Institute for Visualization and Interactive Systems

University of Stuttgart
Universitätsstraße 38
D–70569 Stuttgart

Bachelorarbeit

Leveraging Electromyography to Enhance Musician-Instrument Interaction Using Domain-Specific Motions

Hendrik Schuff

Course of Study:	Medieninformatik
Examiner:	Prof. Dr. Albrecht Schmidt
Supervisor:	Jakob Karolus, M.Sc., Thomas Kosch, M.Sc.
Commenced:	January 27, 2017
Completed:	June 28, 2017
CR-Classification:	H.5.2

Abstract

Manual interaction tasks, such as playing a musical instrument, require certain amounts of training until users are proficient. Electromyography (EMG) can bridge this gap and is able to provide proficiency feedback without the need for supervision. EMG measures the electrical potential that is related to muscular activity and has been used in Human-Computer-Interaction (HCI) in a variety of applications. This thesis explores the usage of EMG together with domain-specific movements, such as playing guitar chords, in the context of musician-instrument interaction. This includes a review of related work, an evaluation of suitable features, and machine learning methods as well as the realization of an EMG guitar tutor system. The results of this thesis show that it is possible to classify guitar chords with an average F1-measure of 87%. We identified a trade-off between classifier-accuracy and window size, which is an important finding regarding real-time interaction. Further, we evaluated a guitar tutor system within a study. The results suggest, that electrodes and wires did not limit the participants in playing the guitar. An analysis of inter-person generalizability shows that dimensionality reduction methods can slightly increase the classifier performance. We propose further solutions to enhance the guitar tutor system from a machine learning perspective as well as from an usability perspective. Ultimately, we discuss how our findings can be transferred to related domains.

Kurzfassung

Manuelle Interaktionen, wie z.B. das Spielen eines Musikinstrumentes, erfordern ein gewisses Maß an Übung, bevor Nutzer diese sicher beherrschen. Elektromyographie (EMG) kann diesen Prozess unterstützen und den Nutzern Feedback geben, ohne diese zu überwachen. EMG misst die, durch Muskelaktivität erzeugten, elektrischen Potentiale und wurde vielfach im Rahmen der Mensch-Computer-Interaktion (MCI) benutzt. Diese Bachelorarbeit untersucht die Verwendung von EMG zusammen mit domän-spezifischen Bewegungen, wie z.B. Gitarren-Akkorden, im Bezug auf die Musiker-Instrument-Interaktion. Diese Untersuchung umfasst einen Überblick verwandter Forschung, eine Evaluierung verschiedener Merkmale (Features) und Methoden des machinellen Lernens, sowie die Entwicklung eines EMG Gitarren-Assistenz-Systems. Die Ergebnisse dieser Bachelorarbeit zeigen, dass Gitarren-Akkorde mit einem durchschnittlichen F1-Wert von 87% klassifiziert werden können. Desweiteren konnten wir eine Beziehung zwischen der Qualität der Klassifikation und der verwandten Fenster-Größe erkennen. Die Auswertung eines Gitarren-Assistenz-Systems führten wir im Rahmen einer Studie durch. Die Ergebnisse legen nahe, dass die angebrachten Elektroden und Kabel die Teilnehmer nicht beim Gitarrenspielen beeinträchtigt haben. Eine Analyse der Inter-Personen-Generalisierbarkeit zeigte, dass Methoden zur Dimensionalitätsreduktion die Qualität der Klassifikationen leicht verbessern können. Wir beleuchten Lösungen zur Verbesserung des Gitarren-Assistenz-Systems im Kontext des maschinellen Lernens als auch im Hinblick auf die Nutzbarkeit (Usability). Schlussendlich diskutieren wir mögliche Anwendungen unserer Ergebnisse auf verwandte Forschungsbereiche.

Contents

1 Introduction

Exploring the usage of biosignals is an emerging topic in the HCI community. One especially promising type of biosignal is the electrical signal that is generated by neural and muscular activities. This type of signal can be used to draw conclusions about brain activity (*electroencephalography*) and muscle activity (*electromyography*).
In the field of HCI a broad range of novel interaction techniques using EMG have been explored. These include sign-language recognition [Abr+16], wheelchair control [Moo+03] or detecting emotions using facial EMG [Vra93]. Moreover, musician-instrument-interaction has been addressed using EMG in multiple ways ([TK02] [Ars+06] [DCT+13]). However, the usage of domain-specific motions, such as playing a chord on a guitar, in combination with EMG, has not been explored yet. This bachelor thesis aims at making a first step in closing this gap. Besides conceptual insights, this topic also offers great real-world applications. These include systems that support assembly workers, help athletes to improve an exercise or tutor aspiring musicians to improve their playing beyond audible characteristics. A prototype of a guitar tutor system will be developed in the scope of this thesis.

Structure

This thesis is structured in seven chapters, each chapter's content is briefly outlined in the following. After the introduction, a presentation of related work will be given in Chapter 2. Covering an explanation of the physiological foundations of EMG and an overview of work on EMG in HCI. In Chapter 3 common analysis techniques of EMG signals will be introduced. In Chapter 4 the findings of Chapter 3 will be applied on classifying guitar chords using muscle signals of the upper forearm. Chapter 5 addresses the design and realization of an EMG guitar tutor system. Within Chapter 6 interesting questions for future work that arised throughout the work on this thesis will be highlighted. Finally a conclusion of this thesis will be given in Chapter 7.

2 Background and Related Work

This chapter briefly introduces the physiological foundations of EMG. Further, it presents work in the field of HCI that is related to this thesis.

2.1 Physiological Foundations

In order to interpret EMG signals correctly, one first has to understand how they are produced and what one can infer from a given signal regarding the muscular activity. Therefore, we start with a look at which processes and units are involved in the activation of a muscle.

Motor Unit Action Potential (MUAP)

The motor unit can be considered "[...] the basic functional unit in muscular contraction." [Hof84]. This term comprises "[the] ensemble of a motor neuron and the muscle fibers it innervates." [Hof84]. "Each muscle is made up of many motor units." [Sap+08]. A depiction of motor units and connected components is shown in Figure 2.1.
In order to activate a muscle, a signal is sent from the brain along the nervous system to the motor neurons. The motor neurons then, activate the muscle fibers to which they are connected via so called *action potentials*. These action potentials are electrical signals that cause a contraction of the muscle fibers [Sap+08].
"The sum of all the electrical activity in a motor unit during contraction is referred to as a *motor unit action potential* (MUAP)" [Sap+08]. This MUAP is what EMG measures [Sap+08]. In particular, this means that EMG precisely measures the *isometric* muscle activity [TK02], that is the contraction of the muscle without a movement of the body (e.g. pushing against a wall). In contrast to that, EMG is less appropriate to measure *isotonic* muscle activity, that is a contraction with a motion but no change in tension [TK02].

