Tone and Minimising the Error

Once the first period of the tone is retrieved, the Fourier series coefficients are computed to estimate its spectral peaks, X_k. Setting the total number of harmonics K will be discussed in Section 4.2.4. Fig. 4.5 shows the Fourier series coefficients of the first cycle of the tone in Fig. 4.4 compared with the electric guitar model calculated using Eq. (2.11). For the first 9 harmonics, the observed data and the model are nearly identical. The higher harmonics of the model have higher amplitudes compared to the observed data because the model does not account for the low-pass filtering effect of the pickup width, plucking dynamics and plectrum width.

The grid search minimises the mean squared error between the spectral peaks of the observed data and electric guitar model (Eq. (2.11) for single pickup data and Eq. (2.12) for in-phase mixed pickup data) for plucking points and pickup positions ranging from 25 mm to 180 mm with a spatial resolution of 1 mm, where the string length L is known. Thus, the pair giving the minimum mean squared error corresponds to the estimated plucking point $\hat{\rho}$ and pickup position $\hat{d}$.

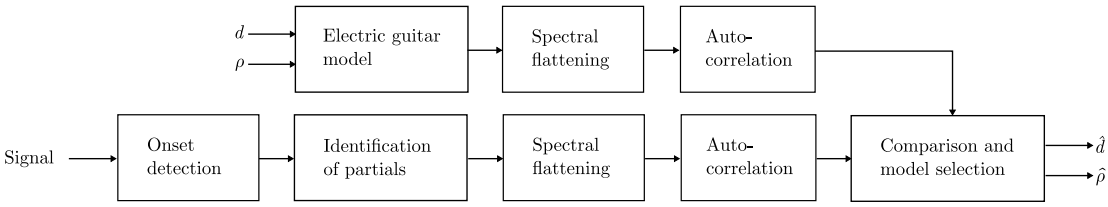


Figure 4.6: Block diagram for estimating pickup position and plucking point of an electric guitar using the Autocorrelation of Spectral Peaks method (AC-SP)

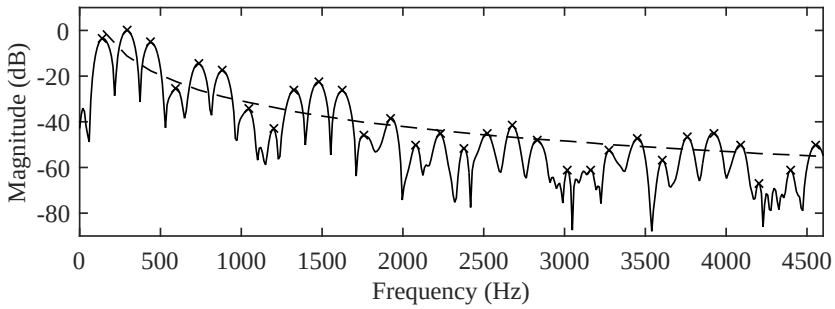


Figure 4.7: The spectrum of the first 3 cycles of the electric guitar tone in Fig. 4.4 (solid line), the magnitudes of each of its partials in decibels (crosses) and the slope of the spectrum estimated using linear regression (dashed line). Note that the linear regression that is performed to find the slope of the spectrum deals with the log-frequency domain, see Eq. (4.2).

4.2.2 Autocorrelation of Spectral Peaks Method (AC-SP)

The overview of the second method is shown in Fig. 4.6. The onset time is first detected and a window of three periods is taken for STFT analysis. The algorithm for detecting the spectral peaks differs from the SP method, and is discussed in this section. The spectral peaks are flattened before calculating the autocorrelation. The grid search algorithm to find the estimates $\hat{\rho}$ and $\hat{d}$ is similar to the SP method, but the point giving minimum error between the autocorrelation of the observed data and the electric guitar model is found instead to yield the estimates.

Onset Time Estimation and Spectral Peak Extraction

The start time of the electric guitar tone and its fundamental frequency f_0 are estimated using the method discussed in Section 4.2.1, but the window size is set to be longer than T_0

seconds, so that there is sufficient frequency resolution for the STFT analysis. An example is given in Fig. 4.4 where a window size of $3T_0$ (the region between the first and third dashed lines) is taken for analysis. Only the first few cycles are taken, in order to capture the initial conditions of the pluck before information is lost due to decaying harmonics. Furthermore, the window size should be as small as possible so that time modulation effects such as reverberation and delay will not be prevalent. The STFT analysis is performed on the signal using a Hamming window and zero padding factor of 4; an example spectrum is shown in Fig. 4.7. Each spectral peak is searched by taking its maximum value in windows of ± 30 cents around expected partial frequencies f_k (see Eq.(2.7)) using empirical measurements of inharmonicity coefficient B for each string provided by Barbancho et al. (2012). The magnitudes and the frequencies of the spectral peaks are then refined using quadratic interpolation (Smith, 2011).

A more precise estimate of the inharmonicity of the tone can be obtained from each pair of estimated partial frequencies. Given the frequencies f_i and f_j of any two partials i and j, Eq. 2.7 is rearranged to obtain an estimate of B for each pair (Dixon et al., 2012):

$$\hat{B}_{i,j} = \frac{i^2 f_j^2 - j^2 f_i^2}{j^4 f_i^2 - i^4 f_j^2} \tag{4.1}$$

Some spectral peaks might not be correctly detected because the initial inharmonicity coefficient that is set may be more or less than the actual inharmonicity coefficient of the tone. Therefore, a robust measurement that discards outliers is needed to obtain the final estimate of the inharmonicity. The median is a reasonable robust measure of central tendency in the presence of some measurement errors, unlike the mean which is prone to bias from noise. Thus, the median of all $\hat{B}_{i,j}$ values is taken as the estimated inharmonicity coefficient $\hat{B}$. Fig. 4.7 shows the detected spectral peaks represented by crosses. For this example, the estimated inharmonicity for the open 4th string is 7.4×10^{-5} which is taken from the median of Eq. (4.1) using estimated partial frequencies in Fig. 4.7.

Some of the falsely detected spectral peaks can be corrected using the estimated inharmonicity $\hat{B}$ of the tone. A threshold is set to identify any falsely estimated partial frequencies. The target frequencies are calculated using Eq. (2.7) with coefficient $\hat{B}$ estimated earlier. Then, the estimated partial frequencies are identified as false if any of them deviate more than ± 30 cents from their target frequencies. The corrected spectral peak is found in the revised window, and further refined using quadratic interpolation. The corrected partial frequency is set equal to its target frequency if no peak is detected in the window.

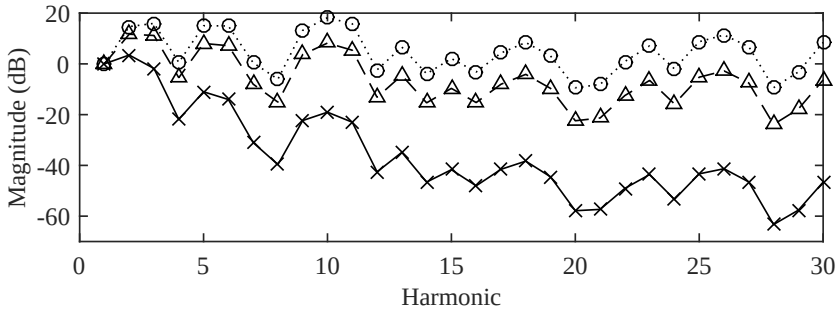


Figure 4.8: The spectral peaks of Fig. 4.7 (crosses) adjusted to have slopes of -3 (triangles) and 0 dB/octave (circles).

Linear Regression Spectral Flattening (LRSF)

The spectrum needs to be flattened in order to compensate for the energy losses due to nonrigid end supports (e.g. bridge and fingers), and the low-pass filtering effect due to pickup width, plectrum width and plucking dynamics. Flattening the spectrum could reverse these effects by amplifying the level of higher partials.

The slope of the spectrum is estimated by fitting a line relating log-magnitude to log-frequency, where the best fitting line can be written as:

$$\log(X_k) = \phi \log(k) \tag{4.2}$$

where the spectral peak X_k for harmonic k is normalised to a maximum of 0 dB. The parameter ϕ is estimated using least squares solution. Thus, the variable power of the harmonics gives the slope of the spectrum where k^{ϕ} has a 6ϕ dB/octave slope. Once the parameter ϕ_{dB} in decibels per octave is obtained, the slope of the spectrum can be adjusted to a target slope φ_{dB}. The flattened spectrum is written as:

$$\bar{X}_k = \frac{X_k}{k^{\Delta}} \tag{4.3}$$

where $\Delta = \frac{\phi_{\mathrm{dB}} - \varphi_{\mathrm{dB}}}{6}$. Fig. 4.7 shows the slope of the spectrum estimated using linear regression, where the spectral slope is estimated to be -11.17 dB/octave. Fig. 4.8 shows two examples of the spectral peaks in Fig. 4.7 adjusted to have slopes of -3 and 0 dB/octave. This spectral flattening technique will be later referred to as the Linear Regression Spectral Flattening LRSF method.

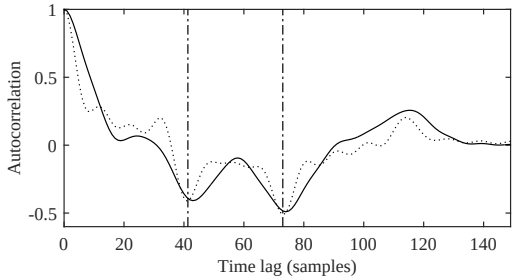


Figure 4.9: The autocorrelations of the spectral peaks of Fig. 4.7 (solid line) and the electric guitar model (dotted line), where both of their spectral slopes are adjusted to -3 dB/octave beforehand. The vertical dash-dotted lines at lags τ_ρ and τ_d (from left to right) correspond to the plucking and pickup positions respectively.

Autocorrelation

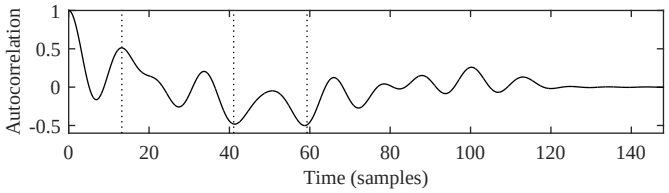
By using the frequency domain approach, the autocorrelation of the signal is equivalent to the inverse transform of the power spectrum. The amplitudes of the flattened spectrum $\bar{X}_k$ are used to calculate the autocorrelation (Traube and Depalle, 2003):

$$\Gamma(\tau) = \sum_{k=1}^{K} \bar{X}_k^2 \cos(\frac{2\pi k\tau}{T_0}) \tag{4.4}$$

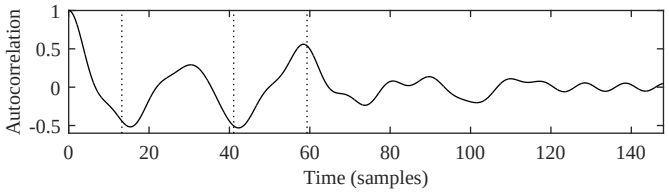
It is known that the autocorrelation compares the signal to its shifted copy, and shows peaks at multiples of the pitch period T_0 as discussed in Section 2.2.1. Other information can be found from the short-term evolution of the autocorrelation function, such as the pickup and plucking locations. This can be explained as follows: two waves travel in opposite directions from where the guitar string is plucked. The two pulses and their inverted reflections arrive at the pickup at different times. These time differences cause two strong negative correlations at time lags τ_d and τ_ρ that correspond to the pickup and plucking positions respectively. Note that the pickup and plucking positions have similar effects, both producing similar troughs but at different locations. Distinguishing between the two could be performed using post-processing techniques as discussed later in Section 5.3.1. The relationship between the lags, the relative positions and the parameter R in the electric guitar model can be expressed as:

$$\frac{\tau_d}{T_0} = \frac{d}{L} = R_d \tag{4.5}$$

$$\frac{\tau_\rho}{T_0} = \frac{\rho}{L} = R_\rho \tag{4.6}$$



(a)



(b)

Figure 4.10: The autocorrelations of the electric guitar model with (a) in-phase and (b) out-of-phase mixed pickup selected plucked on the open 4th string at 90 mm from the bridge. The spectral slope is adjusted to 0 dB/octave beforehand. The mixed pickup is a mix between the neck and middle pickup. The vertical lines (from left to right) at lags τ_β, τ_ρ and τ_α correspond to distances β, ρ and α respectively.

This means that the time lags also correspond to the parameters for the two comb filters. It is worth mentioning that the string length L does not have to be known when using the estimates as parameters for electric guitar synthesis if the distances are expressed as ratios R_d and R_ρ. Once the estimated time lags are found, the estimates $\hat{d}$ and $\hat{\rho}$ are calculated as follows:

$$\hat{d} = \frac{\hat{\tau_d}}{T_0} L \tag{4.7}$$

$$\hat{\rho} = \frac{\hat{\tau_\rho}}{T_0} L \tag{4.8}$$

Fig. 4.9 shows the autocorrelation of the electric guitar tone in Fig. 4.4 calculated from the spectrum that is flattened to -3 dB/octave. The two dominant troughs can be seen at time lags 43 and 73 samples. This means that the plucking and pickup position estimates are 94.25 mm and 160.01 mm, using Eq. (4.7) and (4.8) where $L = 651$ mm and $T_0 = 297$ samples. This gives absolute errors less than 5 mm for both estimates. Fig. 4.9 also shows the autocorrelation of the model which shows similar trough at lags τ_ρ and τ_d.

For in-phase mixed pickup signals, the electric guitar model in Eq. (2.12) and (2.13) predicts two troughs in the autocorrelation with lags corresponding to the locations of the

pluck ρ and the average of the two pickup locations α, and predicts one peak with lag corresponding to one half of the distance between the two pickups β. Fig. 4.10a shows an example of the electric guitar model with mixed pickup configuration $\mathrm{PU_{N+M}}$ plucked at 90 mm from the bridge. The two troughs at lags τ_ρ and τ_α, and the peak at lag τ_β can be seen. For out-of-phase pickups, the model in Eq. (2.14) and (2.15) predicts opposite effects for lags τ_α and τ_β as shown in Fig. 4.10b. When observing real mixed pickup data, there are some cases where the expected peak at τ_β is less prominent and cannot be distinguished. There is a way to emphasise this peak by using the LC-SP method which is explained later in Section 4.2.3. Also, the autocorrelation of a mixed pickup signal where the peak is not distinguishable is shown (see Fig. 4.17).

For a humbucker pickup which can be considered as a mixed pickup (explained in Section 2.1.1 page 26), it will be useful to treat it as a wide single pickup for practical purposes, where the lag τ_α will correspond to the middle of the humbucker pickup. This is because the value of β of a humbucker is very small which is around 4.5 mm (the width of a Seymour Duncan double coil pickup is around 36 mm), thus, it is very difficult to estimate τ_β when the expected peak could be around $\tau_\beta = \frac{4.5}{651} = 2$ samples for $T_0 = 297$ samples. So, estimating the distance α (located at the middle of the humbucker pickup) is more practical for this case. The distance α could be used to distinguish between guitar models which will be explained later in Section 4.5.1. For a known guitar with humbuckers, distances α can act as target locations to identify the pickup selection (this will be discussed later in Section 4.5.2). In the case of a known guitar with single coil pickups, if the estimate $\hat{\alpha}$ is located between two pickups, it can be assume that a mixed pickup configuration is selected.

Troughs near zero lag represent pickup or pluck locations closer to the bridge. Flattening the spectrum emphasises the higher harmonics, which in turn enhances detection of troughs that correspond to positions near the bridge. Over-flattening the spectrum, however, could create unwanted troughs near the zero lag. Fig.4.11 shows three autocorrelations of the same electric guitar tone where the slope of its spectrum is adjusted differently. There is an unwanted trough which can be seen near zero lag when the spectrum is adjusted to 0 dB/octave. On the other hand, by not flattening the spectrum, the two troughs are seen to be merged into a single trough which is dominated by the pickup or plucking position that is further away from the bridge. This is one of the reasons that estimating the pickup and plucking position using the time domain approach proposed by Penttinen and Välimäki (2004) cannot produce accurate results for electric guitar tones, where troughs related to

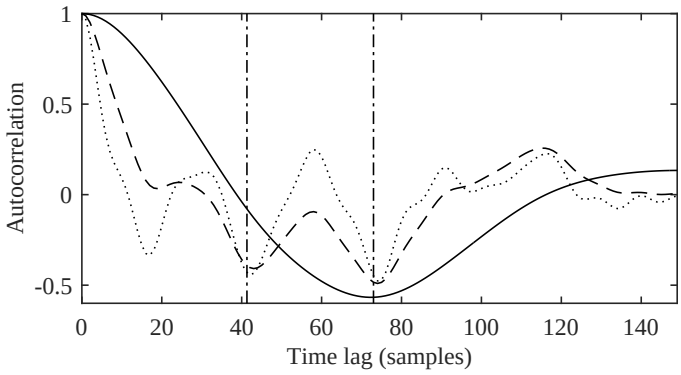


Figure 4.11: The autocorrelations of the spectral peaks of Fig. 4.7 (solid line) and with its spectral slope adjusted to -3 (dashed line) and 0 dB/octave (dotted line) beforehand. The vertical dash-dotted lines at lags τ_ρ and τ_d (from left to right) correspond to the plucking and pickup positions respectively.

pickup or plucking positions that are further away from zero lag are more dominant than the ones near zero lag.

Finding the Minima or Minimum of Autocorrelation: Grid Search

In the case where the plucking point is at or near the pickup, the troughs will merge into one, making it difficult to estimate the two locations independently from the time lags of the troughs. Estimating the plucking point of an acoustic guitar is therefore easier, because the autocorrelation of an acoustic guitar signal only produces one trough (Traube and Depalle, 2003; Penttinen and Välimäki, 2004). In order to solve the problem of merged troughs, where the pickup and plucking positions are close to each other, a grid search is employed to estimate the values. The mean square error between the autocorrelations of the observed data and the model for plucking points and pickup positions ranging from 25 mm to 180 mm with a spatial resolution of 1 mm is calculated. The electric guitar model is calculated using Eq. 2.11 which avoids using more parameters such as the plectrum and pickup width. The spectral slopes of the observed data and the electric guitar model are flattened to -3 dB/octave beforehand. The minimum mean square error yields the estimated pluck $\hat{\rho}$ and pickup locations $\hat{d}$. For in-phase mixed pickups, the minimum error yields the estimates $\hat{\rho}$ and $\hat{\alpha}$.

Further improvements can be obtained for estimates that are located near the bridge.

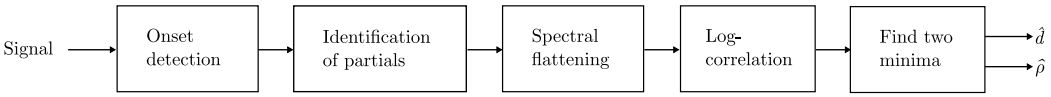


Figure 4.12: Block diagram for estimating pickup position and plucking point of an electric guitar using the Log-correlation of Spectral Peaks method (LC-SP)

While flattening the spectrum to -3 dB/octave may suppress unwanted troughs near zero lag in the autocorrelation, any correct estimates near the bridge will have a less sharp trough near zero lag. To compensate for this problem, the spectral slope is adjusted to 0 dB/octave for any pluck or pickup estimates that are less than 60 mm from the bridge. Then the grid search procedure described above is repeated, where the range of the search is from 25 mm to the previously estimated value.

4.2.3 Log-correlation of Spectral Peaks Method (LC-SP)

The third method (LC-SP) is similar to the AC-SP method. The overview of the LC-SP is shown in Fig. 4.12. The method used to find the onset time and spectral peaks of the electric guitar tone are the same as for the AC-SP as discussed in Section 4.2.2. The spectral slope is adjusted to 0 dB/octave using the LRSF method, and the log-correlation is calculated using the log amplitude of the spectral peaks. The two troughs in the log-correlation that correspond to the pickup and plucking positions are then found. In this section, further details are discussed for calculating the log-correlation and locating the minima.

Log-correlation

Traube and Depalle (2003) introduced a plucking point estimation on an acoustic guitar using the log version of the autocorrelation function, which the authors called it the "log-correlation". The log-correlation of the signal emphasises the contributions of low amplitude harmonics that are situated around the notches in the comb-filter by introducing large negative weighting coefficients. The log-correlation of the electric guitar tone is calculated as:

$$\Gamma'(\tau) = \sum_{k=1}^{K} \log(\bar{X}_k^2) \cos(\frac{2\pi k\tau}{T_0}) \tag{4.9}$$

For single pickup data, it is also expected to see two troughs in the log-correlation where the time lag of one trough τ_d indicates the position of the pickup and the time lag of the other τ_ρ indicates the position of the plucking event. For in-phase mixed pickup data, it is

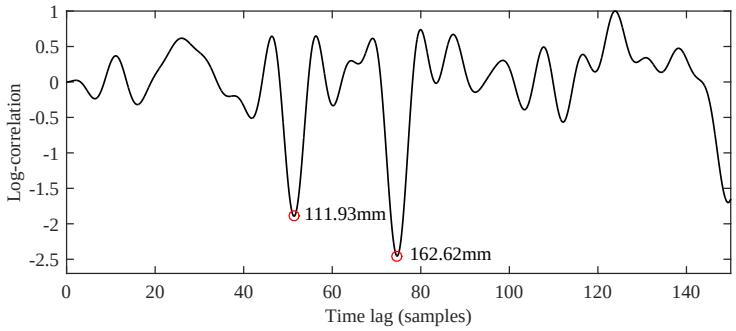


Figure 4.13: Log-correlation of an electric guitar tone where the open 4th string is plucked at 110 mm from the bridge with a pickup located at 159 mm from the bridge.

expected to see two troughs that correspond to distances α (the mid-point of the pickups) and ρ, and a peak that corresponds to β (half of the pickup spacing).

Finding the Minima or Minimum of Log-correlation: Trough Detection

Using the grid search method as described in Section 4.2.2 to find the troughs of the log-correlation of the guitar signal does not yield accurate results. This is because the comb filter notches in the electric guitar model calculated using Eq. (2.11) are emphasised by the large negative weighting of the log-transformation of small values ($V_k \rightarrow -\infty$). This produces sharp troughs at various time lags in the log-correlation whereas the log-correlation of real data exhibits more stable and smaller oscillation, which makes the grid search prone to errors.

Therefore, the two minima are found using a simple trough picking method. A local minimum in the log-correlation is detected when the amplitude of the current sample is less than that of the previous and the next sample. The range for detecting the local minima starts from the lag that corresponds to 25 mm from the bridge until 180 mm. The two lowest troughs that were detected correspond to the pickup and plucking position estimates. As an example, the log-correlation of an electric guitar plucked on the open 4th string at 110 mm from the bridge with a pickup located at 159 mm from the bridge is shown in Fig. 4.13. The log-correlation is calculated until T_0 samples with a time lag resolution of 0.01 samples. The two lowest troughs are visible at time lags of 74.66 samples (or 162.62 mm) and 51.38

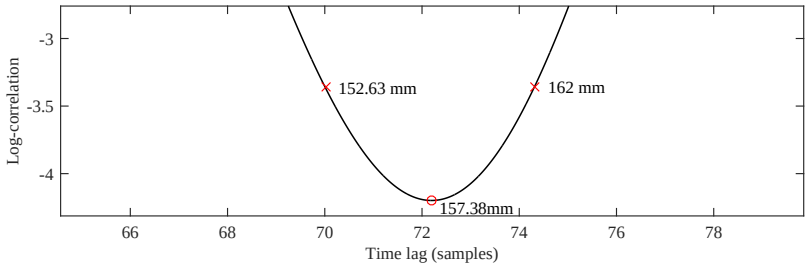


Figure 4.14: Log-correlation of an electric guitar tone where the open 4th string is plucked at 150 mm from the bridge with a pickup located at 159 mm from the bridge, resulting in the two troughs merging into a single trough.

samples (or 111.93 mm). Note that it is not possible to determine which represents the pickup and which is the plucking point from this information alone. The absolute errors for the pickup position and plucking point estimates are 3.62 mm and 1.93 mm respectively.

There are cases where the pluck is above or near the pickup which causes the two troughs to merge together. A threshold is empirically set to distinguish between this case and the case where the pluck is sufficiently far from the pickup. The second lowest trough is selected if the amplitude of the trough is below 40% of the lowest trough (note that the values are negative). An example is shown in Fig. 4.14 where the electric guitar is plucked on the open 4th string at 150 mm from the bridge with the pickup located at 159 mm from the bridge. Only the lowest trough is shown where the time lag is 72.20 samples (corresponding to 157.38 mm). Since the two expected troughs are merged together, it is impossible to obtain an accurate estimate of both the plucking point and the pickup position. The width of the merged trough reflects the distance between the pluck and pickup, thus, the time lags where both are at 80% of the minimum value are taken as the pickup and plucking position estimates. This also applies to plucks that are at the pickup position, where it is observed that the width of the trough is thinner. Fig. 4.14 shows that the estimated pickup position and plucking point are at 162 mm and 152.63 mm which yield 3 mm and 2.63 mm absolute errors respectively.

Parameter Estimation for Mixed Pickup Signals: Distinguish between Single and Mixed Pickups

Estimating the distance α may be sufficient for distinguishing between popular guitar models but not enough for guitar synthesis with mixed pickup configurations which also requires the parameter β. If the guitar is known and a mixed pickup configuration is known to be selected, the parameter β is taken based on the distances between the pickups.

There might be a case where the guitar is not known and a mixed pickup selection is identified. Typical single coil pickups are susceptible to noise at 50 Hz and in-phase mixed pickups suppress this effect. So, determining whether there is a 50 Hz hum/noise might distinguish between the two cases. Thus, in the case where the guitar is unknown and a mixed pickup is known to be selected, the parameter β needs to be estimated.

In audio forensics, there are techniques to detect whether a recording has been tampered by analysing the 50 Hz hum (Grigoras, 2005; Cooper, 2008). Under normal operating conditions, the Electric Network Frequency (ENF) is maintained within strict limits with frequency deviations up to 50 mHz. Since the ENF values may vary, the values can be extracted by finding the maximum of the power spectrum in a very narrow window. Then, the extracted ENF values are compared with a database of known ENF values at that particular time-stamp. The recording can be regarded as authentic if the values match. The detection of ENF is still an ongoing research and little efforts are made for ENF detection without prior knowledge. In a recent article, a possible approach to detect a recording with or without ENF is proposed by Hua et al. (2017). It is shown that a recording with ENF has a continuous energy in the frequency band of interest, while a recording without ENF has random patterns in that frequency band. However, no automatic detection to distinguish between the two cases are proposed but suggested further research directions to tackle this problem such as feature extraction and machine learning techniques may be applied to a large amount of recordings with or without ENF.

Fig. 4.15 shows spectrograms of the electric guitar plucked on the open 1st string at 30 mm from the bridge for each pickup configuration. The window length is set to 8192 samples with a hop size of 4096 samples, a hamming window and a zero padding factor of 4. The spectrograms for single pickups are shown in Figs. 4.15a, 4.15b and 4.15c and for mixed pickups are shown in Figs. 4.15d and 4.15e. It can be seen that the spectral peaks at around 50 Hz are more prominent for single pickups than mixed pickups. Note that a broadband

noise appears during the onset of the note, which might be due to the plucking noise.

The amplitude around 50 Hz is calculated for each case in Dataset II. For each signal in Dataset II, STFT analysis is performed, where the window length is 8192 samples with a hamming window, a hop size of 4096 samples and a zero padding factor of 4. For each frame, the maximum amplitude is searched in a narrow window starting from $f_{hum} - 50$mHz to $f_{hum} + 50$mHz, where f_{hum} is 50 Hz. Note that the length of this narrow window is chosen according to the normal operating conditions, where the ENF is maintained within strict limits with frequency deviations up to 50 mHz (Grigoras, 2005; Cooper, 2008). The maximum amplitude in the narrow window for each frame is averaged to give an estimate of the amplitude of the 50 Hz hum.

Fig. 4.16 shows the estimated amplitudes of the 50 Hz hum for Dataset II. There is a clear distinction between the single pickup and mixed pickup data. The amplitudes of the 50 Hz hum are all above -105 dB for single pickups and below -112 dB for mixed pickups.

The strength of the ENF might vary due to the proximity between the pickup and the power supply. Further investigations could test on different recording conditions such as recording the electric guitar in a studio and outdoor performances with various pickup to power supply distances. Furthermore, different electric guitars with in-phase mixed pickups or humbuckers might suppress the ENF even better than the electric guitar used in this experiment. So, further investigations should also test on several electric guitars, particularly the ones with different pickup configurations e.g. single, in-phase mixed pickups (humbuckers) and out-of-phase pickups.

Parameter Estimation for Mixed Pickup Signals: Parameter β Estimation

As mentioned in Section 4.2.2, the autocorrelation of a mixed pickup signal shows two troughs that correspond to the pickups' mid-point α and plucking position ρ, and a peak that corresponds to position β. There are some cases where the peak is not prominent. An example is shown in Fig. 4.17 where the peak at τ_β of the autocorrelation cannot be distinguished. The peak can be emphasised by taking the log-correlation of the signal calculated using Eq. (4.9). The peak can now be seen at 21.43 samples which gives the estimated distance $\hat{\beta} = 35$ mm. Thus, the absolute error of the estimate $\hat{\beta}$ is 6 mm.

Given the log-correlation Γ' of a mixed pickup signal, the lag τ_β is estimated as follows:

Firstly, the first peak and the first trough that are closest to zero lag are detected. There are two cases that should be considered when finding the lag τ_β, where one is when

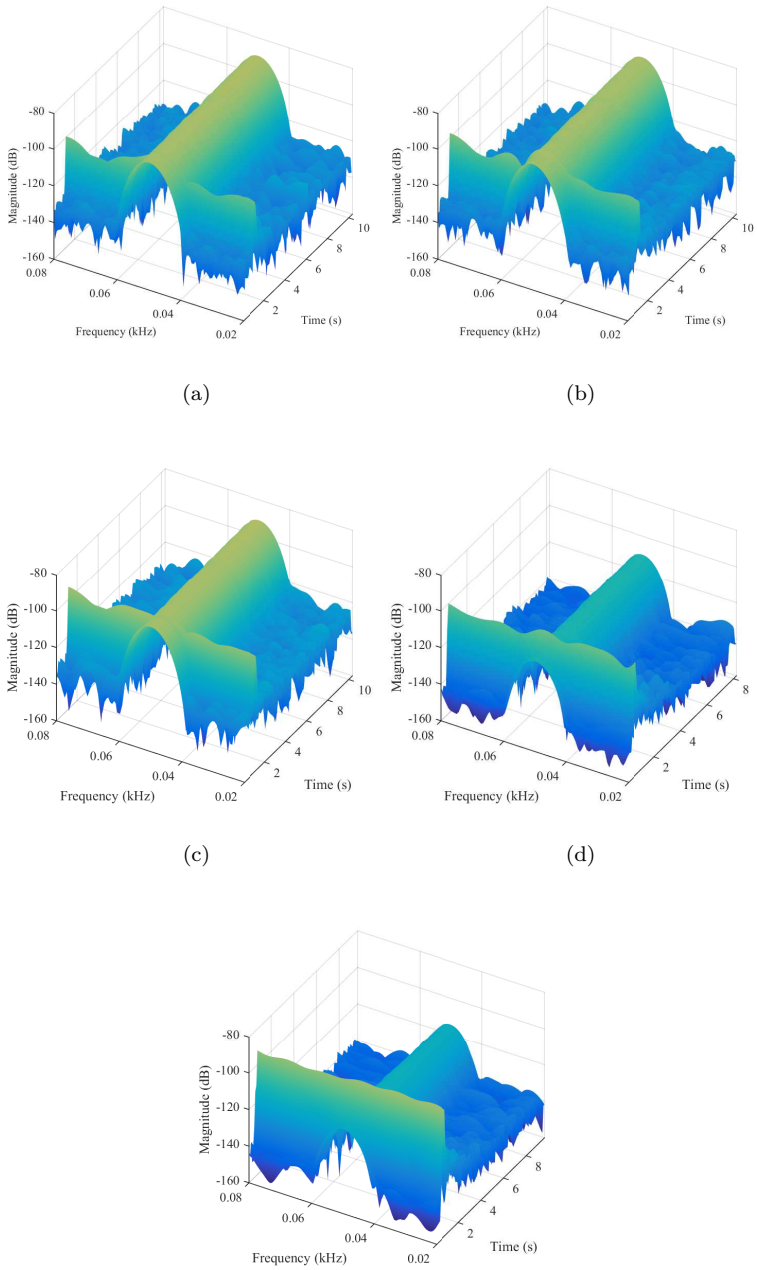


Figure 4.15: The time-frequency analysis of the electric guitar string plucked on the open 1st string at 30 mm from the bridge captured from the (a) neck pickup, (b) middle pickup, (c) bridge pickup, (d) a mix between neck and bridge pickups and (e) a mix between middle and bridge pickups in Dataset II.

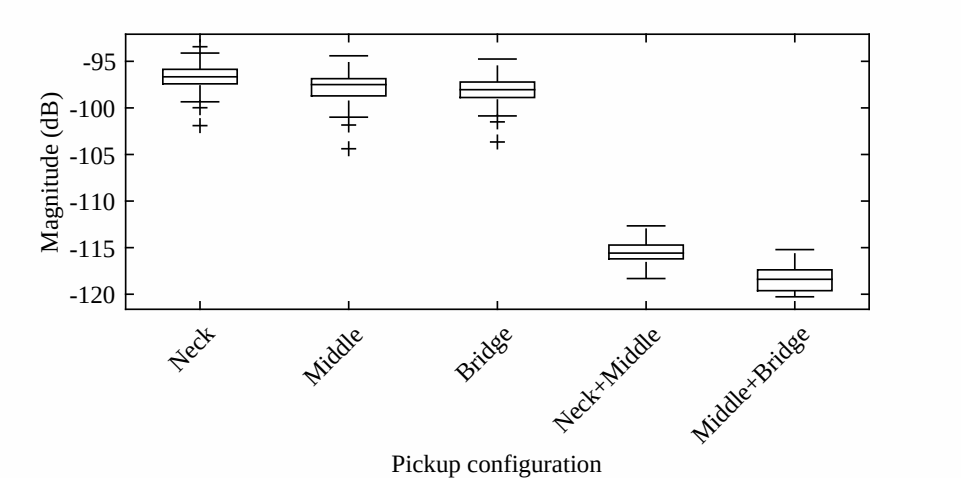


Figure 4.16: The amplitudes of the 50 Hz hum for neck pickup, middle pickup, bridge pickup, a mix between neck and bridge pickups and a mix between middle and bridge pickups.

Algorithm 1 Parameter β estimation algorithm

procedure EstimateBeta(Γ')
  $\Gamma'_{pk}, \tau_{pk} \leftarrow$ findPeakClosestToZeroLag(Γ')
  $\Gamma'_{tr}, \tau_{tr} \leftarrow$ findTroughClosestToZeroLag(Γ')
  if $|\Gamma'_{tr}| > |\Gamma'_{pk}|$ **then**
    $\hat{\tau}_\beta \leftarrow \tau_{tr}$
  else
    $\hat{\tau}_\beta \leftarrow \tau_{pk}$
  end if
  return $\hat{\tau}_\beta$
end procedure

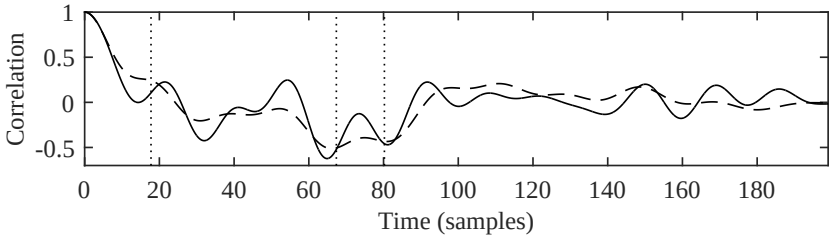
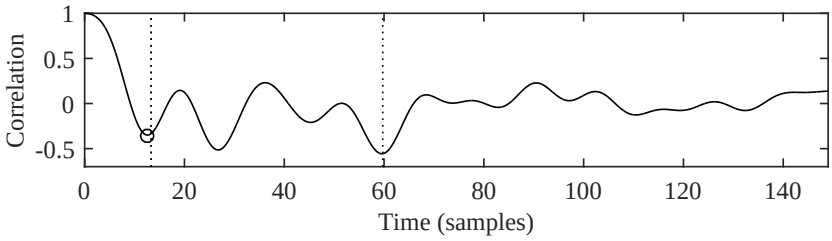


Figure 4.17: Autocorrelation (dashed line) and log-correlation (solid line) of a mixed pickup signal $\mathrm{PU_{N+M}}$ plucked on the open 5th string at 110 mm from the bridge. The vertical dashed lines (from left to right) at lag τ_β, τ_ρ and τ_α correspond to distance β, ρ and α respectively.

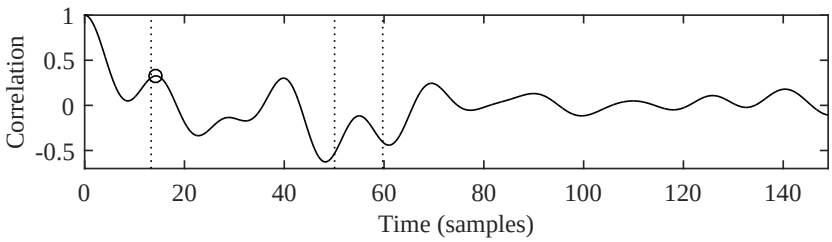
the plucking point is at or near distance β and the other is when the plucking point is away from distance β. Fig. 4.18 illustrates two log-correlations with the same mixed pickup configuration where the string is plucked on the open 4th string ($f_0 = 147$ Hz) at 30 mm and 110 mm from the bridge, and their lags τ_ρ are at 14.15 and 51.87 samples respectively. For this example, the distance β is 29 mm and lag τ_β is at 13.33 samples.

Fig. 4.18a shows the log-correlation of the electric guitar plucked near distance β. The peak that corresponds to the half of pickup spacing β and the trough that corresponds to the plucking point ρ could cancel each other out. However, the trough at τ_ρ seems to be more dominant than the expected peak at τ_β in practical cases as shown in Fig. 4.18a. So, the estimated lag $\hat{\tau}_\beta$ is set equal to τ_{tr} if the absolute log-correlation $|\Gamma'_{tr}|$ at the first trough is higher than the absolute log-correlation $|\Gamma'_{pk}|$ at the first peak.

Fig. 4.18b shows the log-correlation of the guitar signal where the plucking point is further away from distance β, and the expected peak at τ_β can be seen. Unfortunately, for the case where the plucking point is away from distance β, a false trough near zero lag is detected because the log-correlation always starts with a trough. In order to avoid this problem, the false trough is ignored if the absolute log-correlation $|\Gamma'_{tr}|$ at τ_{tr} is less than the absolute log-correlation $|\Gamma'_{pk}|$ at τ_{pk}. For this case, the estimated lag $\hat{\tau}_\beta$ is at τ_{pk}. Fig. 4.18a shows the expected trough at 13.01 samples is successfully detected. The false trough at 8.6 samples in Fig. 4.18b is ignored and the expected peak at 14.71 samples is successfully detected. The absolute errors of estimates $\hat{\beta}$ in Figs. 4.18a and 4.18b are 1.62 and 2.12 mm respectively.



(a)



(b)

Figure 4.18: Log-correlations of an electric guitar tone played on the open 4th string ($f_0 = 147$ Hz) plucked at (a) 30 mm and (b) 110 mm from the bridge, with mixed pickup configuration $\mathrm{PU_{N+M}}$ is selected where its distance β is 29 mm and distance α is 130 mm. For (a), the vertical dashed lines (from left to right) at lag τ_ρ and τ_α correspond to distance ρ and α respectively. Note that distance $\beta \approx \rho$ in this case. For (b), the vertical dashed lines (from left to right) at lag τ_β, τ_ρ and τ_α correspond to distance β, ρ and α respectively. The circles indicate the estimated τ_β.

4.2.4 Setting the Total Number of Harmonics

Setting the total number of harmonics K depends on a number of factors. If the number of harmonics is too low, the pluck or pickup positions that are close to the bridge cannot be properly estimated. For an example, if $K = 20$ harmonics and the string length L is 648 mm, any pluck or pickup positions below $L/K = 32.4$ mm cannot be estimated correctly. Also, the total number of harmonics should not be higher than the Nyquist rate. For an example, if T_0 is 66 samples, K cannot be set more than 33 harmonics. Furthermore, the number of harmonics also depends on the fret position at which the string is stopped. The number of harmonics available on an open string is twice the number for the same string played at the 12th fret. Also, when the string is fretted, the string length is shortened but the pickup width remains constant which decreases the number of harmonics available (see Eq. (2.18) and (2.20)). Other factors that produce a low pass filtering effect such as softer plucks and greater pickup width could lessen the number of available harmonics that can be analysed.