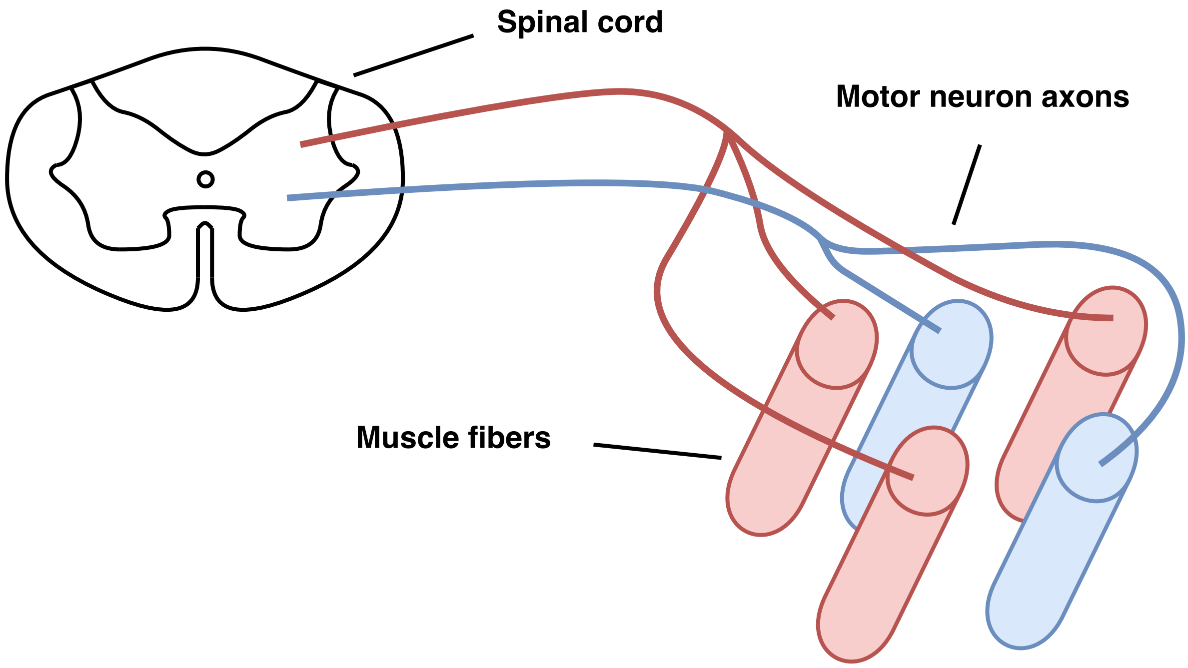


Figure 2.1: The figure shows two motor units (one colored in red and one in blue) based on Figure 12.18, Dee U. Silverthorn, Physiologie, 4., updated version, Pearson 2009 [Sil09]. Usage permission granted by Pearson Deutschland.

Firing Rate and Recruitment

The exerted muscle force depends on two parameters, the *firing rate* wich corresponds to the "firing frequency of the motor units" and the *recruitment* which corresponds to the "number of active motor units" [Hof84]. Hof describes this with the following analogy:

> *This effect is analogous to what happens when one listens to the applause of an audience, at various levels of participation ('recruitment') and enthusiasm ('firing rate').*

- A. L. Hof, *[Hof84]*

As Tanaka notes, this in particular means that "[...] EMG is ultimately not a continuous signal, but the sum of discrete neuron impulses" [Tan]. This should be kept in mind when one processes and analyzes EMG signals.
Furthermore, it is important to note that the single "[motor] units fire asynchronously" [Hof84].

Invasive and Non-Invasive Electrodes

EMG signals can be acquired using two different techniques – invasive and non-invasive electrodes. Both methods measure an electrical potential between one or more electrodes and a reference electrode.

Invasive electrodes are inserted into the muscle using an injection needle [Hof84]. This method is also called *needle EMG* and ensures that one obtains a very selective signal within a range of approximately 0.3 mm [Hof84]. Nevertheless, its main disadvantage is the need for an injection which makes it barely usable for HCI.
Non-invasive electrodes are applied onto the surface of the skin. This is also called *surface EMG (sEMG)*. Due to their placement on the skin surface, the signals are influenced by different tissues between the muscle fibers and the skin. This degrades the signal and makes it more challenging to draw detailed conclusions about the activity of single muscles.

Within this thesis we will only address *sEMG*. If not further specified, EMG refers to sEMG throughout all chapters.

Related Biosignals

There are several other biosignals that are linked to EMG closely. As some of them are mentioned multiple times within this thesis, Table 2.1 gives a brief overview of the most common signals.
Another related biosignal is the *MMG (mechanomyogram)*[1]. In contrast to EMG, it is not an electrical signal, but an acoustical one. It can be measured as the muscle fibers oscillate and thereby produce a signal that can be measured using contact microphones [Ort12]. MMG and EMG have been successfully combined for musical practice, making use of their complementing relation [DCT+13].

[1]Throughout literature MMG is also known as *acoustic myogram (AMG)*, *phonomyogram (PMG)* and *viromyogram (VMG)* [Ort12].

Signal Type	Source	Amplitude	Frequency
EMG (*electromyogram*)	muscles	50 μV - 5 mV	2 Hz - 500 Hz
EEG (*electroencephalogram*)	brain	2 μV - 100 μV	0.5 Hz - 100 Hz
ECG (*electrocardiogram*)	heart	1 mV - 10 mV	0.05 Hz - 100 Hz
EOG (*electrooculogram*)	eye movements	0.5 μV - 5 mV	up to 100 Hz

Table 2.1: Overview of a selection of electrical biosignals based on [Sch15].

2.2 EMG in HCI

EMG has been extensively studied in medical diagnosis. Although EMG is relatively new to the HCI community, it has been used in a variety of applications and scientific work which is presented in this section.

Prostheses

Controlling robotic prostheses, in particular hand-prostheses, is one of the most popular and often repeated applications of EMG ([SG82], [AOY96], [BVDS06], [Ten+07], [CS09]). The most prominent approaches are either to classify movements or gestures based on the EMG signal [SG82], or to continuously map characteristics of the signal to a parameter of the prosthesis (e.g. the exerted force or the angle of a finger) [AOY96]. The same techniques can be applied to related modalities. For example, Moon et al. [Moo+03] use the signals of a muscle at the back of the neck to map the commands "Go", "Stop", "Left" and "Right" of a wheelchair.

Gesture Classification

The problem of predicting the correct gesture given the EMG signal is directly linked to prostheses. Nevertheless, gesture classification comprises a much wider field of applications. These include controlling a portable music player using hand gestures [Sap+09], recognizing the different gestures of "Rock-Scissors-Paper" [Hir+16] and classifying sign language gestures [Abr+16]. Besides hand-gestures, EMG has also been applied to facial expressions. For example, Vrana [Vra93] explored the muscle pattering of the face during different emotional states.

Music and Instrument Control

One field of applications that is particularly interesting within the context of this thesis is the usage of EMG to control music or instruments. Tanaka et al. [TK02] used EMG together with relative postion sensing to control electronic musical devices. Arslan et al. ([Ars+05], [Ars+06]) combined EMG with EEG and created the so called "bio-orchestra". Furthermore, Donnarumma et al. [DCT+13] combined EMG and MMG to modulate two instrument-components using only one complex arm gesture. Saponas et al. [Sap+09] developed a real-time finger-gesture classification that they tested in different scenarios. In their scenario "Air-Guitar-Hero", users used four pinching gestures that were mapped to four buttons in the game "Guitar Hero".[2] An image from the video of the "Air-Guitar-Hero" system can be found in Figure 2.2.

Figure 2.2: Image of the demonstration video to "Air-Guitar-Hero" [Sap+09]

Similar to the work regarding prostheses, the presented EMG-instruments can be divided into instruments that first classify the signal and then trigger an event or switch a state on the one hand and instruments that adjust their parameters continuously, depending on the EMG signal on the other.

Altogether, all previous work known to the author focused on either inventing completely new interaction-patterns or introducing an intermediate gestural representation that is

[2] A demonstration of the system can be found at https://youtu.be/pktVSTwC8qo (followed last on June 26, 2017).

mapped to instrument parameters afterwards. The usage of domain-specific movements such as the hand gesture of playing a guitar chord or the finger positions on a saxophone or a keyboard, has – to the best of our knowledge – not been investigated yet.
Nevertheless, domain-specific gestures offer multiple advantages. If one uses the same movements as they are used to play a real instrument, there is no need to learn the interaction for users that know how to play the instrument. Similarly, an interaction can be practiced using the virtual system and then be used on a real instrument. This interchangeability enables the interaction with the virtual system to be implicitly performed while interacting with a real instrument.
Apart from implicitly interacting with musical instruments, the usage of domain-specific movements can be used in any manual interaction task, such as recognizing handwritings or pointing gestures, or detecting the exerted muscle force in a virtual reality environment.

Consumer Market

In addition to the progress made in research, there is also a noticeable progress in the consumer market regarding EMG. The Myo armband[3] features eight electrodes, a gyroscope and an accelerometer [NHJ15]. At the time of this thesis it is available for 200$. This development suggests that EMG will take a more dominant role in the future of consumer electronics and every-day HCI.

Given all these properties, we think exploring domain-specific movements using EMG is a promising field of research that can open up new possibilities for more natural interactions.

[3]https://www.myo.com (followed last on June 26, 2017)

3 Technical Realization

In this chapter we showcase how an EMG signal can be acquired and processed. Furthermore, it addresses the selection of features and machine learning methods that are commonly used to classify EMG signals. All these steps are consecutive parts of our machine learning pipeline. Figure 3.1 shows the pipeline for our EMG system. The details of the muscle activity have been discussed in section 2.1, the post processing will arise again in Chapter 5. In this chapter, we address the steps from signal acquisition to classification.

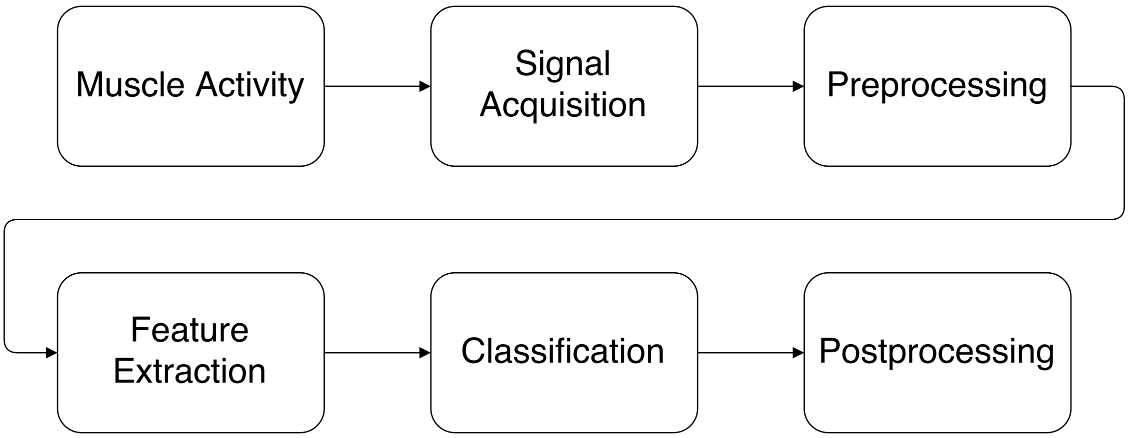


Figure 3.1: Machine learning pipeline for the classification of an EMG signal

3.1 Signal Acquisition

As a first step, the analogue neuron impulses have to be converted into a digital signal. This process involves amplification, sampling and quantization of the EMG signal. We used an actiCHamp recorder[1] to handle these steps. This recorder is designed to be used for EEG recordings, which also makes it suitable for EMG recordings as the EMG amplitude is higher than the EEG amplitude (see Table 2.1). Due to active surface

[1] http://brainvision.com/actichamp.html (followed last on June 26, 2017)

electrodes, the recorder provides a signal with a low noise level and high signal-to-noise quality. The electrodes are attached to the skin using adhesive rings to prevent them from moving on the skin. To improve the conductivity between the skin and the electrodes, a conductive gel is injected between the electrodes and the skin. Similar to Saponas et al. [Sap+08], we attached the electrodes in a narrow band around the left upper forearm (see Figure 3.2). In total we used ten electrodes including one ground- and one reference-electrode. This gave us eight channels.