Thus, the total number of harmonics for open string, fifth fret and twelfth fret are empirically selected to be 25, 20 and 15 respectively. Ideally, the number of harmonics should be automatically detected, but for electric guitar tones where several of their harmonics are suppressed by the pickup and plucking positions, it may be difficult to detect the total number of harmonics that are available. This is left for future work to automatically find the total number of harmonics.

4.3 Just-Noticeable Difference

The pickup and plucking position estimates can be used as parameters for electric guitar synthesis to replicate the sound from an audio recording. It is important to evaluate the effects of the accuracy of the estimates on the the produced sound. A subjective evaluation is performed in order to determine how accurate the estimates should be without hearing any difference. This hearing threshold is called the just-noticeable difference (JND).

An ABX test described by Clark (1982) is performed in order to determine the JND using the Web Audio Evaluation Tool designed by Jillings et al. (2015). Synthesised sounds with various plucking points using the Karplus-Strong algorithm proposed by (Jaffe and Smith, 1983) are used in this listening test. The synthesised guitar sounds are plucked on all six open strings at 40, 100 and 150 mm from the bridge with string length of 650 mm. For each

plucking position ρ, the string is plucked at 0.25, 0.5, 1, 2, 3, 4, 5, 10, 15 and 20 mm towards the nut away from the plucking position. The ABX test compares these deviations $\Delta\rho$ to their normal plucking positions ρ to identify detectable differences between them. A listener with 12 years of experience of playing electric guitars is chosen for this test. For each question, the listener is presented with two synthesised sounds, where one is plucked at ρ (sample A) and the other is plucked at $\Delta\rho$ away from ρ (sample B) followed by one unknown sample X that is randomly selected from either A or B. The listener is then required to correctly identify X as either sample A or B. This will lead up to 180 questions, where the questions are randomly arranged. If the difference between sample A and B is not noticeable, random guessing would incur a 50% chance of choosing the correct answer. So, some trials should be performed in order to have some degree of confidence which is considered as statistically significant. A 95% confidence level is usually considered statistically significant as suggested by Clark (1982), meaning that the sound samples are significantly different, while a 75% level is a common threshold for the JND (University of South Dakota, accessed March 21, 2018), suggesting that the differences between the two sound samples are "just" noticeable. The listening test is performed two more times by the same listener giving 54 trials for each $\Delta\rho$ to distinguish between the two samples.

Fig. 4.19 shows the percentage of how many times the listener chose the right answer for each $\Delta\rho$. As expected, sounds that are very similar to each other incurs random guessing. It can be seen that up until $\Delta\rho = 1$ mm, the correct answers given by the listener are less than 60%, meaning that the listener is not able to distinguish between the two test samples. The bar graph also shows that the percentage of correct answers increases gradually after $\Delta\rho = 1$ mm, and the listener starts to notice some differences at $\Delta\rho = 2$ mm, which is above the JND threshold. After $\Delta\rho = 4$ mm, the percentage of correct answers reaches the 95% confidence level, meaning that the amount of correctly distinguished sound samples are statistically significant. Furthermore, when the plucking points are 20 mm apart, the listener is able to distinguish all of the test samples. Since the pickup position also produces comb-filtering effect like the plucking point, it assumed that the JND for Δd is also 2 mm, where Δd is the distance between two pickup positions.

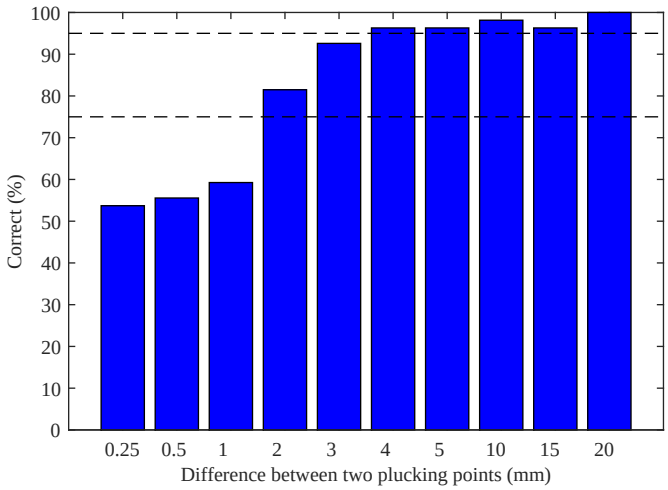


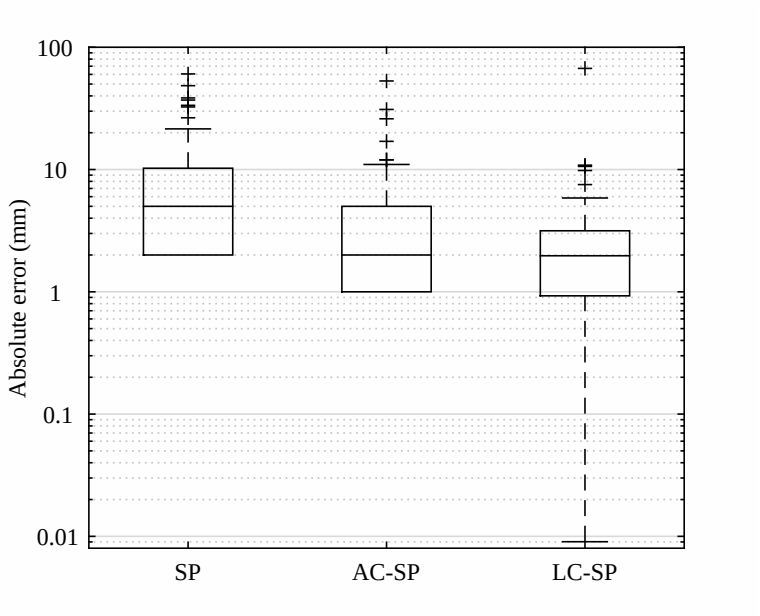
Figure 4.19: The percentage of correct answers. The horizontal dashed lines at 75% and 95% are the JND threshold and the significance level respectively.

4.4 Results: Pickup and Plucking Position Estimates

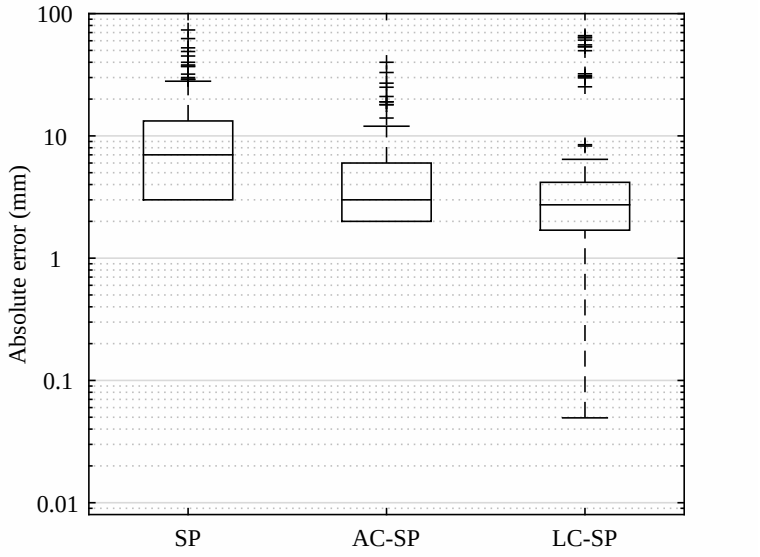
4.4.1 Single Pickup Data: Comparison of SP, AC-SP and LC-SP

The results for estimating the pickup and plucking position of the electric guitar are first presented from each single pickup. The single pickup subset of the Dataset II is used, comprising data from 3 single pickup configurations leading to a total of 144 audio samples (6 strings $\times$ 3 single pickups $\times$ 8 plucking points). It is not possible to distinguish between estimates belonging to the plucking point and the pickup position from this information alone. To test the accuracy of the estimates, the estimated value that is closest to the ground truth pickup position is taken as estimated pickup position and the other value as the plucking point. The total number of harmonics is set to 25 as discussed in Section 4.2.4. For autocorrelation methods i.e. AC-SP and LC-SP methods, the first three cycles of each tone is taken for analysis. Note that this window size is used for every experiment in this chapter.

To assess the accuracy of the estimates, the absolute error ε between the estimated and ground truth values is calculated. Fig. 4.20a shows the absolute errors for pickup position estimates ε_d using SP, AC-SP and LC-SP methods. The median absolute errors for pickup position estimates are 5.00 mm, 2.00 mm and 1.97 mm respectively. This suggests that each

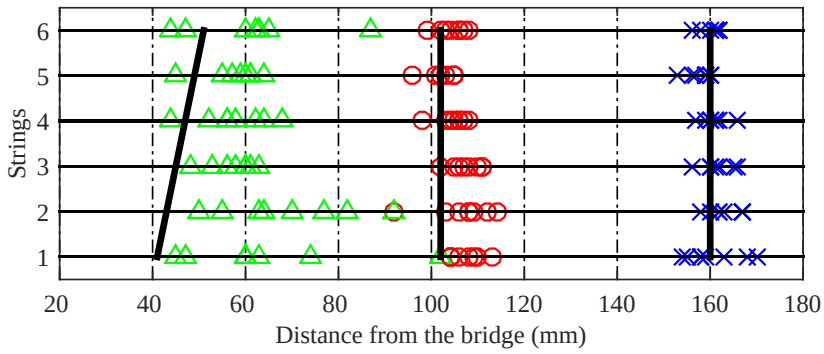


(a)

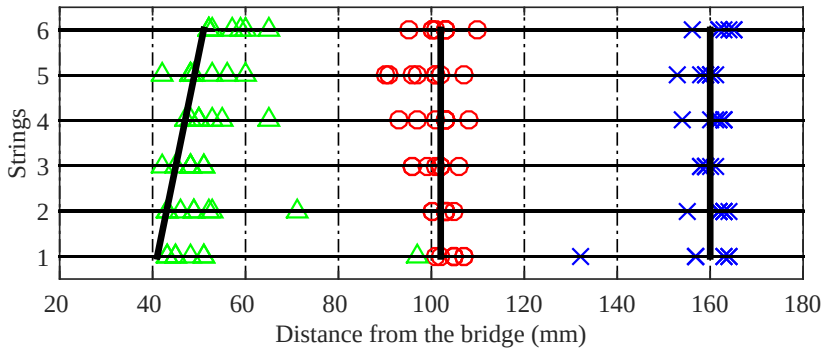


(b)

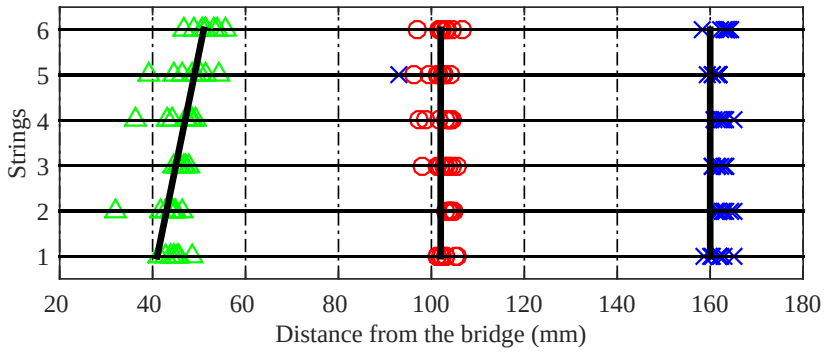
Figure 4.20: Box plot of absolute error in (a) pickup position and (b) plucking point estimates for each method on single pickup data. Note that the y-axis is in log scale.



(a)



(b)



(c)

Figure 4.21: Pickup position estimates using (a) SP, (b) AC-SP and (c) LC-SP methods. The estimates of bridge, middle and neck pickup locations are represented by triangles, circles and crosses respectively. The thick lines represent the ground truth pickup locations. Note that the bridge saddle positions are adjusted to the same position and the offsets are taken into account when displaying the estimates.

method produces quite accurate pickup position estimates. It also suggests that estimating the pickup positions using autocorrelation methods i.e. AC-SP and LC-SP is more precise and has less errors that are higher than 10 mm compared to the SP method. Comparing between autocorrelation methods, LC-SP is the most accurate which has lower minimum, first quartile, median, third quartile and maximum (excluding outliers) absolute errors. The median absolute errors for pickup position estimates using the autocorrelation methods are around the JND threshold, meaning that most of the errors may be noticeable. Nevertheless, the median errors are still less than 4 mm, meaning that the audible differences are not significant.

Fig. 4.20b shows the absolute errors for plucking point estimates using SP, AC-SP and LC-SP methods. As expected, the trend is similar to the pickup position estimates where the SP method performs the worst and LC-SP method is the most accurate. The median absolute errors for the SP, AC-SP and LC-SP are 7.00 mm, 3.00 mm and 2.73 mm. The median absolute errors for plucking position estimates using the autocorrelation methods are above the JND threshold, but the audible differences are not significant.

Although the LC-SP method appears to give more accurate results than the AC-SP method, by performing the paired-sample t-test, it was found that the differences are not statistically significant. In the case for single pickup data, the results for LC-SP are at least as good as those for AC-SP, and the LC-SP method has the additional advantage of faster computation. Note that the results from the AC-SP and LC-SP methods are significantly different when compared to the SP method by performing paired-sample t-tests.

The pickup position estimates and their ground truths are shown in Fig. 4.21. The bridge pickup estimates are less accurate for the SP method as shown in Fig. 4.21a. The SP method uses Fourier series to estimate the spectral peaks of the signal, which is not accurate for higher partials due to the inharmonicity of the strings. Any pickup or pluck that is close to the bridge suppresses high partials and the important nulls in the spectrum should be correctly identified. The Fourier series does not identify high partials correctly which affects the results for pickup and plucking positions near the bridge. The LC-SP method yields the best result, where most estimates are grouped together with small deviations from their ground truths compared to other methods.

The next experiments will look into assessing the accuracy on mixed pickup data and the effects of changing plucking dynamics and fret positions. The SP is not used for later experiments because it is expected to not perform well. One of the reasons is that the SP

method relies heavily on the accuracy of the electric guitar model where extra parameters are needed to include the effects of pickup width, plectrum width, plucking dynamics etc.

4.4.2 Single Pickup Data: Comparison of Existing Methods

Existing methods that estimate the plucking point of a guitar are discussed in Section 2.2.3 such as Penttinen and Välimäki (2004) and Traube and Depalle (2003) methods.

Penttinen and Välimäki (2004) proposed a method to locate the minimum of the autocorrelation of an acoustic guitar signal which corresponds to the plucking point. However, given an electric guitar signal, its autocorrelation will have two local minima that correspond to the pickup and plucking locations. Therefore, by using the method proposed by Penttinen and Välimäki (2004), the minimum of the autocorrelation might correspond to the pickup location instead of the plucking point producing a large error.

Traube and Depalle (2003) used a similar approach by analysing the log-correlation of a classical guitar tone. The first approximation of the plucking point is found by locating the minimum of the log-correlation. Then, the plucking point estimate is further refined by using weighted least squares. This only works well if the first approximation of the plucking point is close to the expected plucking point. Similar to the problem when using the method proposed by Penttinen and Välimäki (2004), the minimum of the log-correlation of an electric guitar tone might be the pickup location instead of the plucking point. This will also cause a large error for the plucking point estimate. Furthermore, the weighted least squares used by Traube and Depalle (2003) minimises the mean squared error between the ideal string equation (only accounts for the plucking position effect and does not include the pickup position effect) and the observed data. The weighted least squares will not converge to a feasible solution when given an electric guitar tone.

The plucking point results produced by the methods proposed by Penttinen and Välimäki (2004) and Traube and Depalle (2003) are compared with the LC-SP method. Note that only the first approximations of the plucking points are taken when using the method proposed by Traube and Depalle (2003) because the weighted least squares did not converge below the criterion error. This is caused by the large errors from the first approximations. Also, the least squares does not converge because the cost function that it is minimising does not consider the pickup position effect. The errors are shown in Fig. 4.22 showing that the LC-SP method produces significantly less errors than the methods proposed by Penttinen and Välimäki (2004) and Traube and Depalle (2003). As expected, the method

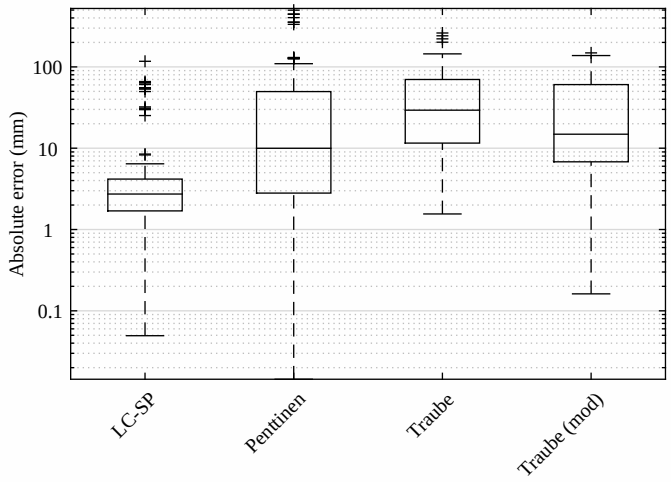


Figure 4.22: Box plot of absolute error in plucking point estimates on single pickup data for (from left to right) the LC-SP method, the method proposed by Penttinen and Välimäki (2004), the method proposed by Traube and Depalle (2003) and the method proposed by Traube and Depalle (2003) with some modification. Note that the y-axis is in log scale.

proposed by Penttinen and Välimäki (2004) has a larger interquartile range because the minimum of the autocorrelation might correspond to the pickup position instead of the plucking point. The method proposed by Traube and Depalle (2003) has larger errors because the log-correlation is calculated from the sustain part of the signal. More information can be found in the attack part of the signal because partials decay rapidly over time. Therefore, the method proposed by Traube and Depalle (2003) is modified, where the attack part of the signal is taken for calculating its log-correlation. Fig. 4.22 shows the modified version of the method proposed by Traube and Depalle (2003) showing that the results improve significantly.

4.4.3 Single Pickup Data: Effects of Onset Detection

Other types of onset detection methods are also considered. The LC-SP method is modified with various onset detection methods to test how they will affect the accuracy of the results. The onset detection techniques under test are the Spectral Flux (SF), Penttinen's High Frequency Content (P-HFC) and Rectified Complex Domain (RCD) methods.

By performing paired-sample t-tests, the results produced by LC-SP with the three onset

detection methods are not significantly different. The median errors for pickup position estimates with SF, P-HFC and RCD onset detections are 1.97, 2.06 and 1.97 mm respectively. The median errors for plucking point estimates with SF, P-HFC and RCD onset detections are 2.73, 2.76 and 2.73 mm respectively.

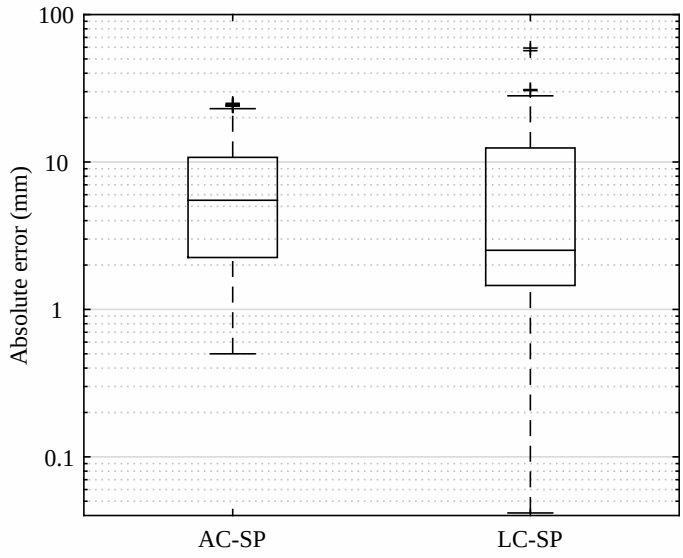
4.4.4 Mixed Pickup Data: Comparison of AC-SP and LC-SP

In order to assess the accuracy of the estimates on mixed pickup data, only autocorrelation methods should be used because these methods can estimate position α which can then be used to identify whether a mixed pickup configuration is selected. The identification of pickup selection is discussed later in Section 4.5.2. The mixed pickup subset of Dataset II is used, comprising data from 2 single pickup configurations $\mathrm{PU_{N+M}}$ and $\mathrm{PU_{M+B}}$ leading to a total of 96 audio samples (6 strings $\times$ 2 mixed pickups $\times$ 8 plucking points).

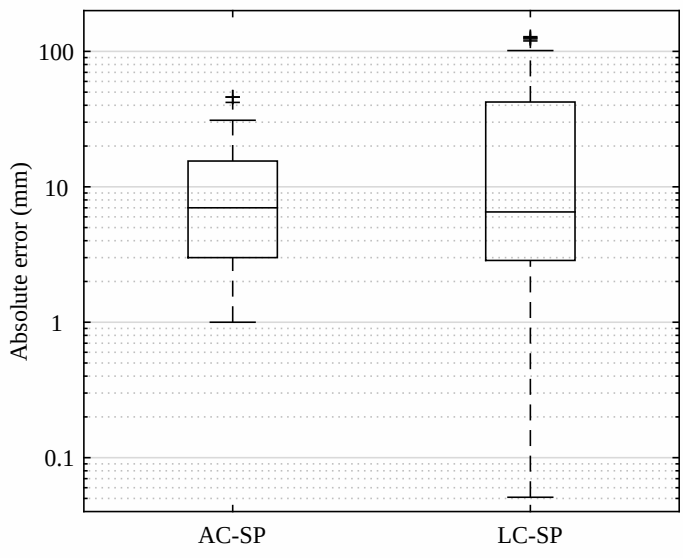
Fig. 4.23a shows the absolute errors for position $\hat{\alpha}$ estimates using the two autocorrelation methods. The median absolute errors for the AC-SP and LC-SP methods are 5.50 mm and 2.52 mm respectively. The number of outliers is higher for LC-SP which is mostly due to unwanted troughs that are lower in value than the expected troughs in the $\mathrm{PU_{M+B}}$ data.The log-correlation and over-flattening the spectrum might have enhanced these unwanted troughs. When using the AC-SP the unwanted troughs are less prominent. By performing the paired-sample t-test, the difference in the results for the AC-SP and LC-SP methods is not statistically significant. This means that the results for AC-SP are at least as good as those for LC-SP even though LC-SP produces a lower median absolute pickup error.

In Fig. 4.23b, the absolute errors for plucking point estimates using the two methods are shown. The same trend is observed as for the estimated position $\hat{\alpha}$ errors where the median absolute error is lower for LC-SP but the method produces some unwanted troughs in $\mathrm{PU_{M+B}}$ data, and thus a greater number of outliers. The plucking point results for AC-SP and LC-SP have a statistically significant difference. The interquartile range of the results for AC-SP has a smaller spread than LC-SP which means that AC-SP produces better plucking point estimates overall.

Assuming that a mixed pickup configuration is known to be selected, the distance β is estimated as discussed in Section 4.2.3. The median absolute errors for estimates $\hat{\beta}$ in $\mathrm{PU_{N+M}}$ and $\mathrm{PU_{M+B}}$ data are 1.81 mm and 1.9 mm respectively, which suggests that the method can estimate distance β accurately.



(a)



(b)

Figure 4.23: Box plot of absolute error in (a) pickup pair mid-point α and (b) plucking point estimates for AC-SP and LC-SP methods on mixed pickup data. Note that the y-axis is in log scale.

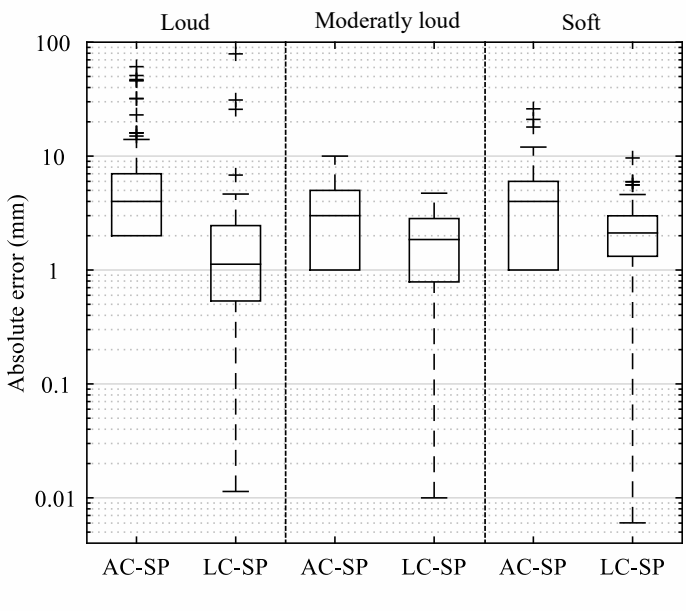
4.4.5 Effects of Plucking Dynamics

Section 2.1.1 discusses the effects of plucking dynamics on the spectrum of the electric guitar tone, where the level of high harmonics reduces when the string is plucked softly.

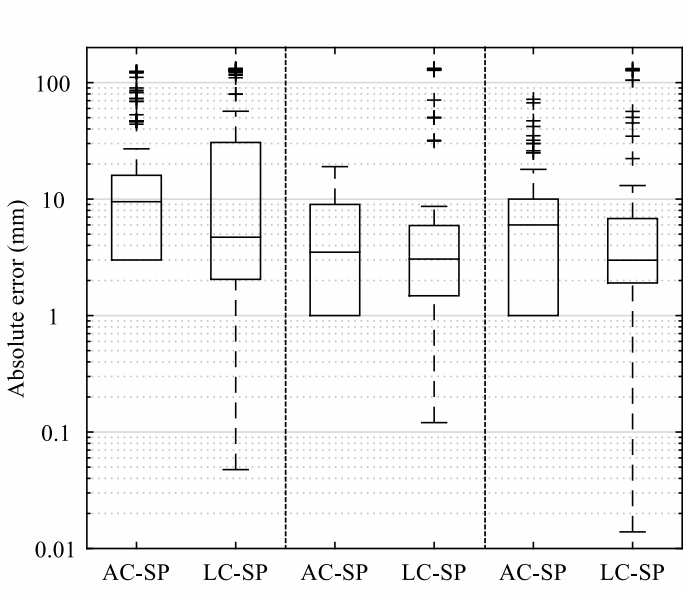
In this section, the effects of different plucking dynamics on the estimates are examined and compared between the AC-SP and LC-SP methods. Dataset I (Section 3.1) is used for this experiment. Fig. 4.24a shows the absolute errors for pickup position estimates ε_d using AC-SP and LC-SP methods for each plucking dynamic when the guitar is played on the open string. Each case has a total of 162 audio samples (6 open strings $\times$ 3 pickups $\times$ 3 plucking points $\times$ 3 instances). For each plucking dynamic, the median absolute errors when using AC-SP and LC-SP are less than 4 mm and 3 mm respectively. The third quartile of the absolute errors using the AC-SP and LC-SP methods are less than 6 mm and 3 mm, which suggests that the LC-SP produces more precise pickup position estimates. The AC-SP and LC-SP produce significantly different results for each plucking dynamic which is supported by the paired-sample significance test. The lower median, third and first quartile absolute errors in the LC-SP results confirm that the method produces more robust pickup position estimates towards different plucking dynamics than the AC-SP.

It can be seen in Fig. 4.24a that the number of outliers increased for louder tones, and to a lesser extent for softer tones, compared with the very robust results for *mezzo-forte* tones. This could be due to the nonlinear behaviour of the string when plucked at a higher force. The notches in the comb filter produced by the plucking point effect are less deep due to the nonlinear coupling between vibrating modes (Legge and Fletcher, 1984), where this effect can be more prominent when the string is strongly plucked (Penttinen, 1996, p. 8). This will depress the expected troughs in the autocorrelation affecting the results. The errors higher than 10 mm for softer tones when using AC-SP method are due to the grid search failing to find the troughs of the autocorrelation even though the troughs are around the expected time lag. The LC-SP method is more robust to the changes in plucking dynamics than the AC-SP method where 98.15% and 87.65% of forte results respectively have less than 10 mm error. All pickup position estimates using the LC-SP method for piano results have less than 10 mm error, whereas 95.06% of the estimates using the AC-SP method have less than 10 mm error.

Fig. 4.24b shows the absolute errors for plucking point estimates ε_ρ. A similar trend is shown where the number of outliers for both extremes of the dynamic range increases.



(a)



(b)

Figure 4.24: Box plot of absolute error in (a) pickup position and (b) plucking point estimates for AC-SP and LC-SP methods on single pickup data with three plucking dynamics. Each two boxes from the left represent errors where the guitar is played loudly, moderately loudly and softly respectively. Note that the y-axis is in log scale.

Nevertheless, the median absolute errors for the plucking point estimates when using the AC-SP and LC-SP methods are less than 10 mm and 5 mm respectively. The results for AC-SP and LC-SP in *forte* and *piano* data are not significantly different. On the other hand, the *mezzo-forte* results for AC-SP and LC-SP are significantly different which is caused by the outliers in the LC-SP results.

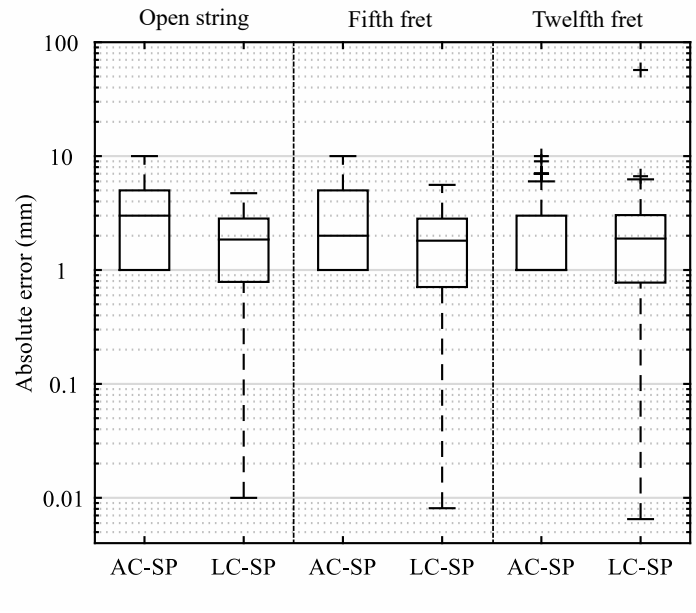
4.4.6 Effects of Fret Positions

So far, the experiments have estimated pickup positions and plucking points on open strings. In this section, the accuracy and robustness of the pickup and plucking position estimates on fretted strings when using AC-SP and LC-SP methods are investigated. The length of the string is shortened when the guitar is fretted at fret $\mathcal{F}$ by a factor of $2^{-\mathcal{F}/12}$. Therefore, a pickup at a fixed position suppresses different harmonics when the string is fretted than when it is open.

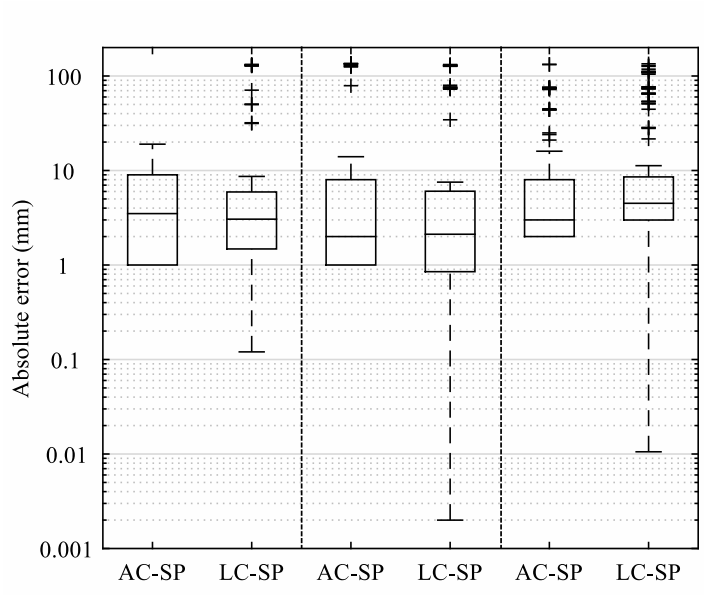
Recordings from Dataset I are used for this experiment. In Fig. 4.25, the absolute errors for pickup ε_d and plucking position estimates ε_ρ using AC-SP and LC-SP methods are shown for each fret position when the guitar is played moderately loud. Each case has 162 audio samples (6 strings $\times$ 3 single pickups $\times$ 3 plucking positions $\times$ 3 instances).

As shown in Fig. 4.25a, almost all of the errors for pickup position estimates are less than 10 mm. This suggests that both systems produce pickup position estimates that are highly robust to changes in fret positions. The median absolute errors using AC-SP and LC-SPfor all cases are less than 3 mm and 2 mm respectively. The paired-sample t-test indicates that the results for AC-SP and LC-SP on open string and fifth fret data are significantly different. The lower median, first and third quartile absolute errors in LC-SP results compared to AC-SP confirm that LC-SP produce better results. For the twelfth fret data, the t-test shows that the results for the two methods are not significantly different from each other. This means that the AC-SP is as good as the LC-SP when estimating the pickup position on twelfth fret data.

For the plucking point estimates, the number of outliers increases with the fret position, as shown in Fig. 4.25b. For the fifth fret, the outliers are due to unwanted troughs near zero lag. For the twelfth fret, the string length is halved ($L_{12} = L/2$) which causes a problem for the detection of pickup and plucking positions. Due to symmetry in the autocorrelation using a frequency domain approach, it is not possible to detect a pickup or a plucking position that is located more than $L_{\mathcal{F}}/2$ from the bridge. For cases where open strings and low fret



(a)



(b)

Figure 4.25: Box plot of absolute error in (a) pickup position and (b) plucking point estimates for AC-SP and LC-SP methods on single pickup data with three fret positions. Each two boxes from the left represent errors where the guitar is played on the open string, fifth fret and twelfth fret respectively. Note that the median absolute error for the 12th fret using AC-SP (second box plot from the right) is the same as its first quartile. Also note that the y-axis is in log scale.

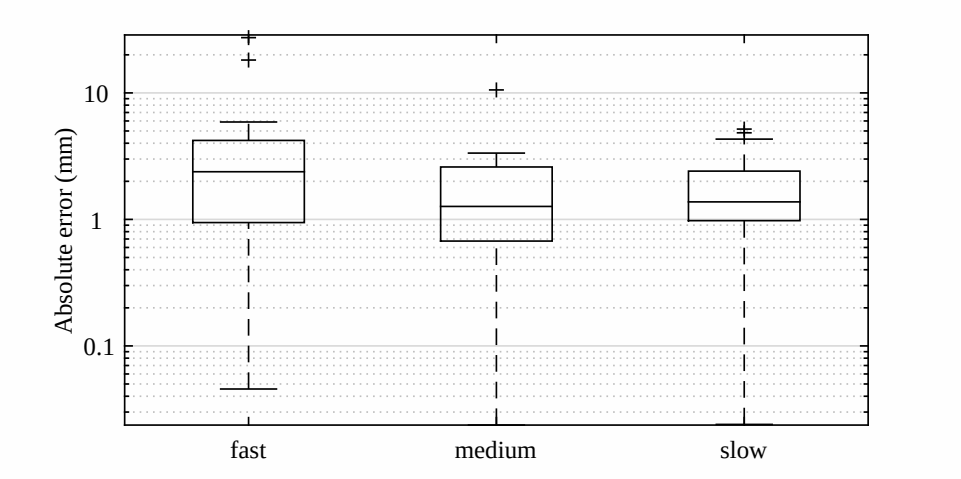
positions are played, the pickup and pluck can safely be assumed to be located in the half of the vibrating string nearest the bridge, but for higher fret positions, it is possible that the pickup or pluck are nearer to the stopped end of the string than the bridge. In the case for twelfth fret, positions that are around more than 162 mm could not be estimated correctly. The plucking points can vary and might go over the limit, which explains most of the outliers observed for the twelfth fret data. Improvements can be performed by initially estimating a position using the autocorrelation of the tone via a time domain approach as proposed by Penttinen and Välimäki (2004). Note that the time domain approach is biased to positions that are closer to the nut of the guitar, where the spectral flattening helps mediate this problem by enhancing the troughs that are closer to zero lag. The t-test in Matlab shows that the results for AC-SP and LC-SP on open string and twelfth fret data are significantly different, but for fifth fret data, the t-test suggests otherwise. This is mainly due to the outliers in the LC-SP results influencing the t-test. Nevertheless, the system can generally find accurate plucking positions, where the median absolute errors using AC-SP and LC-SP for all cases are less than 4 mm and 5 mm respectively.

4.4.7 Test on Chords

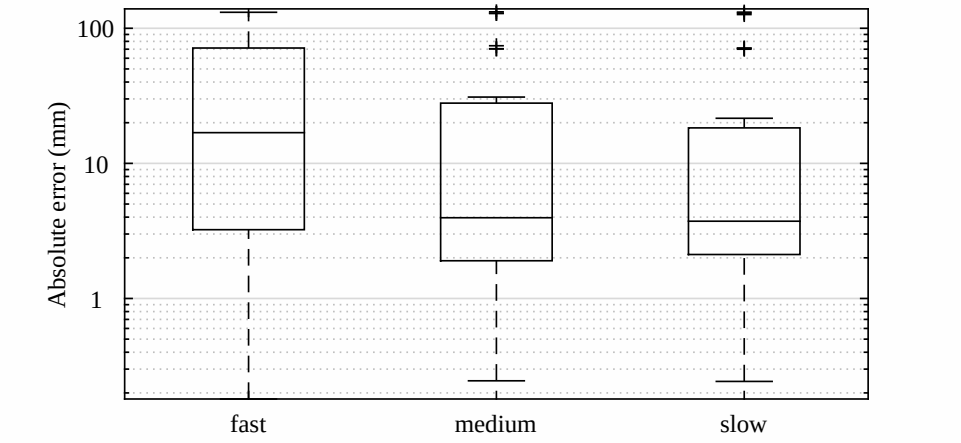
This section discusses the accuracy of the estimates on strummed chords. E major, A major and G major chords are strummed at 3 different speeds and 3 positions. The first string to be struck is the 6th string for all chords (downstrokes). The pickup and pluck positions are unlikely to change during the strum, so the method only requires the first few pitch periods of the electric guitar tone in order to estimate the pickup and plucking positions, where the second note is struck after a few cycles of the first note. The LC-SP method is chosen in this experiment for its faster computation and accuracy.

For the method to be unaffected by the strum, the shortest time allowed between the first and second note would be 36.4 ms (3 cycles of note 82.41 Hz) for the worst case scenario of the first pitch being E2 which is the lowest pitch on the guitar. In this experiment, the time between the first and second note, t_c is manually measured, where natural strumming of a guitar leads to values of t_c of 80 ms for slow strums and 20 ms for fast strums.

The fundamental frequencies of the first note struck on each chord are known beforehand, which are 82.41 Hz (E major and A major) and 98.00 Hz (G major). Note that the A major chord is played in second inversion (i.e. with a low E in bass) in order to present the worst case scenario. A method proposed by Klapuri (2003) that estimates multiple pitches of a



(a)



(b)

Figure 4.26: Box plot of absolute error in (a) pickup position and (b) plucking point estimates with three strumming speeds. Note that the y-axis is in log scale.

polyphonic recording could be used to find the fundamental frequency of the first note.

A shorter window is needed for faster strums so that less of the second note is included in the STFT analysis so that it does not affect the results. Therefore, for t_c shorter than 40 ms, the first 2 cycles are taken for the STFT analysis and the first 3 cycles are taken when t_c is longer than 40 ms. The total number of harmonics K is set to 25.

Fig. 4.26a and 4.26b show the absolute errors for pickup and plucking position estimates respectively. The plucking position estimates are less reliable for faster strums, where the median absolute error is 17 mm. Nevertheless, the median absolute errors are less than 4 mm when the chords are strummed at a medium and slower speeds. Almost all of the pickup position errors are less than 6 mm for all strumming speeds even though the second note starts to bleed into the window.