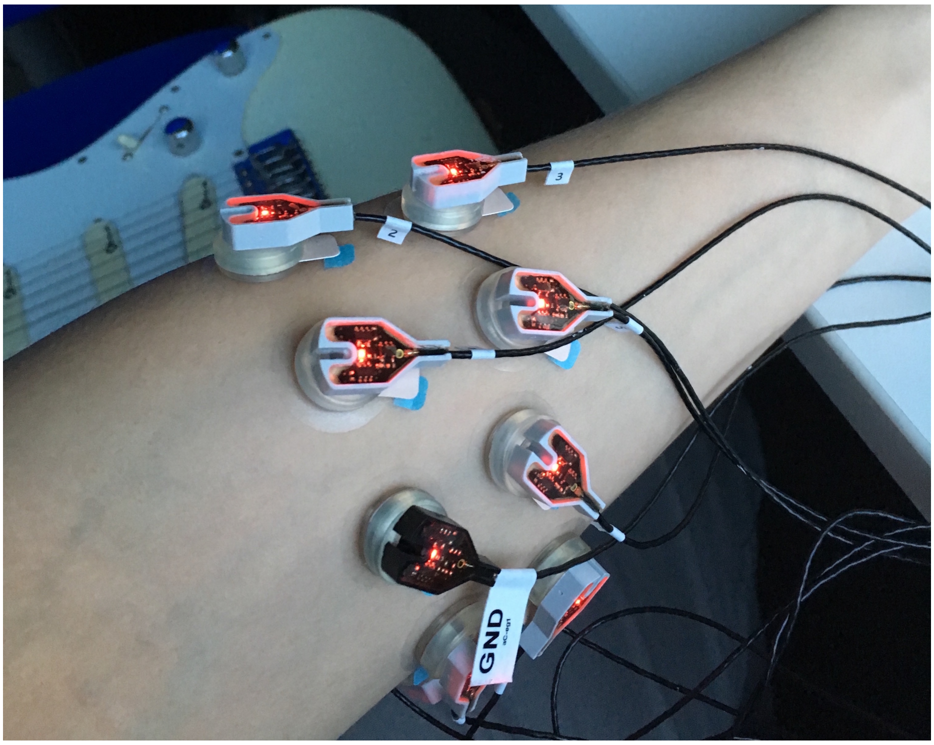


Figure 3.2: Attachment of the electrodes in form of two rings around the upper forearm

Figure 3.3 shows a single-channel EMG recording of the biceps. One can recognize where the muscular activity begins and where the muscle is relaxed by looking at the signal plot. This representation of the signal is not that clear to a computer or more precisely, to a machine learning system. Therefore, one extracts features based on the signal. These features are discussed in Section 3.3. However, before one can extract features from the signal, the signal has to be cleaned. This is done by preprocessing the signal.

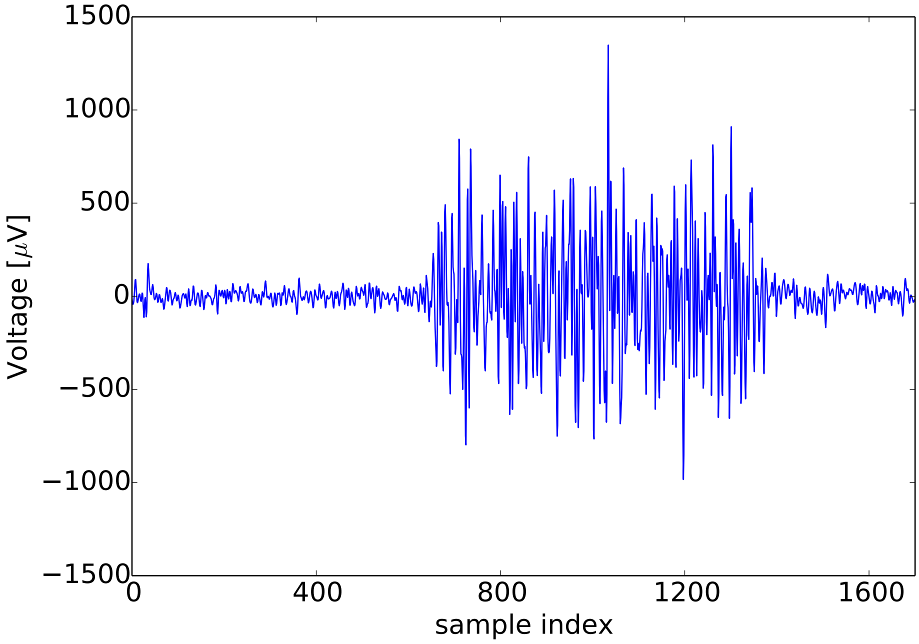


Figure 3.3: EMG signal of a biceps contraction. The signal was recorded using a 500 Hz sampling rate and cleaned using the preprocessing that will be explained in Section 3.2.

3.2 Preprocessing

In order to enhance the signal, two filters are applied in consecutive order:

1. *Bandpass filter*: Similar to Saponas et al. [Sap+08], we apply a bandpass filter between 2 Hz and 100 Hz as their pilot studies and our own tests on the data acquired in Chapter 4 showed that this is the most usable spectrum. A Butterworth filter [But30] of order 5 was used. The removal of low frequencies is important to remove long-term drifts and the DC offset (see Figure 3.5(c)).
2. *Bandstop filter*: Also based on the experiences of Saponas et al. [Sap+08], we filter out the band between 49 Hz and 51 Hz to reduce the typical recording noise (e.g. due to main line frequency). This type of filter is called bandstop. It is also realized using a Butterworth filter of order 5.

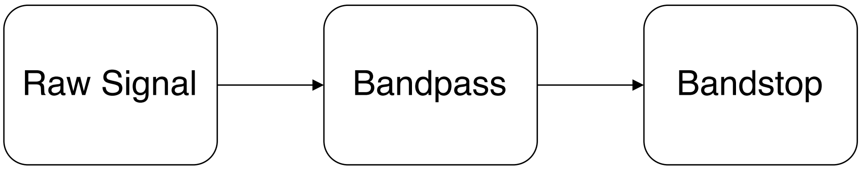
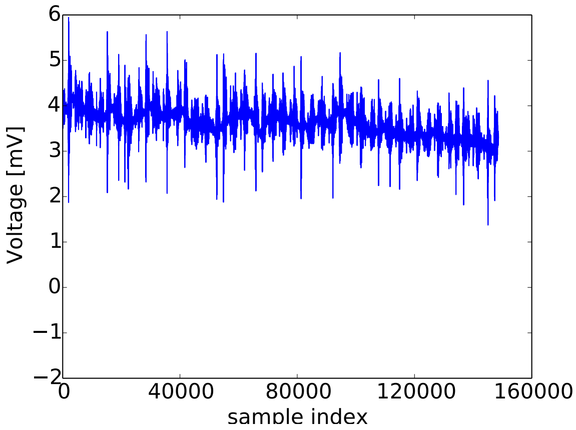
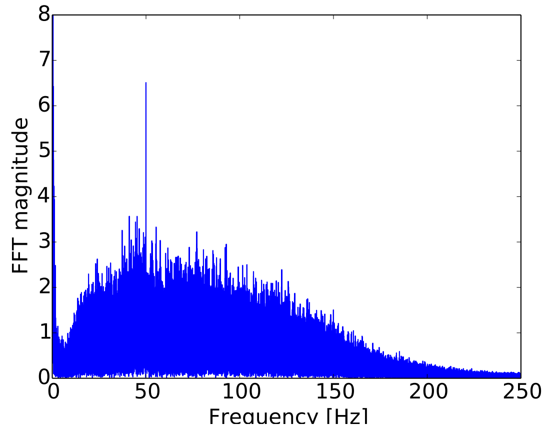