4.5 Real-world Applications

There are several practical uses for finding the pickup and plucking positions of an electric guitar. Other string instruments that use pickups to capture the vibrations of its strings could find similar applications such as the electric bass guitar. Fig 4.27 shows a complete system of estimating the pickup and plucking locations of an electric guitar signal and the uses of these estimates. First, the onset times of an electric guitar signal are estimated, and the fundamental frequency for each note is then estimated. The techniques for onset and pitch detection are discussed in Section 4.2.1. Then, the string and fret position for each note should be detected because the string length is shortened by a factor of $2^{-\mathcal{F}/12}$ when the fret is pressed, thus, the ratio estimate $\hat{R}$ of a pickup or a plucking position for fretted strings is reduced by the same factor. Dividing the ratio estimates $\hat{R}$ by $2^{\mathcal{F}/12}$ will normalise each estimate, which can help with post-processing techniques i.e. distinguishing the pickup from a pluck, guitar identification and pickup selection identification. Also, if the fret position is known, the total number of harmonics K that are available reduces for fretted strings as discussed in Section 4.2.4, hence, this can be defined accordingly. The fourth step is to estimate the pickup and plucking positions using either AC-SP or LC-SP. Each note will have two ratio estimates $\hat{R}_1$ and $\hat{R}_2$ which does not distinguish pickup from pluck estimates immediately. A post-processing technique is required to distinguish pickup from pluck estimates, which is later explained in Section 5.3.1. Then, the method yields the estimated pickup $\hat{R}_d$ or $\hat{R}_\alpha$ and plucking positions $\hat{R}_\rho$.

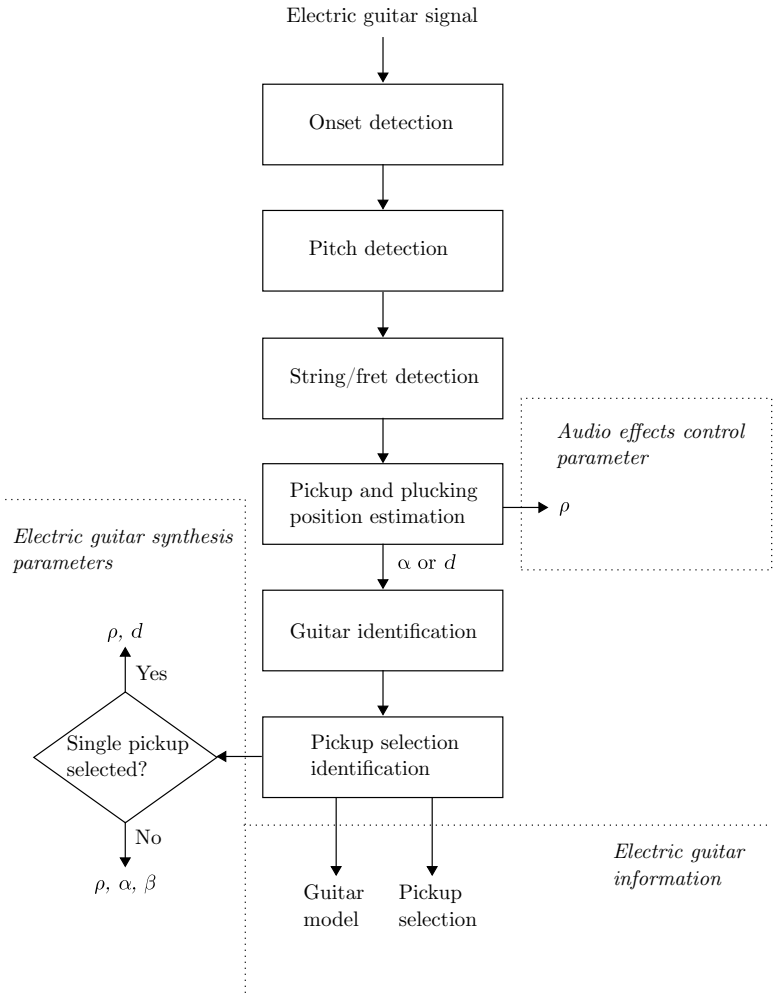


Figure 4.27: A complete system of estimating the pickup and plucking positions and its applications.

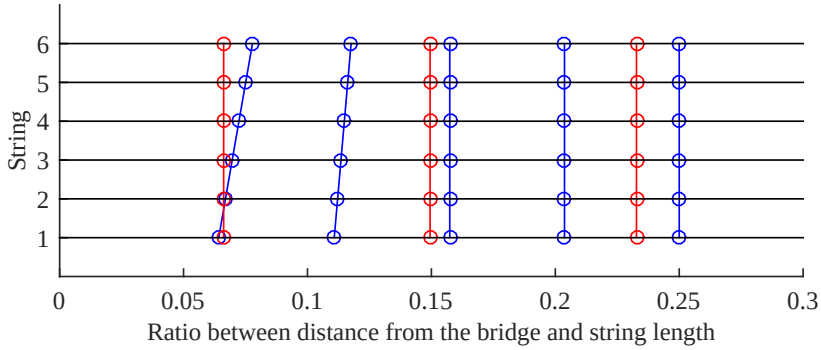


Figure 4.28: A Fender Stratocaster's (blue) and a Gibson Les Paul's (red) target locations. The target locations consist of the guitars' pickup locations R_d and pickup pair mid-points R_α. Note that a Fender Stratocaster has 3 pickups and a Gibson Les Paul has 2 pickups.

One example of the practical uses of these estimates is that the plucking position estimates can be used to control audio effects parameters in real-time as discussed by Penttinen et al. (2005). For an example, the gain of a distortion effect can be changed according to the user's plucking positions. For the case of electric guitars, two minima that correspond to the pickup and plucking positions in the autocorrelation are found. The pickup selection can be defined by the user beforehand telling the system that the other trough should correspond to the position of the pluck. This enables the system to predict the plucking position in real-time.

Alternatively, the pickup estimates can be used to distinguish between popular guitar models, where popular guitar models have different pickup positions. For a known guitar, the pickup estimates can be used to identify which pickup configuration is used in a recording as demonstrated in Section 4.5.2. If the pickup configuration (i.e. single or mixed pickup) is known, the estimates can be used as parameters for electric guitar synthesis to replicate the sound of guitarists from their recordings.

4.5.1 Identification of Electric Guitar Model

The pickup position estimates could be used to distinguish between popular guitar models because their pickup locations are quite distinct. As an example, a Fender Stratocaster's pickup locations and angles are different from those of a Gibson Les Paul. The pickup locations R_d and their mid-points R_α of a Fender Stratocaster and a Gibson Les Paul

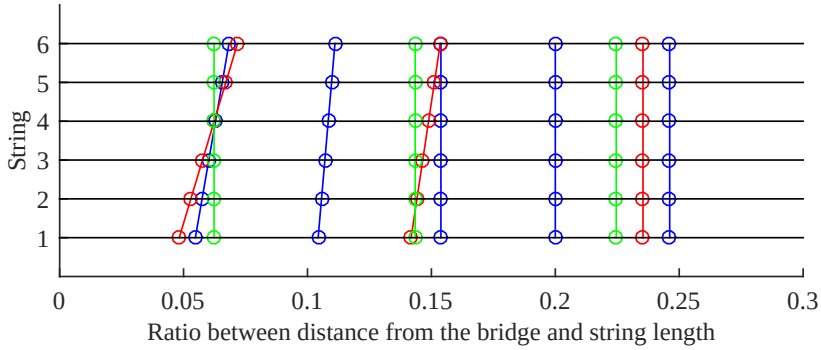


Figure 4.29: A Fender Stratocaster's (blue), a Fender Telecaster's (red) and a Fender Jaguar's target locations. The target locations consist of the guitars' pickup locations R_d and pickup pair mid-points R_α. Note that a Fender Stratocaster has 3 pickups, and a Fender Telecaster and Jaguar have 2 pickups.

are shown in Fig. 4.28. These locations can be used as target locations to identify the guitar model and its pickup selection. So, if the pickup position estimates are closer to a Fender Stratocaster's target location, a Fender Stratocaster might be used in the recording. Comparing between the target locations of a Fender Stratocaster and a Gibson Les Paul, most of the locations are distant from each other. However, the bridge pickups of a Fender Stratocaster and a Gibson Les Paul (see target locations that are closest to the bridge in Fig. 4.28) seem to overlap with each other, but the two target locations can be differentiated by their angles. If the information in the recording is sufficient i.e. the number of strings plucked is more than 2, they could be distinguished easily.

Another example is to use the estimates to distinguish between guitar models of a famous brand e.g. Fender. Fig. 4.29 shows the target locations of a Stratocaster, a Telecaster and a Jaguar manufactured by Fender. The bridge pickups of the three guitar models are located very close to each other, especially the locations at the 4th and 5th strings. Nonetheless, they can be distinguished by their angles. Other target locations are easier to distinguish, where the spacings between them are larger.

4.5.2 Identification of Pickup Selection

If the guitar is known, the pickup position estimates can be used to identify which pickup is selected in the recording. A three-pickup guitar such as the Fender Stratocaster typically

	Correct (%)				
Method	PU_B	PU_M	PU_N	PU_{N+M}	PU_{M+B}
AC-SP	97.92	100.00	91.67	75.00	89.58
LC-SP	97.92	100.00	100.00	91.67	66.67

Table 4.1: The percentage of pickup selections identified correctly using AC-SP and LC-SP methods.

has 5 pickup configurations which are 3 single pickup selections (bridge, middle and neck pickup) and 2 mixed pickup selections (a mix of bridge and middle pickups and a mix of middle and neck pickups). Therefore, 5 target locations can be allocated to be distinguished from each other. The target locations for mixed pickup signals are in between the two pickups (distance α). It is also worth noting that a two-pickup guitar such as the Gibson Les Paul should have 3 target locations, where it typically has 2 single pickup and one mixed pickup selection. The pickup selection is identified when the pickup estimate is closest to its target location compared to others.

The pickup configuration that is selected can be successfully identified using AC-SP and LC-SP methods. The first three cycles of each guitar tone in Dataset II are analysed and the total number of harmonics K is set to 25. Table 4.1 shows the number of correctly identified pickup selections as percentages using the AC-SP and LC-SP methods. Overall, most pickup selections are correctly identified. The mixed configuration PU_{M+B} identified using LC-SP has a lower percentage than others, which is due to unwanted troughs in the log-correlation (see Section 4.4.4). Note that these identified pickups are based on only single notes. For real-world applications, pickup position estimates from sequences of notes in a song may help remove outliers, allowing the pickup selection to be identified correctly.

Chapter 5

Pickup and Plucking Position Estimation in Real World Settings

Frequency domain approaches to estimate the pickup and plucking positions on an electric guitar were discussed in the previous chapter, where only direct input electric guitar signals were used in the experiments. This chapter explores the effects on the pickup and plucking position estimates in a real world setting, where various guitar signal chains are applied to the direct input signals. The LC-SP method is chosen for the experiments in this chapter because overall it produces accurate estimates with fast computation compared to other approaches.

In order to estimate the pickup and plucking positions of an arbitrary electric guitar from a published recording, the changes to the frequency contents of the electric guitar signal made by guitar effects, amplifiers, loudspeakers, microphones and post-processing audio effects should be taken into account. The LC-SP can be modified to compensate these effects, where another spectral flattening method is introduced. The proposed spectral flattening method will be discussed in Section 5.1. The accuracy of the estimates are then evaluated on the electric guitar with several different presets i.e. different guitar effects, amplifiers, loudspeakers and microphones with different settings. The results and discussions of the experiment are presented in Section 5.2. Section 5.1 and 5.2 are based on a published literature by Mohamad et al. (2017b).

Two cases of estimating the pickup and plucking positions of published guitar recordings are presented in Section 5.3,where a method to distinguish between pickup and pluck estimates is presented. Also, the identification of the electric guitar and its pickup selection based on the pickup position estimates is discussed in Section 5.3.

5.1 Polynomial Regression Spectral Flattening (PRSF)

It is known that the effects, amplifier, loudspeaker and microphone alter the frequency response of the guitar tone as discussed in Section 2.1, and this can affect the estimation of performance parameters. In order to address this problem, another approach to flatten the spectrum is introduced using polynomial regression. Not only can this reverse the low-pass filtering effect due to finite widths of the plectrum and pickup, plucking dynamics and nonrigid end supports, but it can also approximate the frequency response produced by the guitar signal chain and mitigate its effects as well.

The best fitting curve for the log magnitude X_k in the log-frequency domain is calculated

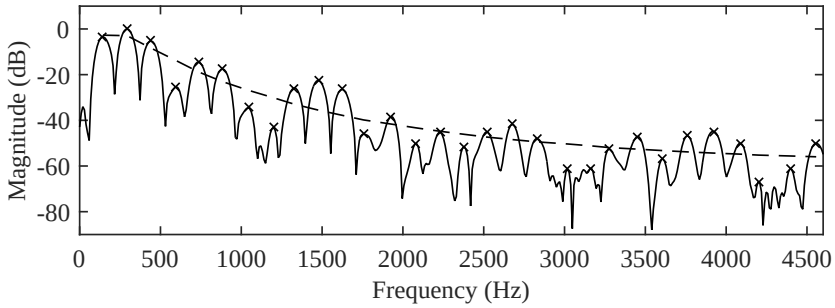


Figure 5.1: The spectrum of the first 3 cycles of the electric guitar tone in Fig. 4.4 (solid line), the magnitudes of each of its partials in decibels (crosses) and the envelope of the spectrum estimated using polynomial regression (dashed line).

via polynomial regression:

$$\log(X_k) = g(\log(k)) = g_0 + g_1 \log(k) + \cdots + g_G \log(k)^G \tag{5.1}$$

where g_i are the polynomial regression coefficients and G is the order of the polynomial regression. Section 5.2 shows that a third-order polynomial regression is sufficient to follow the curve in the frequency response produced by a typical guitar signal chain. The polynomial regression coefficients g are obtained using least squares solution. Fig. 5.1 shows the spectrum of an electric guitar tone and the slope of the spectrum estimated using polynomial regression.

The spectrum X_k can then be flattened as follows:

$$\bar{X}_k = \frac{X_k}{e^{g\left(\log(k)\right)}} \tag{5.2}$$

This spectral flattening technique is later referred to as the Polynomial Regression Spectral Flattening (PRSF) method.

5.2 Test on Various Guitar Signal Chains

5.2.1 Audio samples

The 144 direct input electric guitar signals recorded in Dataset II discussed in Section 4.1 are taken to test the accuracy of the estimates on various guitar signal chains. In this dataset, the Squier Stratocaster is plucked moderately loud (mezzo-forte) on 6 open strings at 8 plucking positions using 3 single pickup configurations. In order to test the effects of

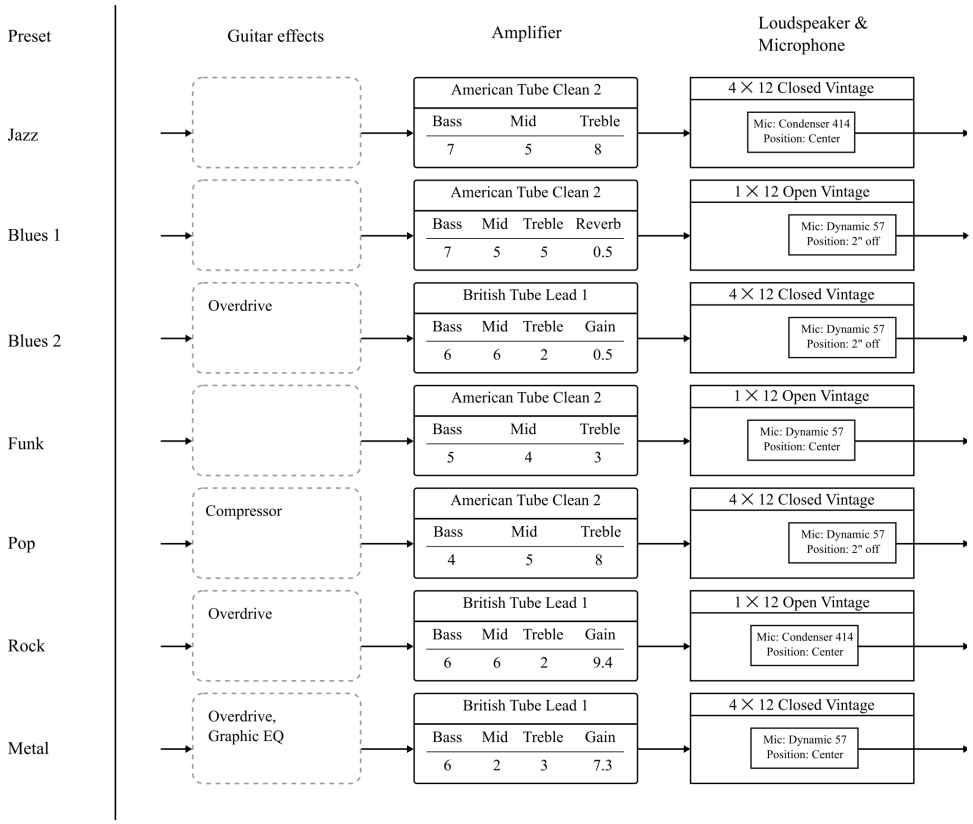


Figure 5.2: Seven combinations of emulated electric guitar effects, amplifier, loudspeaker and microphone, as published by Mohamad et al. (2017b).

different guitar signal chains, each signal is processed through 7 combinations of digitally emulated electric guitar effects, amplifier, loudspeaker cabinet and microphone using the commercial software Reaper (Cockos, Inc., accessed April 8, 2017). This leads to a total of 1152 audio samples in total (8×144 samples, where direct input signals are also included for the experiments).

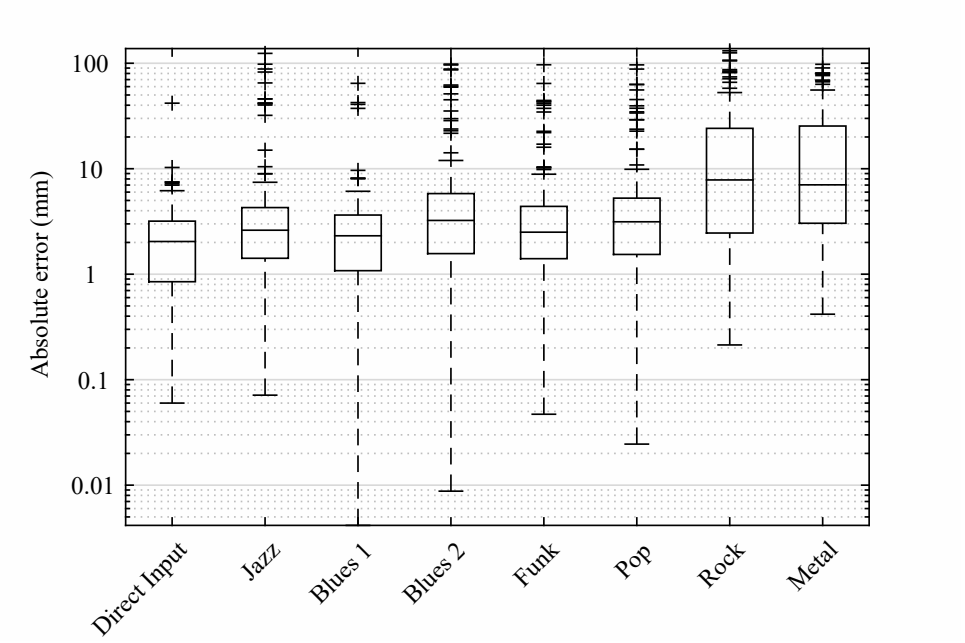
The selected presets are freely available in Amplitube Custom Shop by IK Multimedia, where each of them produces a tone for a certain style (IK Multimedia, accessed April 8, 2017). Each preset represents a typical combination of equipment and settings used for a specific style of music. Popular styles such as Jazz, Blues, Funk, Pop, Rock and Metal are selected for this experiment. Fig. 5.2 shows the 3 emulated guitar effects, 2 emulated amplifiers, 2 emulated loudspeaker cabinets and 2 emulated microphones used in the experiment. Also, note that each preset has different settings and microphone placement.

5.2.2 Results

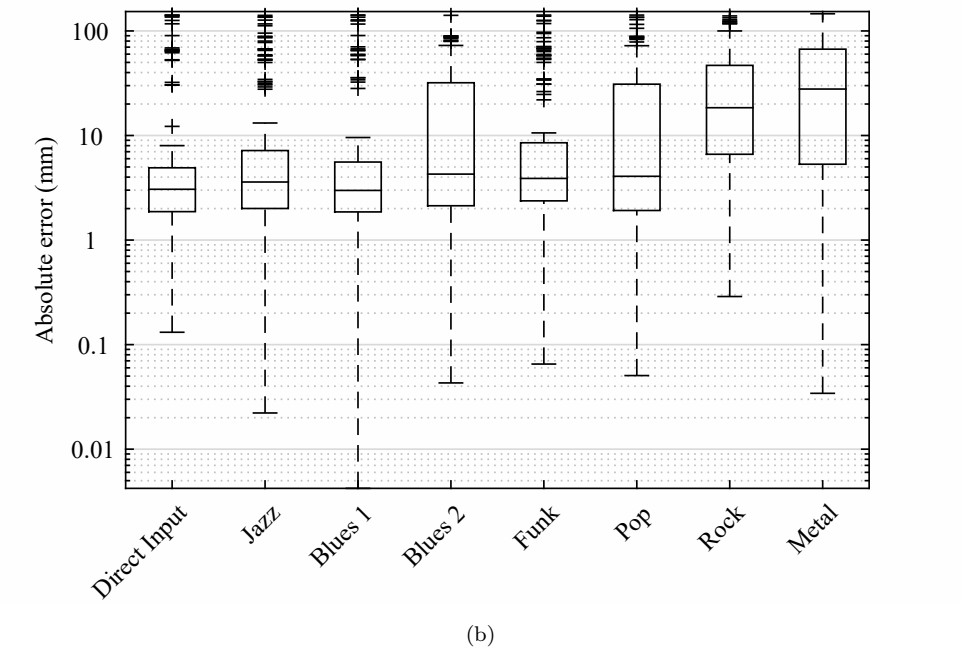
In this section, the effects of different combinations of emulated electric guitar effects, amplifier, loudspeaker and microphone on the estimates are examined. The audio samples mentioned previously in Section 5.2.1 that are used in this experiment consists of clean tones i.e. Direct input, Jazz, Blues 1 and Funk presets, compressed tones i.e. Pop preset, overdriven tones i.e. Blues 2 preset and heavily distorted tones i.e. Rock and Metal presets. The modified LC-SP method is used to find the estimates, where the spectrum is flattened using the PRSF method. The first 3 cycles of each tone is taken for the STFT analysis and the total number of harmonics K is set to 25.

Fig. 5.3a and 5.3b show the absolute errors for pickup and plucking position estimates respectively. Overall, the median absolute errors for pickup position estimates are less than 8 mm ranging from 2.04 mm – 7.83 mm, while the median absolute errors for plucking point estimates are less than 30 mm ranging from 2.98 mm – 27.81 mm. The outliers for the pickup and plucking position estimates increase when audio effects are applied to the direct input guitar signals.

It seems that the choice of guitar amplifiers, loudspeakers and microphones does not have a significant impact on the pickup position estimates shown in Fig 5.3a. For instance, presets Blues 2 and Pop have the same loudspeaker cabinet, microphone model and placement but have different amplifiers where one is the British Tube Lead 1 emulating a Marshall JCM800 and the other is the American Tube Clean 2 emulating a Fender Deluxe Reverb '65 (see



(a)



(b)

Figure 5.3: Box plot of absolute error in (a) pickup position and (b) plucking point estimates for each preset. Note that the y-axis is in log scale.

Fig. 5.2). Both presets show similar results where the median absolute errors for pickup position estimates are 3.23 mm and 3.14 mm respectively.

In Fig. 5.3a, the third quartiles for most presets are less than 10 mm which suggests that the pickup position estimates are robust to most presets. The errors for pickup position estimates are mostly affected by the increasingly distorted signals. Fig. 5.3a shows that the median absolute errors slightly increased when the guitar signal is overdriven or compressed, and the median absolute errors significantly increased when the guitar signal is heavily distorted. For plucking point estimates, a similar trend can be seen in Fig. 5.3b, where errors increase as the electric guitar signal is more distorted. This is due to distortion harmonics introduced by the nonlinear effects which tend to fill in the important nulls and affect relevant spectral peaks at higher harmonics. So, this will affect estimates that are closer to the bridge, where the nulls created by the pickup or pluck near the bridge are at high harmonics. Fig. 5.4 shows an example of the effects of mild and heavy soft clippings on the log-correlation. The guitar is plucked at 50 mm from the bridge which suppresses around every 13th partials ($L = 651$ mm). The spectra in Fig. 5.4b show that the 13th, 14th and 15th partials are greatly affected by the clipping distortions. Fig. 5.4c shows that the depth of the trough near lag τ_p reduces as the signal gets more distorted, and the depth of the trough near lag τ_d is not affected as much.

The plucking position estimates are more strongly affected, for instance, the interquartile range of the errors in Blues 2 (overdriven sound) and Pop (compressed sound) presets are larger compared to the pickup position errors. For mild signal clippings, the depth of the expected troughs that are close to zero lag (which correspond to positions that are close to the bridge) are reduced affecting the results. In practice, signals are clipped differently depending on the amplifier and its settings which affects the spectral peaks in various ways. These variations could also introduce unwanted troughs in the log-correlation. Unfortunately, some troughs that correspond to 30 – 70 mm plucking positions in the Blues 2 and Pop data are not identified correctly which are due to either the troughs are above the 40% threshold (see Section 4.2.3 page 97) or higher than the unwanted troughs (which falsely selects the unwanted troughs instead). So, the plucking positions have more data points that are close to the bridge and are affected by the signal clipping than the pickup position. Thus, the plucking position results have more estimates with large errors causing a larger interquartile range compared to the pickup position results.

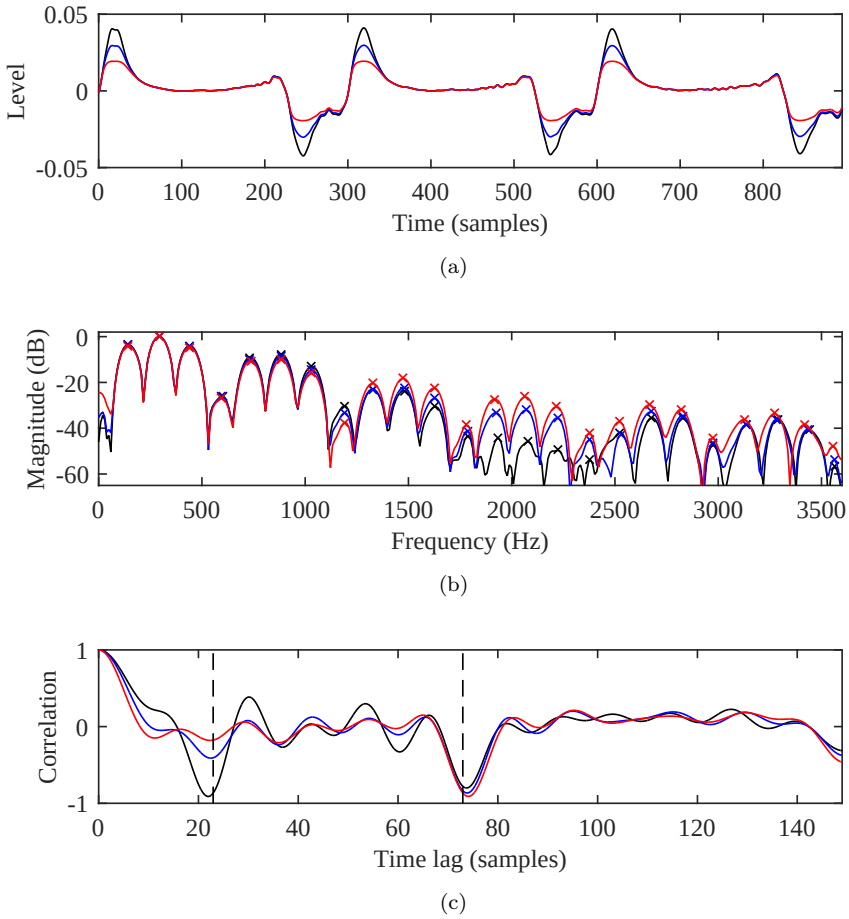


Figure 5.4: (a) The first 3 cycles of an electric guitar tone (black) that is plucked on the open 4th string at 50 mm from the bridge with a pickup at 159 mm. The same tone with mild (blue) and heavy (red) soft-clippings. (b) The spectra of the tone (black) with mild (blue) and heavily (red) soft-clippings, and the crosses represents the identified spectral peaks. (c) The log-correlation of the tone (black) with mild (blue) and heavy (red) soft-clippings. The vertical dashed lines (from left to right) at lags τ_ρ and τ_d correspond to the plucking and pickup positions respectively.

5.2.3 Comparing Spectral Flattening Methods

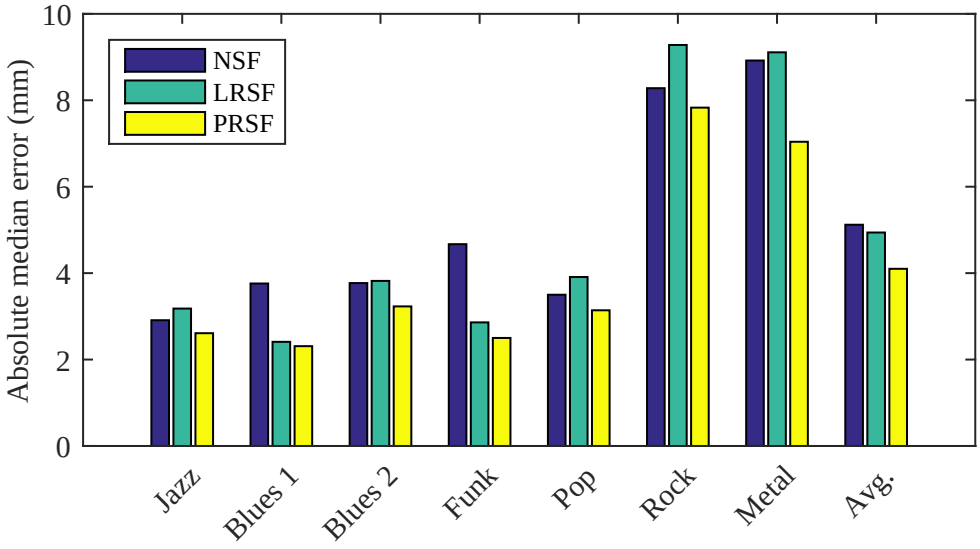
Finally, the experiments are repeated using the Linear Regressing Spectral Flattening (LRSF) method to flatten the spectrum and without any spectral flattening (NSF). This is to test whether the Polynomial Regression Spectral Flattening (PRSF) is making any improvements to the results for electric guitar tones with audio effects.

Figs. 5.5a and 5.5b show the median absolute errors for pickup and plucking position estimates respectively, comparing between the two spectral flattening methods and without any spectral flattening. The results show that without applying any spectral flattening, the median absolute errors for pickup and plucking position estimates are higher on average compared to using spectral flattening methods. However, the errors for NSF and LRSF methods are quite random as to which one is better. After performing the paired-sample t-test, NSF and LRSF produces significantly different pickup and plucking position errors. So for some presets, LRSF performs better than without applying any spectral flattening.

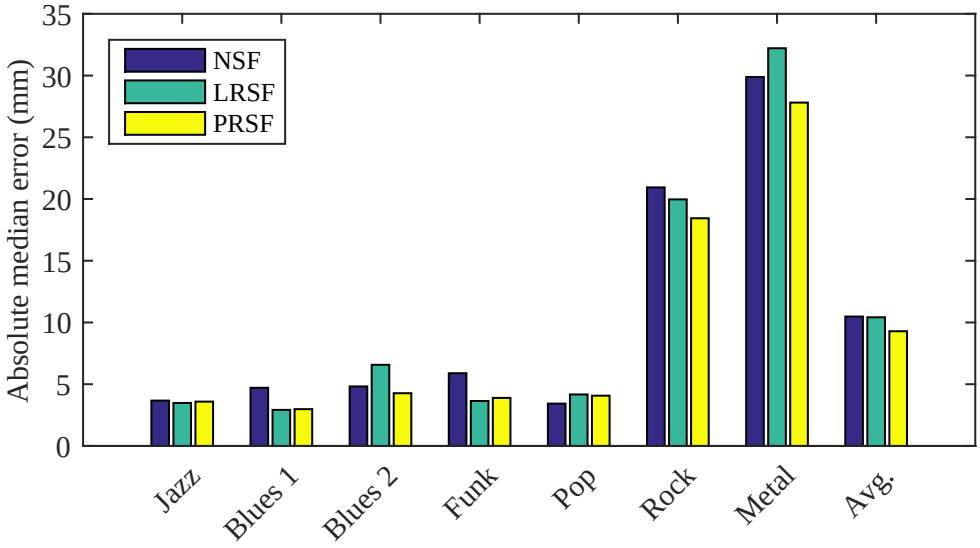
The PRSF method is proposed to mitigate the effects of guitar amplifiers, loudspeakers, microphones and post-processing effects. The results show that PRSF performs the best with the lowest average error compared to others, which suggests that this spectral flattening method works better for electric guitar tones with audio effects. Furthermore, PRSF has lower median pickup position errors for all presets compared to other approaches. In addition to lower errors, the results for PRSF are statistically different from LRSF and NSF after performing the t-test. Using PRSF does improve the results compared to using the LRSF method, where the average median absolute errors across all presets for pickup and plucking position estimates decrease by 0.84 mm and 1.13 mm respectively.

5.3 Test on Real World Recordings

The pickup and plucking positions of real world guitar signals are estimated in this section using the modified LC-SP method, where the spectrum is flattened using the PRSF method. Two recordings are taken for this experiment which are a live performance of ‘Love Me Two Times’ played by The Doors and a studio recording of ‘Day Tripper’ played by The Beatles. A technique to distinguish between pickup from pluck estimates is described in Section 5.3.1, where information from successive notes is used to distinguish between the two. After estimating the pickup positions, the estimates are used to identify which guitar is played and its pickup selection, which will be discussed in Section 5.3.2.



(a)



(b)

Figure 5.5: The median absolute errors for (a) pickup and (b) plucking positions for each preset using the two spectral flattening methods (LRSF and PRSF) and without any flattening method (NSF).

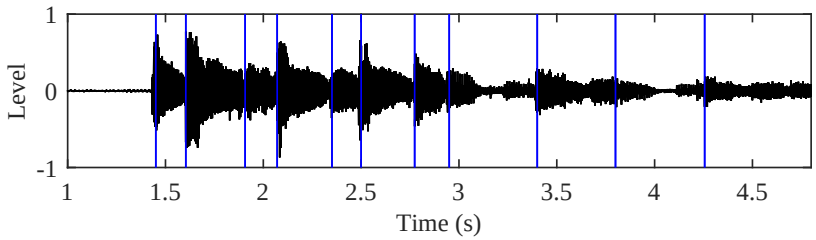


Figure 5.6: Excerpt from a guitar signal (black) played by Robby Krieger of The Doors in a live performance and the estimated onset times (blue).

5.3.1 Distinguishing Between Pickup and Pluck Estimates

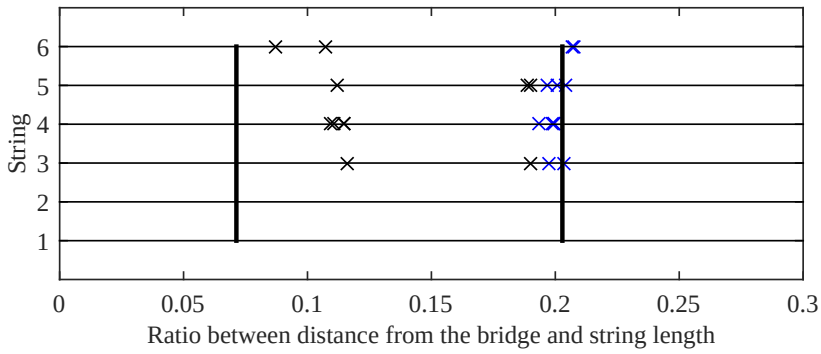
As mentioned in Section 4.2.2, the pickup and plucking positions create similar effects in the autocorrelation and it is not possible to distinguish between the two without any prior information or information from successive notes. This section describes how the two can be distinguished after estimating $\hat{R}_1$ and $\hat{R}_2$ from each note in a song.

The Doors - Love Me Two Times

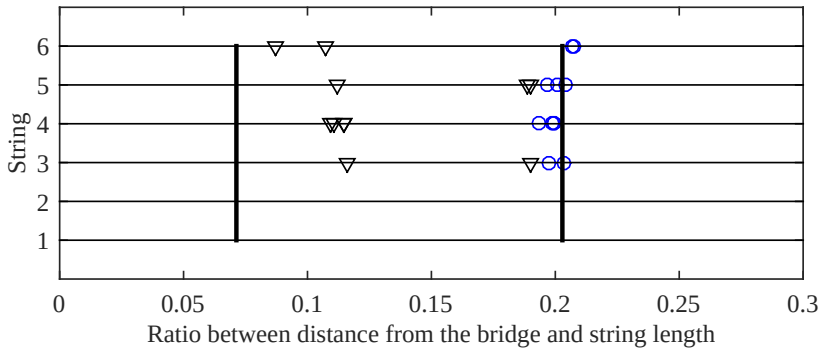
Firstly, the song 'Love Me Two Times' is taken as an example to describe the technique to distinguish between pickup and pluck estimates using information from the notes that are under test. The audio signal that is under test is taken from The Doors' live television performance at Copenhagen, Denmark in 1968 (The Doors, 2002). The first 4.8 seconds of the audio signal (which only has the guitar riff played by Robby Krieger) is taken for analysis as shown in Fig. 5.6. Note that the audio signal has two channels and the guitar tracks from both channels are similar. The left channel is taken to estimate the guitar's pickup and plucking positions. Similar results are obtained for the right channel.

The onset times are estimated using the technique described in the LC-SP method in Section 4.2.3, where the window size is 8192 samples and the hop size is 2048 samples. The peaks of the onset detection function that are above 20% of the maximum peak suggest the onset times, and the estimated onset times are shown as the blue lines in Fig. 5.6.

The strings and frets played by Robby Krieger are known, which makes it possible to normalise the estimates as discussed in Section 4.5. For each note, the modified LC-SP method finds the $\hat{R}_1$ and $\hat{R}_2$ estimates, where the first 3 cycles are analysed and K is set to 25. The subscripts of $\hat{R}$ denote their group numbers, where $\hat{R}_1$ denotes the estimated position that is closer to the bridge than the other. In Fig. 5.7a, the $\hat{R}_1$ estimates are



(a)



(b)

(c)

Figure 5.7: (a) $\hat{R}_1$ (black) and $\hat{R}_2$ (blue) estimates of the electric guitar in the live performance of 'Love Me Two Times', and (b) the distinguished pickup (circle) and plucking position (triangle) estimates. The vertical thick lines from left to right are the bridge and neck pickup positions of a 1964 Gibson SG Special. (c) An excerpt of the live performance of The Doors showing the plucking positions of Robby Krieger on the 1964 Gibson SG Special with P-90 pickups.

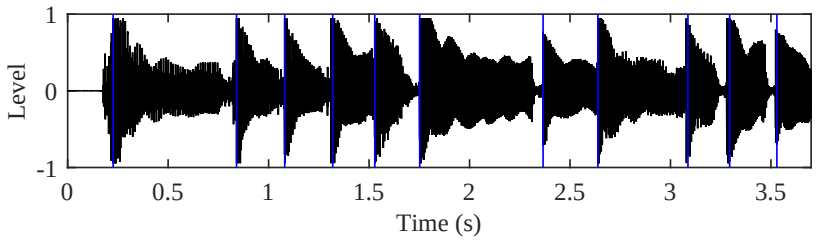
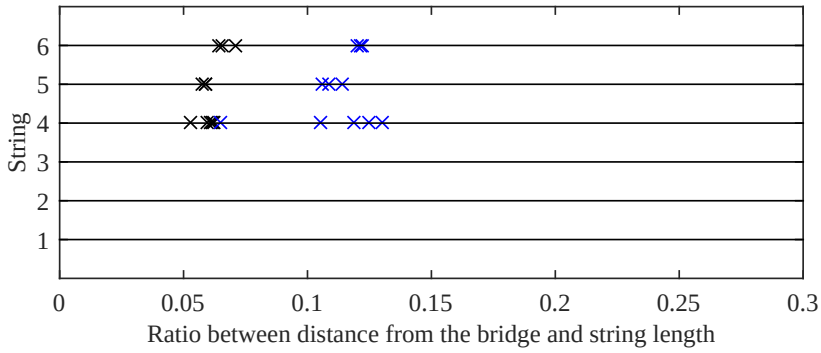


Figure 5.8: Excerpt from a guitar signal (black) played by George Harrison of The Beatles taken from the album '1' and the estimated onset times (blue).