Figure 3.4: Conceptual filter pipeline of the preprocessing step

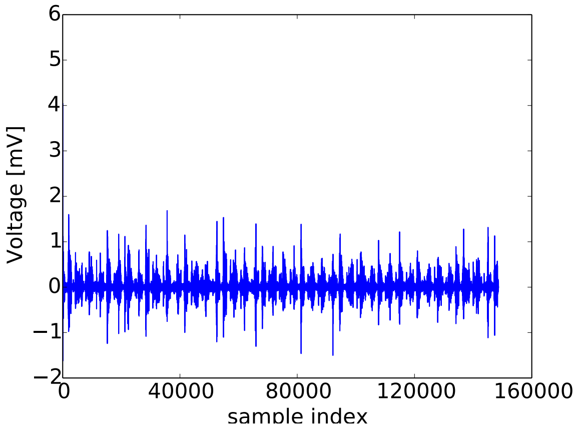
The resulting pipeline is depicted in Figure 3.4. The effects of the single filters can be seen very clearly in the *frequency domain* after applying a *Fourier transform*. Figure 3.5 shows the signal after each step of the pipeline.



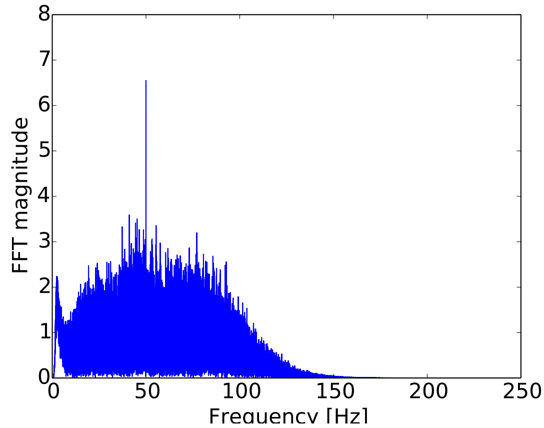
(a) EMG signal without preprocessing



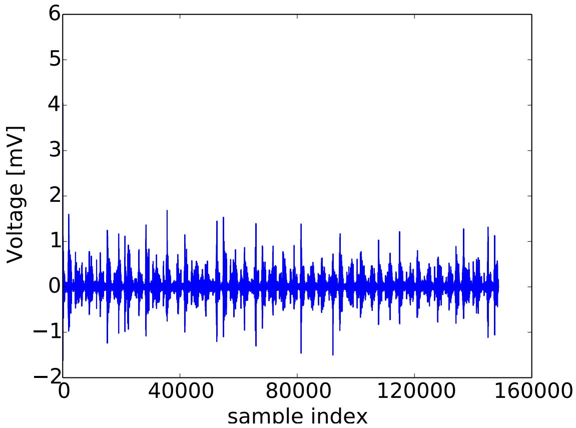
(b) FFT of the unprocessed EMG signal



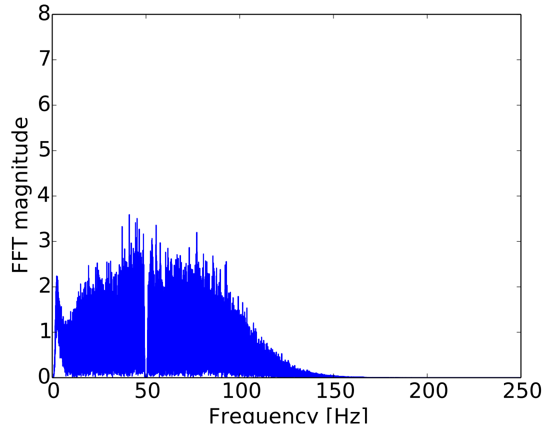
(c) EMG signal after the bandpass filter. One can observe the removal of the DC offset.



(d) FFT after the bandpass filter. Low and high frequencies are attenuated.



(e) EMG signal after the bandpass- and the bandstop-filter. The effects of the bandstop-filter cannot be seen like the effects of the bandpass-filter as the respective frequencies are too high.



(f) FFT after the bandpass and the bandstop filter. The 50 Hz band is cut out.

Figure 3.5: Effects of the preprocessing filters

3.3 Features

Features can be described as properties of an instance. In our case an instance is an EMG signal window. It is a common technique to split the signal into equally sized (e.g. 250 ms) windows to obtain a set of time-independent instances [Sap+08]. For each of these instances a set of of features is evaluated. Figure 3.6 shows a segmented EMG signal on which RMS features have been evaluated. In this case, each of the four windows would be represented by a one dimensional vector (a scalar) with the respective RMS value.

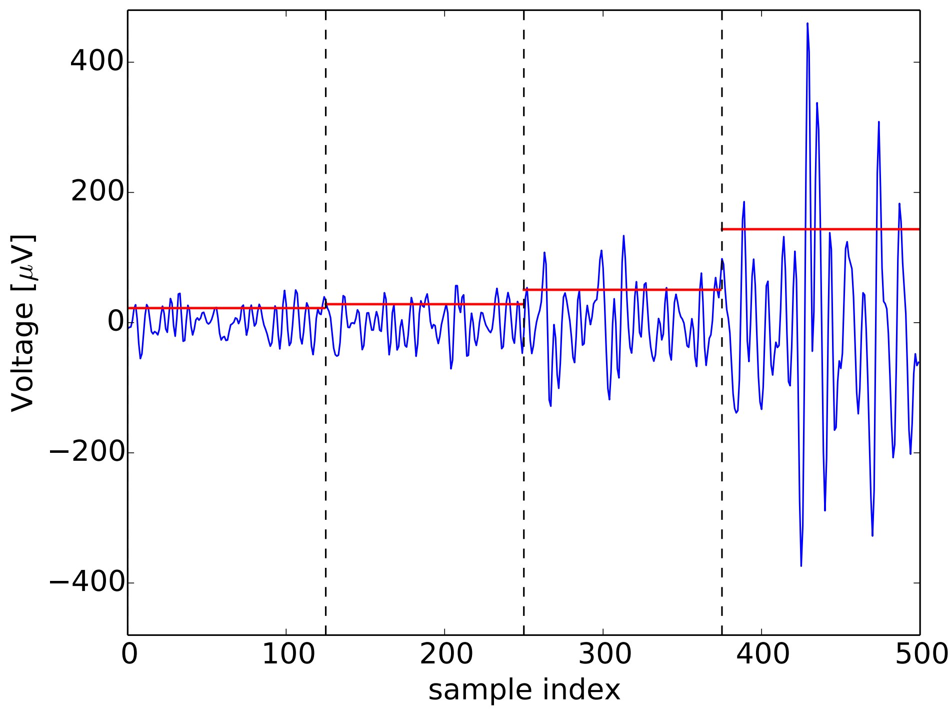


Figure 3.6: Segmented EMG signal using 250 ms windows (125 samples). The red lines show the windows' RMS.