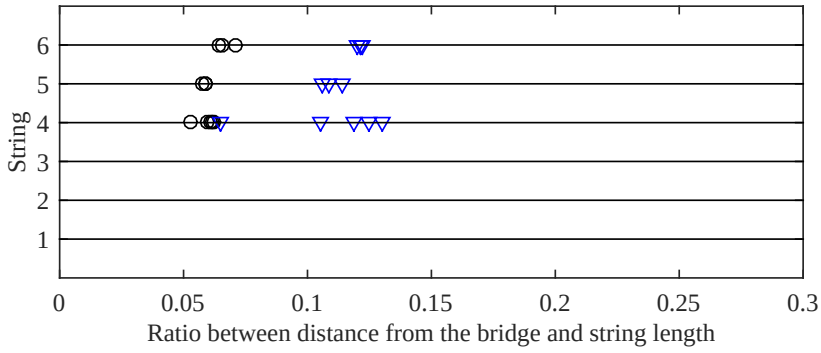
shown as black crosses and $\hat{R}_2$ estimates are shown as blue crosses. For each string that is played, the absolute differences between each estimate in each group are calculated. The median absolute deviation of the estimates in each group is chosen instead of the mean absolute deviation to prevent outliers from affecting the results. Then, the averages of all median absolute deviations across all strings for group 1 and 2 are determined. It is assumed that the plucking point typically varies for each pluck while the pickup remains fixed at a location. Therefore, the group that has the highest average contains the plucking positions and the other is the pickup position. Fig. 5.7b shows the distinguished estimates where the pickup position estimates are shown as circles and the plucking point estimates are shown as triangles. In the video recording of the live performance, Robby Krieger mostly plucks the strings in between his guitar's pickups, where one example is shown in Fig. 5.7c, and the electric guitar model used is a 1964 Gibson SG Special. The plucking point estimates shown in Fig. 5.7b are mostly located in between the two pickups which is where Robby Krieger plays. The pickup selection is unknown, and the pickup selector cannot be seen in the video which means that the estimates can help determine the pickup used in this performance. The pickup position estimates appear to be very close to the neck pickup position compared to other target locations i.e. the bridge pickup position and the mid-point between two pickups which strongly suggest that Robby Krieger uses that pickup in this live performance. The median absolute distance between the pickup position estimates and the neck pickup position is 2.56 mm.

The Beatles - Day Tripper

For the next example, a studio recording of the song 'Day Tripper' by The Beatles is taken to estimate the lead guitar's pickup and plucking positions. It is worth mentioning that there



(a)



(b)

Figure 5.9: (a) $\hat{R}_1$ (black) and $\hat{R}_2$ (blue) estimates of the electric guitar in the studio recording of 'Day Tripper', and (b) the distinguished pickup (circle) and plucking position (triangle) estimates.

are three different electric guitar sounds in the recording. Two of them are played by the lead guitarist, George Harrison, which can be heard in the opening guitar riff (single notes) separated into each channel, and the other electric guitar sound is played by the rhythm guitarist, John Lennon, strumming the E chord shortly after the iconic guitar riff is played twice.

The song is taken from the album '1' (remixed and remastered version of 27 Beatles' hit singles) released in 2015 that was recorded on the 16th of October 1965 at Abbey Road Studios in London (The Beatles, 2015). In this recording, the electric guitar that is used by the lead guitarist is unknown as this detail was not recorded at the time. However, there are a few electric guitar models that have might been used during the recording session. Babiuk (2002, p. 169) mentioned that there were two Fender Stratocasters, a Rickenbacker 325 and an Epiphone Casino available during the session, while Ryan and Kehew (2006, p. 407) added that a Gibson ES-345 and a Gretsch 6120 were also available during that time. The electric guitars that are used in the 'Day Tripper' recording are not mentioned in either book (Babiuk, 2002; Ryan and Kehew, 2006), so the pickup position estimates could be used to determine the model of the guitar played by George Harrison.

There are two guitar tracks that can be heard on two separate channels during the first guitar riff. The guitar tracks were played on separate occasions (Ryan and Kehew, 2006, p. 398). The guitar track on the left channel is more distorted than the right, so the right channel is taken for analysis because less distorted signals are expected to have more accurate results. The opening guitar riff from 0 to 3.7 seconds is taken to find its pickup and plucking positions, where the guitar signal is shown in Fig. 5.8.

The onset times are estimated using a window size of 2048 samples and the hop size of 1638 samples. The peaks of the onset detection function that are above 85% of the maximum peak suggest the onset times, and the estimated onset times are shown as the blue lines in Fig. 5.8. Note that the strings and frets played by George Harrison are known beforehand. The modified LC-SP method finds the $\hat{R}_1$ and $\hat{R}_2$ estimates for each note, where the first 3 cycles are analysed and K is set to 25. Fig. 5.9a shows one of the $\hat{R}_2$ estimates (one of the blue crosses on the 4th string) located further from the centre of its group which could possibly be an outlier in the data.

The pickup and plucking position estimates are then identified using the technique described for the previous example, and the distinguished estimates are shown in Fig. 5.9b. The pickup position estimates appear to be slanted, much like the bridge pickup of a Fender

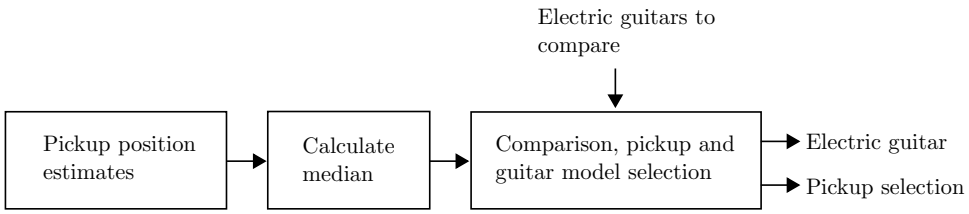


Figure 5.10: Block diagram for electric guitar model and pickup selection identification.

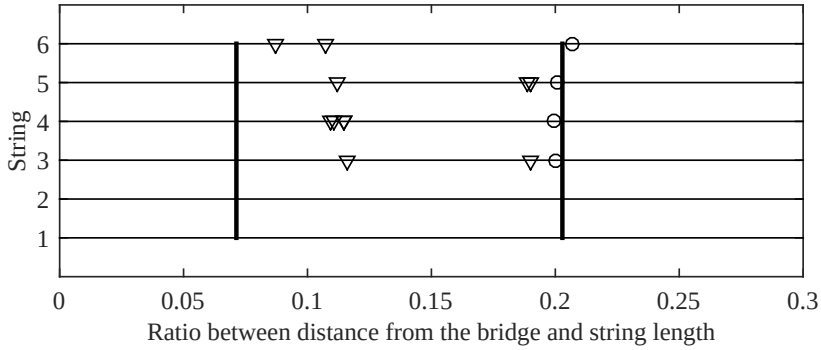


Figure 5.11: The median of pickup position estimates for each string (circles) and the plucking position estimates (triangle) of the electric guitar in the opening riff of 'Love Me Two Times'. The vertical thick lines are the pickup positions of a 1964 Gibson SG Special.

Stratocaster. The next section will discuss the identification of the electric guitar model and its pickup selection based on the pickup position estimates.

5.3.2 Identification of Guitar and its Pickup Selection

This section discusses how to identify the electric guitars that were used in the recordings and their pickup selections based on their pickup position estimates. The process of electric guitar and pickup selection identification is shown in Fig. 5.10.

Firstly, the outliers of the pickup position estimates could affect the identification, thus, the median of all pickup position estimates for each string is calculated to avoid bias from outliers. Fig. 5.11 and 5.12 show the median pickup position estimates for each string played in 'Love Me Two Times' and 'Day Tripper' recordings respectively.

The electric guitar that was used by Robby Krieger is a 1964 Gibson SG Special. In order to identify the electric guitar's pickup selection, the absolute distances between the median pickup position estimate and the target locations (i.e. bridge pickup, neck pickup and the

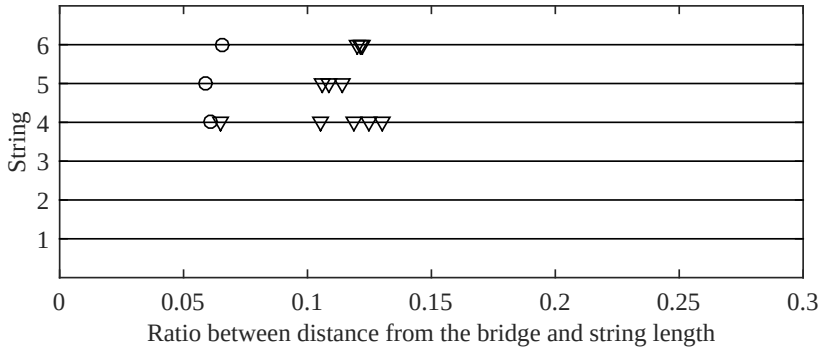


Figure 5.12: The median of pickup position estimates for each string (circles) and the plucking position estimates (triangle) of the electric guitar in the opening riff of 'Day Tripper'.

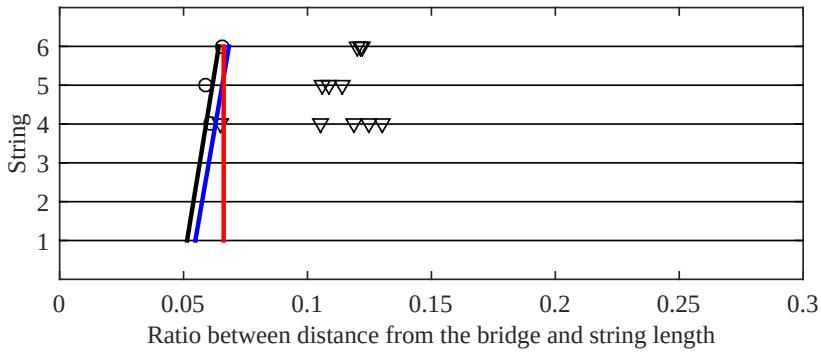


Figure 5.13: The median of pickup position estimates for each string (circles) and the plucking position estimates (triangle) of the electric guitar in the opening riff of 'Day Tripper' with the fitted line (black) and bridge pickup positions of a Fender Stratocaster (blue) and an Epiphone Casino (red).

mid-point between the two pickups) for each string is calculated, and averaged across each string played. The minimum average distance between the estimates and the target locations identifies the pickup selection. Table 5.1 shows the average distances between the estimates and the target locations of an 1964 Gibson SG Special. The distance between the estimates and the neck pickup is the smallest which strongly suggests that Robby Krieger uses the neck pickup position in the live performance.

As described previously in Section 5.3.1, there are several electric guitars that are reported to have been available during the recording session of 'Day Tripper'. In order to identify the electric guitar model, the absolute distances between the median pickup position estimate and the target locations of several electric guitars for each string is calculated, and averaged across each string played. Note that the target locations of an electric guitar are its R_d and R_α distances. The pickup position estimates, as shown in Fig. 5.12, are compared with the target locations of a Fender Stratocaster, Rickenbacker 325, Epiphone Casino, Gibson ES-345 and Gretsch 6120. The minimum distance between the estimates and the target locations of an electric guitar identifies its pickup selection, and the minimum distance between the estimates and the target locations of all electric guitar models finds the predicted electric guitar model used in the recording.

Table 5.2 shows the distances between the estimates and the target locations, where the pickup position estimates are closest to the bridge pickup of a Fender Stratocaster. This suggests that George Harrison used a Fender Stratocaster in the bridge pickup position during the recording session. The Epiphone Casino is almost the same distance due to one of the pickup position estimates has a very small error which overall makes the distance almost as small as the Fender Stratocaster. The pickup position estimates appear to be slanted while the bridge pickup of an Epiphone Casino is straight. A better measurement could perhaps take into account the angle of the pickup position estimates. By fitting a line to the pickup position estimates, it can estimate the pickup position on the 1st, 2nd and 3rd strings as shown in Fig. 5.13. The average distance between pickup position estimates from the fitted line and the bridge pickups of the Fender Stratocaster and Epiphone Casino can then be calculated. The average distance for the Fender Stratocaster and Epiphone Casino is now 2.4 mm and 5.3 mm respectively.

Closest target location	Avg. absolute distance (mm)
Neck pickup	2.2
Neck + bridge pickup	40.5
Bridge pickup	82.0

Table 5.1: Love Me Two Times song: the average absolute distance between the closest target locations of a 1964 Gibson SG Special and the pickup position estimates. The pickup selections are sorted in ascending order based on the average absolute distance.

Guitar model	Closest target location	Avg. absolute distance (mm)
Fender Stratocaster	Bridge pickup	2.5
Epiphone Casino	Bridge pickup	2.9
Rickenbacker 325	Bridge pickup	4.7
Gibson ES-345	Bridge pickup	6.2
Gretsch 6120	Bridge pickup	7.2

Table 5.2: Day Tripper song: the average absolute distance between the closest target locations of selected electric guitars and the pickup position estimates. The electric guitars are sorted in ascending order based on the average absolute distance.

5.4 Summary

The effects of various guitar signal chains on the pickup and plucking position estimates are discussed, where the errors of the estimates increase as the guitar signal is more distorted. The pickup and plucking position estimates are very accurate for clean, overdriven and compressed tones with median absolute errors ranging from 2.31 mm to 3.23 mm and 2.98 mm to 4.27 mm respectively. For distorted tones, the median absolute errors are less than 8 mm. The pickup and plucking position estimates are less robust to distorted tones than clean, slightly overdriven and compressed tones because more distortion harmonics are introduced and tend to fill in the important nulls at higher harmonics making troughs in the log-correlation near zero lag shallower. This strongly affects the estimates that are closer to the bridge when the guitar tones are heavily distorted.

The PRSF method does improve the results when an audio effects chain i.e. guitar effects, amplifier, loudspeaker and microphone is applied to the guitar signal. The PRSF method performs the best compared to using the LRSF method and compared to not using any spectral flattening method.

Furthermore, the modified LC-SP method is then tested on two real-world recordings to estimate the electric guitar's pickup and plucking positions. The estimates for each note are grouped into $\hat{R}_1$ and $\hat{R}_2$ (the subscripts of $\hat{R}$ denote their group numbers), where the estimate that is closest to the bridge is in group 1. This clustering method could affect the results for plucking positions are on top of the pickup, where the pluck can vary around the pickup. Further investigations can look into other clustering methods to group $\hat{R}_1$ and $\hat{R}_2$ estimates and test on cases where the plucking position is near the pickup. The pickup and plucking position estimates can be distinguished by assuming that the plucking point varies for each note while the pickup position remains fixed. So, the variations in the estimates for each group are calculated, and the group with estimates that vary the most is the plucking position estimates.

The method proposed to distinguish between the estimates is correct for the 'Love Me Two Times' example. The video clip of the live performance shows that the player plucks the strings in between the two pickups which agrees with the plucking point estimates. The pickup position estimates are shown to be closer to the neck pickup of a 1964 Gibson SG Special than other electric guitars.

The second example is 'Day Tripper' by The Beatles. The electric guitar and its pickup

selection are unknown, and the pickup position estimates are used to identify which guitar model is played by George Harrison in the recording. The pickup position estimates are shown to be closer to a bridge pickup of a Fender Stratocaster than any of the electric guitars that were reported to be available at that time. This shows that the pickup and plucking position estimation can find applications such as identifying the electric guitar model, its pickup selection and estimating the plucking points on historical recordings.

In these two real-world examples, a set of possible guitars (and their measurements) is assumed to be known. This would not be a problem for most famous artists, where their set of guitars is well recorded. Musicologists could have information about the musician's set of guitars narrowing down the search and help with the identification of the guitar model. Furthermore, the tests are performed on a separate guitar without any other musical instruments interfering with the estimation. Further tests could investigate on how much will other musical instruments e.g. drums, electric guitars, bass guitars and pianos affect the outcome of the estimation. Moreover, the real-world examples contain single notes, and if a guitar signal with multiple notes played simultaneously (e.g. chords) is given, only the first 2 cycles of each note for fast strums are analysed.

Chapter 6

Conclusions

This work has reached its main goal as discussed in Section 1.2, where methods are proposed in Chapter 3, Chapter 4 and Chapter 5 to extract meaningful information from electric guitar recordings. The main motivation behind this research is to replicate the sound of popular guitarists from their recordings, so the focus of this research is extracting relevant parameters from an electric guitar signal to achieve that ambition. Two concepts have been explored; one is about transforming an input guitar sound into a target sound using a linear filter, and the other is about providing parameters for electric guitar synthesis.

In this chapter, the main achievements of this research are highlighted in Section 6.1, and the thesis is concluded with possible directions for future work in Section 6.2.

6.1 Summary

6.1.1 Estimating Filter Coefficients from Electric Guitar Recordings

The research question about how to directly transform a user's guitar sound into another was addressed. A preliminary step was taken, where a method to find coefficients for the FIR filter that transforms the sound of a pickup into another on the same guitar was developed. The filter coefficients were estimated using a least squares estimator given an input signal and a target signal. The method yields good results, where the listening test suggests that the difference between the sound similarity ratings of the hidden reference and the estimated sound is not statistically significant. This suggests that the listeners find it difficult to distinguish between the hidden references and the estimated sounds. Also, the results are supported with a spectral difference measurement, where up to 99% of the spectral difference between the input and target signals is reduced by the learnt filter. However, this method provides some limitations to replicating the target signal in the practical use case, such a guitar synthesiser with a hexaphonic pickup, where the learnt filter is applied to the input. Changing the plucking position, dynamic and fret position will affect the accuracy of the emulation. Random differences between repetitions will have a small degradation on the accuracy. The learnt filters are less robust when the input changes plucking positions and dynamics, and not at all robust when the input played a fret position that was not played by the target signal. This means that the method produces a very accurate emulation but has some limitations, where a user must play the same plucking position, dynamics and notes (string and fret) as the target guitar recording to avoid affecting the emulation.

6.1.2 Estimating Playing Parameters from Electric Guitar Recordings

Another approach to replicate the target sound is by using electric guitar synthesis. In this case, the parameters for electric guitar synthesis need to be known. This approach allows the user the flexibility of changing the plucking position, dynamics and notes. There is other relevant information that can be extracted from an electric guitar recording such as its pickup and plucking positions which can act as synthesis parameters. Three frequency domain approaches are proposed to estimate the pickup and plucking positions on an electric guitar which were described in Chapter 4. The three approaches are called the Spectral Peaks (SP), Autocorrelation of Spectral Peaks (AC-SP) and Log-correlation of Spectral Peaks (LC-SP) methods.

A comparison between the three methods is made on single pickup data, where the SP performs the worst and the AC-SP and LC-SP methods produce accurate estimates. The median absolute errors for pickup position estimates using the SP, AC-SP and LC-SP methods are 5.00 mm, 2.00 mm and 1.97 mm respectively, and the median absolute errors for plucking point estimates are 7.00 mm, 3.00 mm and 2.73 mm respectively. The autocorrelation methods produce accurate results, but in practice, errors of 2 mm and above are "just" noticeable or audible. Nonetheless, the median errors are less than 4 mm, meaning that the differences are not significant.

The autocorrelation methods i.e. the AC-SP and LC-SP methods are then tested on mixed pickup data, where the position in between the two pickups α and the plucking point are estimated for each signal. The median absolute errors for α estimates using AC-SP and LC-SP are 5.50 mm and 2.52 mm respectively, and the median absolute errors for plucking point estimates are 7.00 mm and 6.52 mm respectively.

The effects of plucking dynamics and fret positions on the estimates are then tested using the two autocorrelation methods. For the effects of plucking dynamics, the number of outliers increased for louder and softer tones. Nevertheless, the median absolute errors for pickup position estimates using AC-SP and LC-SP are less than 4 mm, and the median absolute errors for plucking point estimates are less than 10 mm and 5 mm respectively. As for the effect of fret positions, both autocorrelation methods produce pickup position estimates where almost all of them have errors less than 10 mm. This suggests that the pickup position estimates are highly robust to changes in fret position. The number of

outliers of the plucking point estimates increase with the fret position, yet, the median absolute errors are less than 5 mm. However, it was reported that any pickup or plucking position that goes beyond $L_{\mathcal{F}}/2$ cannot be estimated. For multiple notes e.g. chords, the second note is struck after a few cycles of the first note, so there is a small window that can be analysed without the second note affecting the results.