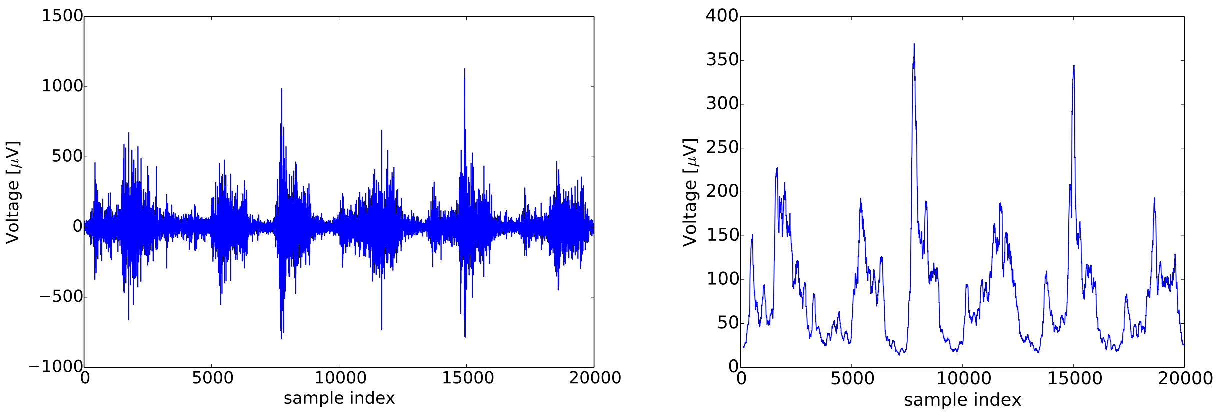
In general, features should express properties of the signal segment by numeric or nominal values. These values are then passed to a machine learning method which we will discuss in Section 3.4.
A broad range of features has been suggested in the context of EMG signals. The following gives an overview over the most commonly used ones.

Root Mean Square (RMS)

Root Mean Square (RMS) is one of the standard features of EMG signals and has been used in a variety of applications ([Sap+08], [KMA08], [Amm+15]). Given the EMG signal as sample points $\{x_1, x_2, ..., x_n\}$, the RMS is defined as

$$x_{RMS} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + ... + x_n^2)}.$$

To visualize the effect of the RMS, Figure 3.7 shows an EMG signal snippet that was preprocessed as described above and the corresponding RMS signal after applying RMS within a rolling window of 125 samples and a hop size of one sample.[2] The motivation behind using RMS is that the EMG amplitude increases with increasing muscular activity as discussed in Section 2.1. Further features can be calculated by taking the relative RMS values between multiple EMG channels [Sap+08]. This will be discussed in Chapter 4.



(a) Preprocessed EMG signal snippet (sampled at 500 Hz)

(b) RMS values of a 250 ms (or 125 samples) rolling window over the signal with hop size 1

Figure 3.7: RMS feature visualization

Frequency Energy

As the motor neurons' firing rate changes with the exerted force (see Section 2.1), the frequency information of the signal encloses valuable information about the muscular

[2] The hop size specifies how far the window is moved (to the right) to generate the next window. We use a hop size of one to show how the RMS evolves over the exemplary signal. In the rest of this thesis we will use distinct windows, corresponding to a hop size equal to the window size.

activity. A fast Fourier transform (FFT) transforms the signal into a representation in the (discrete) Fourier basis. This yields (complex) coefficients corresponding to frequencies up to half the sampling frequency. As we sampled the EMG with 500 Hz, the highest frequency is 250 Hz. Plots of different Fourier spectra are depicted in Figure 3.5(b),(d) and (f). Like Saponas et al. [Sap+08], we used the power of the coefficients and summed them up into 10 bins between 0 Hz and 100 Hz. This yields the (real-valued) energy per frequency bin. One could also consider a more fine- or coarse-grained summation. The advantages and disadvantages of generating many or few features will be addressed in section 4.4.

Phase Coherence

Given multiple EMG channels, another interesting feature is the phase coherence. It measures how synchronous the motor units fire and is a feature that is extensively used in EEG work [Sap+08]. We used the mean coherence as an aggregated feature like Saponas et al. [Sap+08].

ARIMA Models

A typical mean for time series analysis are *ARIMA (auto-regressive integrated moving average)* models. These models are used to predict future values in social sciences or economics (e.g. predicting sales of a certain product). As the name suggests an ARIMA model relies on three basic concepts:

- *Auto-Regressive (AR)*: This part of the model expresses the assumption that one can predict future values as a linear combination of their preceding values. The number of previous values the model considers is often denoted by p.
- *Integrated (I)*: ARIMA models require the time series to be stationary. To fulfill this assumption it is often necessary to difference the time series – that is subtracting each value from its subsequent value. The order of differencing is often denoted by d.
- *Moving Average (MA)*: In contrast to the *AR* component, the *MA* component does not model future values depending on preceding values but only as a linear combination of preceding error terms. The number of the error terms taken into account is often denoted by q.

To extract features from a given EMG signal segment, one fits an *ARIMA* model to the segment and uses the respective *AR* and *MA* coefficients as features. This is conceptually similar to fitting a linear regression and using the slope and the intercept as features. *ARIMA* and the related *AR*, *MA* and *ARMA* have been applied to EMG signals multiple times ([PK93], [RHMY06], [Kar14]).

Wavelet Transform

Like the Fourier transform, the *wavelet transform (WT)* transforms a time series into another basis that yields frequency information. For this transform it is the wavelet basis. In contrast to the Fourier basis, a signal in the wavelet basis is not only localized in frequency but additionally in time. WT coefficients are a common feature in EMG signal classification ([RHMY06], [Kar14]).

3.4 Machine Learning Methods

In machine learning *classification* is the problem of finding the correct class (or label) to a given instance. In the context of EMG signals, an instance corresponds to a window's feature vector. Possible sets of classes in a gesture recognition system could be $\{active, inactive\}$ or $\{pointing, waving, clapping\}$.
Within supervised machine learning, one trains a model using training data that contains a set of instances and their respective labels. This training comes down to minimizing a loss function with respect to the model's parameters. Discussing the different models and methods exceeds the scope of this thesis.
When classifying EMG signals, the most common models are *Support Vector Machines (SVMs)* ([Sap+08], [Abr+16], [PJT16]) and *Artificial Neural Networks (ANNs)* ([AIK11], [Guo+15]). The performance of these two models depends on the choice of features, but overall SVMs and ANNs showed similar results throughout literature ([Kar14], [Guo+15]). Therefore, this thesis will use SVMs as classification model.
In the context of EMG signals, it is especially important to discuss the impact of high dimensional feature vectors and common pitfalls when evaluating the classification performance, which we will address in Chapter 4 and Chapter 5.