The effects of various guitar signal chains on the pickup and plucking position estimates are tested using the LC-SP method for its faster computation compared to the AC-SP method. Different guitar signal chains i.e. different emulated guitar effects, amplifiers, loudspeakers and microphones are applied to the direct input guitar signal. A modified LC-SP method is introduced to compensate the effects of the guitar signal chain. The errors for pickup and plucking position estimates increase as the guitar signal gets heavily distorted. For clean, overdriven and compressed tones, the pickup position estimates are very accurate with median absolute errors ranging from 2.31 mm to 3.23 mm – which the errors are not noticeable in practice, and for distorted tones, the median absolute errors are less than 8 mm. The plucking position estimates are shown to be less robust to distorted tones than clean, overdriven and compressed tones and the median absolute error can lead up to 28 mm. On the plus side, the plucking point estimates are very accurate for clean, overdriven and compressed tones with median absolute errors ranging from 2.98 mm to 4.27 mm. The spectral flattening methods that is introduced i.e. the LRSF and PRSF are reported to improve the results. For all guitar signal chains that are tested, the PRSF method yields lower median absolute errors on average compared to using the LRSF method and without using any spectral flattening method.

Estimating the performance parameters could lead to several possible applications. The pickup locations and angles of popular guitars are quite distinct. Therefore, accurate pickup position estimates could help musicologists or guitar enthusiasts to determine the electric guitar model and pickup selection used in historical recordings where there is insufficient information about the original instrument. Conversely, knowledge from musicologists could be used to distinguish pickup from plucking position estimates. For instance, given a recording of a well-known guitarist where musicologists know that the player has a tendency of plucking the strings near the bridge, the other estimates could be the pickup position. In addition, the pickup and plucking position estimates could be used as parameters for electric guitar synthesis e.g. MIDI guitars or guitar synthesisers with hexaphonic pickups, which opens the possibility of replicating the sound of popular guitarists by extracting relevant parameters

from their recordings.

6.2 Future Work

This section provides a selection of ideas for further work that could be investigated from this research, which includes:

6.2.1 Replicating Electric Guitar Sounds from Audio Recordings

In this research, two concepts are explored in order to replicate a guitar sound from an audio recording: 1) morph the user's guitar sound into another using digital filters and 2) extract relevant information from the guitar sound. However, there are some steps required in order to achieve that ambition, so further work can look into this.

Morphing a Guitar Sound into Another using Digital Filters

By using this concept, an electric guitar with a hexaphonic pickup is required enabling the pickup to output signals from 6 strings separately. So, these signals can be processed individually without affecting the emulation. Mohamad et al. (2015) propose a technique to transform a pickup sound into another on the same guitar using an FIR filter, where the limitations of using this concept are discussed. If the user does not require the flexibility of playing a different plucking dynamic, pickup position and note as the target sound, using this concept may be suitable for the user. This work can be extended, where the technique can be performed on transforming a guitar sound into another e.g. morphing the sound of a Fender Stratocaster into a Gibson Les Paul. Furthermore, further work can look into improving the learnt filters. The filters do not learn anything in between the spectral peaks, so estimating the frequency response in between the partials by smoothing to obtain the spectral envelope could improve the filters. Thus far, only the coefficients for linear filters are estimated. An electric guitar sound has a nonlinear behaviour, so, further investigation can work on estimating the coefficients for nonlinear filters.

Extracting Relevant Information from the Guitar Sound

Further work can investigate on extracting all information that is required in order to replicate the desired guitar sound using machine learning techniques such as deep learning. This

means that all the component in the guitar signal chain (i.e. electric guitar model, effects, amplifier, loudspeaker and microphone) and its settings should be known. This information can help guide the user to purchase the exact model of each component.

Section 5.3.2 describes a technique to identify the electric guitar model used in a real-world recording using the pickup position estimates – which relies heavily on one feature for training. Further work could look into adding more features that can help identify the electric guitar model e.g. pickup width, brightness and dead-spots. Also, further investigations could look into using other techniques to identify electric guitar models without extracting features from recordings.

In-phase mixed pickups such as a humbucker pickup suppresses the noise at 50 Hz due to power supply interference. Further investigation could look into identifying in-phase mixed pickup signals. An electric guitar signal that has a low amplitude noise at 50 Hz could be identified as an in-phase mixed pickup rather than a single pickup or an out-of-phase mixed pickup. The ENF detection should test on several electric guitars, particularly the ones with different pickup configurations e.g. single, in-phase mixed pickups (humbuckers) and out-of-phase pickups. Also, different recording conditions e.g. in a studio and outdoor performances should also be considered, where the distance from the pickups to a power supply may vary the energy of the 50 Hz hum.

Bibliography

J. S. Abel and D. P. Berners. A technique for nonlinear system measurement. In *Audio Engineering Society Convention 121*. Audio Engineering Society, 2006.

J. Abeßer. Automatic string detection for bass guitar and electric guitar. In *International Symposium on Computer Music Modeling and Retrieval*, pages 333–352, 2012.

T. Araya and A. Suyama. Sound effector capable of imparting plural sound effects like distortion and other effects, Oct. 29 1996. US Patent 5,570,424.

A. Askenfelt and E. V. Jansson. From touch to string vibrations. iii: String motion and spectra. *The Journal of the Acoustical Society of America*, 93(4):2181–2196, 1993.

A. Babiuk. *Beatles gear: all the fab four's instruments from stage to studio*. Backbeat Books, 2002.

I. Barbancho, L. J. Tardon, S. Sammartino, and A. M. Barbancho. Inharmonicity-based method for the automatic generation of guitar tablature. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(6):1857–1868, 2012.

E. Barbour. The cool sound of tubes [vacuum tube musical applications]. *IEEE Spectrum*, 35(8):24–35, 1998.

D. Barchiesi and J. Reiss. Automatic target mixing using least-squares optimization of gains and equalization settings. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 7–14, 2009.

D. Barchiesi and J. Reiss. Reverse engineering of a mix. *Journal of the Audio Engineering Society*, 58(7/8):563–576, 2010.

B. A. Bartlett. Tonal effects of close microphone placement. *Journal of the Audio Engineering Society*, 29(10):726–738, 1981.

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*, 13(5):1035–1047, 2005.

S. Bilbao and A. Torin. Numerical simulation of string/barrier collisions: The fretboard. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 137–144, 2014.

S. Bilbao and A. Torin. Numerical modeling and sound synthesis for articulated string/fretboard interactions. *Journal of the Audio Engineering Society*, 63(5):336–347, 2015.

K. Bradley, M.-H. Cheng, and V. L. Stonick. Automated analysis and computationally efficient synthesis of acoustic guitar strings and body. In *IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, pages 238–241, 1995.

W. Bussey and R. Haigler. Tubes versus transistors in electric guitar amplifiers. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 6, pages 800–803. IEEE, 1981.

A. Case. Recording electric guitar – the science and the myth. *Journal of the Audio Engineering Society*, 58(1/2):80–83, 2010.

A. Case, A. Roginska, J. Matthew, and J. Anderson. Electric guitar-a blank canvas for timbre and tone. *The Journal of the Acoustical Society of America*, 133(5):3308–3308, 2013.

P. J. Celi, M. A. Doidic, D. W. Fruehling, and M. Ryle. Stringed instrument with embedded dsp modeling, 2004. US Patent 6,787,690.

D. Chadefaux, J.-L. Le Carrou, B. Fabre, and L. Daudet. Experimentally based description of harp plucking. *The Journal of the Acoustical Society of America*, 131(1):844–855, 2012.

D. Clark. High-resolution subjective testing using a double-blind comparator. *Journal of the Audio Engineering Society*, 30(5):330–338, 1982.

Cockos, Inc. Reaper: Digital audio workstation. Available online at http://www.reaper.fm/, accessed April 8, 2017.

A. J. Cooper. The electric network frequency (enf) as an aid to authenticating forensic digital audio recordings–an automated approach. In *Proceedings of the Audio Engineering Society International Conference on Audio Forensics - Theory and Practice*, 2008.

G. Cuzzucoli and V. Lombardo. A physical model of the classical guitar, including the player's touch. *Computer Music Journal*, 23(2):52–69, 1999.

A. de Cheveigné and H. Kawahara. Yin, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4):1917–1930, 2002.

S. Dixon. Onset detection revisited. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 133–137, 2006.

S. Dixon, M. Mauch, and D. Tidhar. Estimation of harpsichord inharmonicity and temperament from musical recordings. *The Journal of the Acoustical Society of America*, 131(1): 878–887, 2012.

M. Doidic, M. Mecca, M. Ryle, and C. Senffner. Tube modeling programmable digital guitar amplification system, Aug. 4 1998. US Patent 5,789,689.

K. Dosenbach, W. Fohl, and A. Meisel. Identification of individual guitar sounds by support vector machines. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, 2008.

G. Evangelista. Physical model of the string-fret interaction. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 345–351, 2011.

G. Evangelista and F. Eckerholm. Player–instrument interaction models for digital waveguide synthesis of guitar: Touch and collisions. *IEEE transactions on audio, speech, and language processing*, 18(4):822–832, 2010.

H. Fleischer and T. Zwicker. Investigating dead spots of electric guitars. *Acta Acustica united with Acustica*, 85(1):128–135, 1999.

H. Fletcher. Normal vibration frequencies of a stiff piano string. *The Journal of the Acoustical Society of America*, 36(1):203–209, 1964.

N. H. Fletcher. Plucked strings – a review. *Catgut Acoustic Society Newsletter*, 26:13–17, 1976.

W. Fohl, A. Meisel, and I. Turkalj. A feature relevance study for guitar tone classification. In *The International Society of Music Information Retrieval (ISMIR)*, pages 211–216, 2012.

S. Foster. Impulse response measurement using Golay codes. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 11, pages 929–932. IEEE, 1986.

T. Goetze. *CAPS Audio Plugin Suite*. 2017. Available online at: https://quitte.de/dsp/caps.html.

C. Grigoras. Digital audio recording analysis–the electric network frequency criterion. *International Journal of Speech Language and the Law*, 12(1):63–76, 2005.

F. Gustaffsson, P. Connman, O. Öberg, N. Odelholm, and M. Enqvist. System and method for simulation of non-linear audio equipment, Apr. 24 2012. US Patent 8,165,309.

D. E. Hall and A. Askenfelt. Piano string excitation v: Spectra for real hammers and strings. *The Journal of the Acoustical Society of America*, 83(4):1627–1638, 1988.

R. O. Hamm. Tubes versus transistors – is there an audible difference. *Journal of the audio engineering society*, 21(4):267–273, 1973.

N. G. Horton and T. R. Moore. Modeling the magnetic pickup of an electric guitar. *American Journal of Physics*, 77(2):144–150, 2009.

A. Hoshiai. Musical tone signal forming device for a stringed musical instrument, 1994. US Patent 5,367,120.

G. Hua, G. Bi, and V. L. L. Thing. On practical issues of electric network frequency based audio forensics. *IEEE Access*, 5:20640–20651, 2017.

IK Multimedia. Amplitube custom shop. Available online at https://www.ikmultimedia.com/products/amplitubecs/, accessed April 8, 2017.

C. Issanchou, J.-L. Le Carrou, C. Touzé, B. Fabre, and O. Doaré. String/frets contacts in the electric bass sound: Simulations and experiments. *Applied Acoustics*, 129:217–228, 2018.

D. A. Jaffe and J. O. Smith. Extensions of the Karplus-Strong plucked-string algorithm. *Computer Music Journal*, 7(2):56–69, 1983.

H. Järveläinen and M. Karjalainen. Perceptibility of inharmonicity in the acoustic guitar. *Acta Acustica United with Acustica*, 92(5):842–847, 2006.

H. Järveläinen and V. Välimäki. Audibility of initial pitch glides in string instrument sounds. In *Proceedings of the 2001 International Computer Music Conference*, pages 282–285, 2001.

N. Jillings, D. Moffat, B. De Man, and J. D. Reiss. Web Audio Evaluation Tool: A browser-based listening test environment. In *12th Sound and Music Computing Conference*, July 2015.

T. Jungmann. Theoretical and practical studies on the behavior of electric guitar pick-ups. Master's thesis, Helsinki University of Technology, Finland, 1994.

M. Karjalainen and J. Pakarinen. Wave digital simulation of a vacuum-tube amplifier. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 5, pages 14–19. IEEE, 2006.

M. Karjalainen, T. Mäki-Patola, A. Kanerva, and A. Huovilainen. Virtual air guitar. *Journal of the Audio Engineering Society*, 54(10):964–980, 2006.

K. Karplus and A. Strong. Digital synthesis of plucked-string and drum timbres. *Computer Music Journal*, 7(2):43–55, 1983.

C. Kehling, J. Abeßer, C. Dittmar, and G. Schuller. Automatic tablature transcription of electric guitar recordings by estimation of score and instrument-related parameters. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 219–226, 2014.

A. P. Klapuri. Multiple fundamental frequency estimation based on harmonicity and spectral smoothness. *IEEE Transactions on Speech and Audio Processing*, 11(6):804–816, 2003.

R. Kuroki and T. Ito. Digital audio signal processor with harmonics modification, Nov. 24 1998. US Patent 5,841,875.

Y. Lai, S.-K. Jeng, D.-T. Liu, and Y.-C. Liu. Automated optimization of parameters for fm sound synthesis with genetic algorithms. In *International Workshop on Computer Music and Audio Technology*, page 205, 2006.

M. Laurson, C. Erkut, V. Välimäki, and M. Kuuskankare. Methods for modeling realistic playing in acoustic guitar synthesis. *Computer Music Journal*, 25(3):38–49, 2001.

N. Lee, J. O. Smith, J. Abel, and D. Berners. Pitch glide analysis and synthesis from recorded tones. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, 2009.

K. A. Legge and N. H. Fletcher. Nonlinear generation of missing modes on a vibrating string. *The Journal of the Acoustical Society of America*, 76(1):5–12, 1984.

N. Lindroos, H. Penttinen, and V. Välimäki. Parametric electric guitar synthesis. *Computer Music Journal*, 35(3):18–27, 2011.

D. G. Manolakis, V. K. Ingle, and S. M. Kogon. *Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing*. McGraw-Hill Boston, 2000.

P. Masri. *Computer modelling of sound for transformation and synthesis of musical signals.* PhD thesis, University of Bristol, 1996.

A. Mehrabi, S. Dixon, and M. Sandler. Towards a comprehensive dataset of vocal imitations of drum sounds. In *Audio Engineering Society Workshop on Intelligent Music Production*, 2016.

Z. Mohamad, S. Dixon, and C. Harte. Digitally moving an electric guitar pickup. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 284–291. Citeseer, 2015.

Z. Mohamad, S. Dixon, and C. Harte. Pickup position and plucking point estimation on an electric guitar. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 651–655, 2017a.

Z. Mohamad, S. Dixon, and C. Harte. Estimating pickup and plucking positions of guitar tones and chords with audio effects. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 420–426, 2017b.

Z. Mohamad, S. Dixon, and C. Harte. Pickup position and plucking point estimation on an electric guitar via autocorrelation. *The Journal of the Acoustical Society of America*, 142(6):3530–3540, 2017c. doi: 10.1121/1.5016815. URL https://doi.org/10.1121/1.5016815.

M. Mustonen, D. Kartofelev, A. Stulov, and V. Välimäki. Experimental verification of pickup nonlinearity. In *Proceedings of the International Symposium on Musical Acoustics (ISMA 2014)*, volume 1, 2014.

A. Novak, L. Guadagnin, B. Lihoreau, P. Lotton, E. Brasseur, and L. Simon. Non-linear identification of an electric guitar pickup. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 5–9, 2016.

A. Novak, L. Guadagnin, B. Lihoreau, P. Lotton, E. Brasseur, and L. Simon. Measurements and modeling of the nonlinear behavior of a guitar pickup at low frequencies. *Applied Sciences*, 7(1):50, 2017.

R. C. Paiva, J. Pakarinen, and V. Välimäki. Acoustics and modeling of pickups. *Journal of the Audio Engineering Society*, 60(10):768–782, 2012.

J. Pakarinen and D. T. Yeh. A review of digital techniques for modeling vacuum-tube guitar amplifiers. *Computer Music Journal*, 33(2):85–100, 2009.

A. Paté, J.-L. Le Carrou, and B. Fabre. Predicting the decay time of solid body electric guitar tones. *The Journal of the Acoustical Society of America*, 135(5):3045–3055, 2014.

H. Penttinen. Acoustic timbre enhancement of guitar pickup signals with digital filters. Master's thesis, Helsinki University of Technology, 1996.

H. Penttinen and V. Välimäki. A time-domain approach to estimating the plucking point of guitar tones obtained with an under-saddle pickup. *Applied Acoustics*, 65(12):1207–1220, 2004.

H. Penttinen, J. Siiskonen, and V. Valimaki. Acoustic guitar plucking point estimation in real time. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 3, pages 209–212. IEEE, 2005.

E. K. Pritchard. Semiconductor emulation of tube amplifiers, Feb. 19 1991. US Patent 4,995,084.

Radiocommunication Sector of ITU (ITU-R). Method for the subjective assessment of intermediate quality levels of coding systems. Available online at https://www.itu.int/rec/R-REC-BS.1534/en, accessed March 6, 2018.

J. Rauhala, H.-M. Lehtonen, and V. Välimäki. Fast automatic inharmonicity estimation algorithm. *The Journal of the Acoustical Society of America*, 121(5):184–189, 2007.

L. Remaggi, L. Gabrielli, R. C. D. de Paiva, V. Välimäki, and S. Squartini. A pickup model for the clavinet. In *Proceedings of the 15th International Conference on Digital Audio Effects (DAFx-12), York, UK*, volume 1721, 2012.

K. Ryan and B. Kehew. *Recording the Beatles: the studio equipment and techniques used to create their classic albums*. Curvebender, 2006.

B. Santo. Volume cranked up in amp debate. *Electronic Engineering Times*, (817):24–35, 1994.

M. Senior. *Guitar Amp Recording*. 2017. Available online at: https://www.soundonsound.com/techniques/guitar-amp-recording.

J. O. Smith. *Spectral audio signal processing*. 2011. Available online at: https://ccrma.stanford.edu/∼jos/sasp/.

M. Stein. Automatic detection of multiple, cascaded audio effects in guitar recordings. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 4–7, 2010.

M. Stein, J. Abeßer, C. Dittmar, and G. Schuller. Automatic detection of audio effects in guitar and bass recordings. In *Audio Engineering Society Convention 128*. Audio Engineering Society, 2010.

C. R. Sullivan. Extending the Karplus-Strong algorithm to synthesize electric guitar timbres with distortion and feedback. *Computer Music Journal*, 14(3):26–37, 1990.

The Beatles. *Day Tripper*. In: *1* [CD]. Calderstone Productions Limited/ Apple Corps Limited, 2015.

The Doors. *Love Me Two Times*. The Doors Music Company/ Eagle Rock Entertainment Limited, 2002. Available online at https://www.youtube.com/watch?v=QdCZR9M5EKY.

J. D. Tillman. *Response effects of guitar pickup position and width.* 2002. Available online at: http://www.till.com/articles/PickupMixing/index.html.

C. Traube and P. Depalle. Extraction of the excitation point location on a string using weighted least-square estimation of a comb filter delay. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, 2003.

C. Traube and J. O. Smith. Estimating the plucking point on a guitar string. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 153–158, 2000.

University of South Dakota. Weber's law of just-noticeable difference. Available online at http://apps.usd.edu/coglab/WebersLaw.html, accessed March 21, 2018.

M. van Walstijn and J. Bridges. Simulation of distributed contact in string instruments: a modal expansion approach. In *Signal Processing Conference (EUSIPCO), 2016 24th European*, pages 1023–1027. IEEE, 2016.

A. von dem Knesebeck and U. Zölzer. Comparison of pitch trackers for real-time guitar effects. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, 2010.

D. T. Yeh and J. O. Smith. Simulating guitar distortion circuits using wave digital and nonlinear state-space formulations. *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 19–26, 2008.

D. T. Yeh, J. Abel, and J. O. Smith. Simulation of the diode limiter in guitar distortion circuits by numerical solution of ordinary differential equations. *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, pages 197–204, 2007.

D. T. Yeh, B. Bank, and M. Karjalainen. Nonlinear modeling of a guitar loudspeaker cabinet. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*, 2008.

U. Zölzer. *Digital audio signal processing.* John Wiley & Sons, 2008.