4 Evaluation of Features and Parameter Choices

We introduced features that are commonly used in EMG classification in Chapter 3. This chapter evaluates these features and further methods for a special set of gestures – that is guitar chords. Guitar chords differ from usually used gestures in the exerted force and the level of granularity. This imposes challenges on the classification of these gestures. To quantify the suitability of the presented features for this task, we evaluate their discriminative power leveraging SVMs. Apart from evaluating the features, we investigate the effects of different parameter choices on the classifier's performance. These include parameters such as the size of a signal-segment and hyperparameters of the classifier.

4.1 Method

Design

In order to acquire labeled data, we implemented a recording system that gathers the EMG signals of eight electrodes together with the chords that the left hand[1] plays.
We used two laptops to handle the recording and the experimental setup, handling the synchronization of the EMG and marker signals via *Lab Streaming Layer (LSL)*[2]. LSL was designed to handle near real-time data exchange within research experiments. A validation[3] showed that LSL introduces timing errors under 1ms which is sufficiently low for our use case. The detailed marking procedure will be explained in the following. A schematic depiction of the whole setup is shown in Figure 4.1.

[1]We used an electric guitar where the chords are played with the left hand and the strings are strummed with the right hand.
[2]https://github.com/sccn/labstreaminglayer (followed last on June 26, 2017)
[3]https://sccn.ucsd.edu/~mgrivich/LSL_Validation.html (followed last on June 17, 2017)

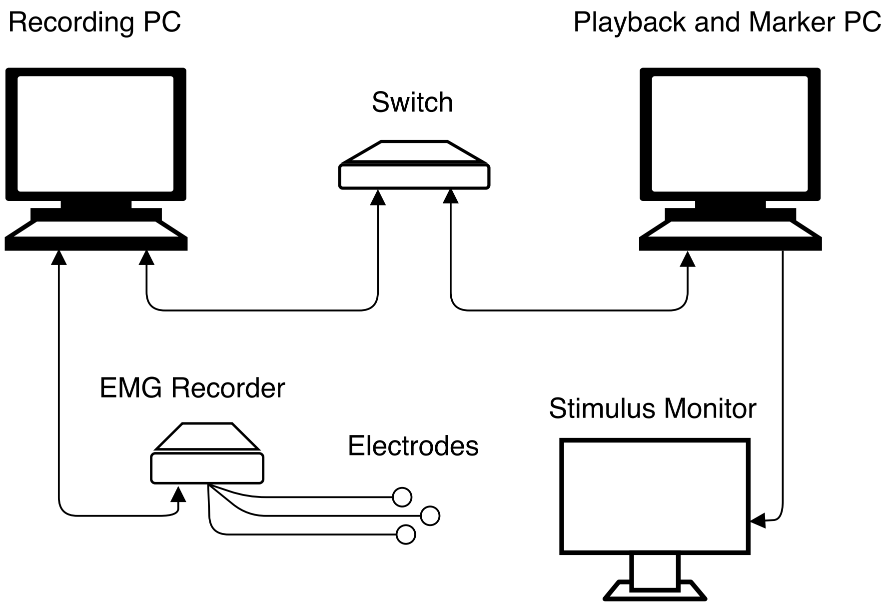


Figure 4.1: Schematic depiction of the recording setup

We decided to task the participants to play the chord sequence:

-, Am, C, -, G, D , -, Em, F, -, Fm.

"-" refers to no chord being played by the left hand. This yields eight classes of chords. From a music theory perspective, in the context of the C major scale, these chords include I (C), III (Em), IV (F), V (G), VI (Am). This choice of chords is realistic, as a vast amount of songs can be played using these chords or transposed versions of them. The five mentioned scale degrees cover a large majority of the distribution found within the RS 5 x 20 corpus. This corpus contains annotated scale degrees of popular rock music [DCT11].

A musical piece also contains rests. During this time, no chord is played and thus, we include the "-" state. As we cannot assume a distribution of chords and pauses, we chose to introduce a pause at the beginning and after two chords each. Based on the selected chords we chose the pattern shown in Figure 4.2. The chords were presented to the subjects using the simple UI depicted in Figure 4.3. The current chord is colored red, whereas the next chord is shown in black. In addition, a bongo sound was triggered each beat to help the subjects to stay in the right rhythm. The first beat was highlighted using a different bongo sound. To realize the playback, we sent MIDI notes to the digital audio workstation Ableton Live. A LSL marker containing a chord identifier was sent each time a new chord was displayed.

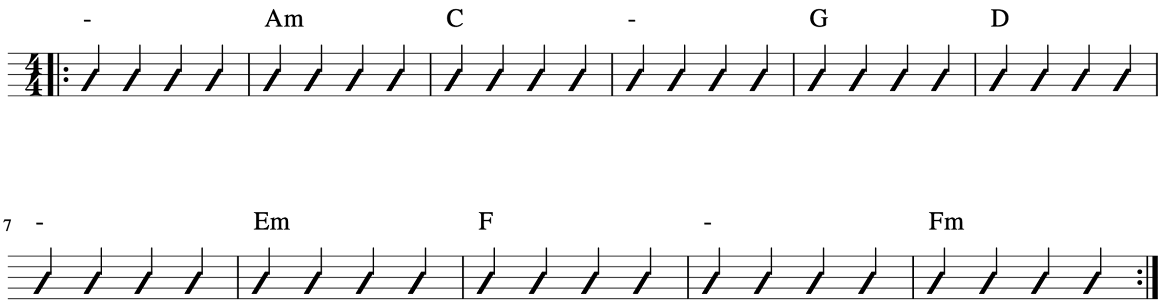


Figure 4.2: The chord pattern the participants were tasked to play. "-" indicates that no chord is to be played (all strings are open).

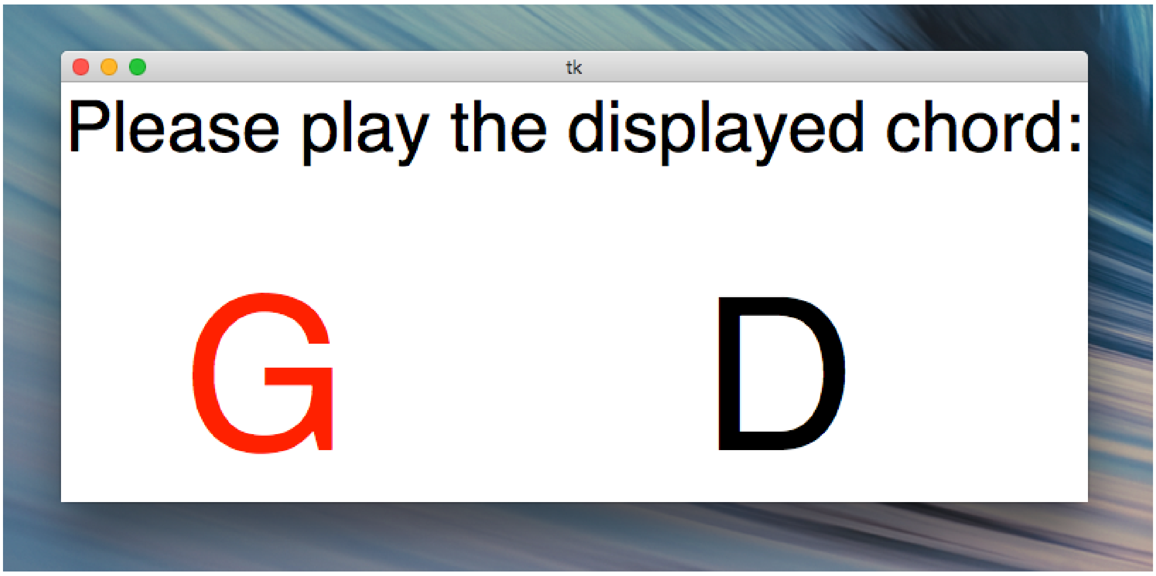


Figure 4.3: Recording UI that presented the chords the participants played.

To keep the task as easy as possible, the subjects were asked to strum each chord of the sequence four times with a tempo of 100 beats per minute, thereby minimizing the amount of false annotations due to erroneous play. Before starting the experiment, the subjects practiced the chord sequence until they felt confident playing it. Each subject repeated the pattern for five minutes. After a subject had finished, the EMG signal and the marker data were stored.

4.2 Evaluation Procedure

We applied the preprocessing steps explained in Section 3.2 to the recorded EMG signals. These included a bandpass- and a bandstop filter. Given the cleaned signals and the chord markers, we divide each EMG recording into segments corresponding to one bar of the same chord. Subsequently, we split each segment into fixed-size, non-overlapping windows. We used 250 ms windows for the standard window size as suggested by Saponas et al. [Sap+08][4].

Since the window size is not an integer divisor of the length of one bar, we dismiss the partial window at the end of each bar. Similar to Saponas et al. [Sap+08], we additionally neglect a fixed number of full windows at the end of the remaining windows (see Figure 4.4), as they may be distorted by the change of chords. At a tempo of 100 beats per minute (BPM), a quarter note has the length of 600 ms. Thus, a chord change has to take less than 600 ms. Therefore, we chose to neglect two full windows (500 ms) and a partial window (100 ms) at the end of each segment throughout this chapter.

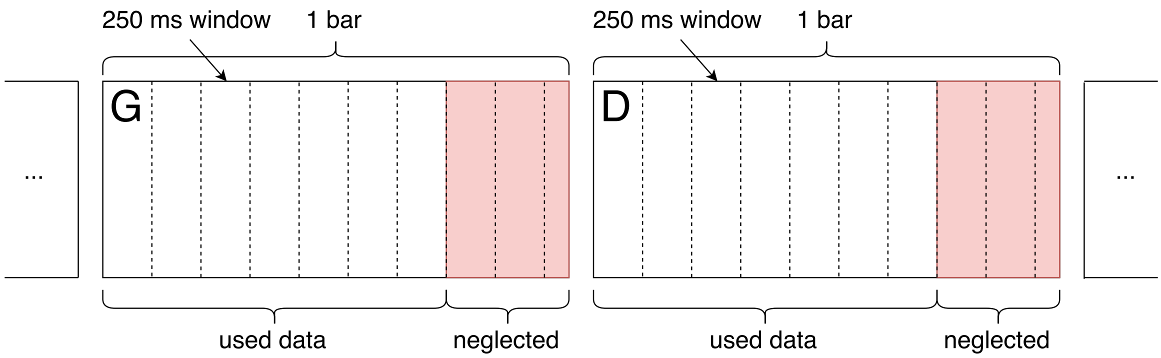


Figure 4.4: Depiction of the neglected and used signal parts, each box represents one bar. The dotted lines mark 250 ms windows. We neglect two full windows and a partial one at the end.

The extracted windows are transformed into instances using different subsets of the feature functions presented in Section 3.3. The resulting data set is split into training data (the first 80%) and test data (the remaining 20%). Whereas one often shuffles the data set before dividing it into train- and test-sets, it is important to keep the temporal order of instances of EMG signals as shuffling could artificially boost the classifier's performance [Sap+08]. This also reflects the real-world procedure of calibrating the system before using it.

[4]If not stated differently, we used 250 ms windows throughout all results in this chapter.

Before fitting a SVM with the training data, we first normalize the whole data set. Therefore, we estimate the mean and the standard deviation on the training data[5]. We center the feature values around zero by subtracting the estimated mean. Additionally, we scale the feature values to unit variance by dividing each component with its estimated standard deviation. A zero mean and unit variance are often assumed, for example when using the RBF kernel[6].

We use the training data to fit a SVM. In essence, this is an optimization problem where we maximize the margin between two classes in the feature space. This adjustment happens within a fixed setting of specific parameters such as a regularization coefficient, a specific kernel and more. These parameters are called *hyperparameters* and influence the performance of a classifier. For best performance, optimization of hyperparameters can be approached in multiple ways. We are using a *grid-search cross-validation*. This method performs an exhaustive search over all combinations of hyperparameters[7] and chooses the best combination with respect to some metric (such as accuracy or F1-measure) using a *cross-validation (CV)*.

A cross validation estimates the performance on unseen data, that is data that was not used for training. The *k-fold cross validation* splits the training data into k sets and uses one set for validation and only the remaining $k-1$ sets for training[8]. This is repeated for all k combinations, resulting in k iterations. An example for $k = 3$ is shown in Figure 4.5.

As the EMG signal of one bar (and hence one chord) is transformed into multiple instances, the instances are not distributed independently. Therefore, we want to avoid that the instances of one bar are used for training *and* validation during one iteration as this could artificially improve the classifier's performance. To this end, we apply a *group k-fold* that ensures that instances of one bar stay together during the k-fold. Throughout our experiments we chose $k = 3$ as higher k values induced an unacceptably high computation time.

[5] We do not include the test data for the estimation, as this would corrupt the separation of train and test data.

[6] The radial basis function (RBF) kernel is commonly used in SVMs.

[7] We tested, linear, polynomial and RBF kernels, the gamma values (kernel coefficients for RBF and polynomial kernels) 0.0000001, 0.001, 0.01, 1.0, 10.0 and 100.0 and the penalty parameters for the error term 0.0000001, 0.001, 0.01, 0.1, 1.0, 10.0 and 100.0.

[8] This means that all k splits are part of the 80% training data, the data of the 20% test set is not used in the CV.

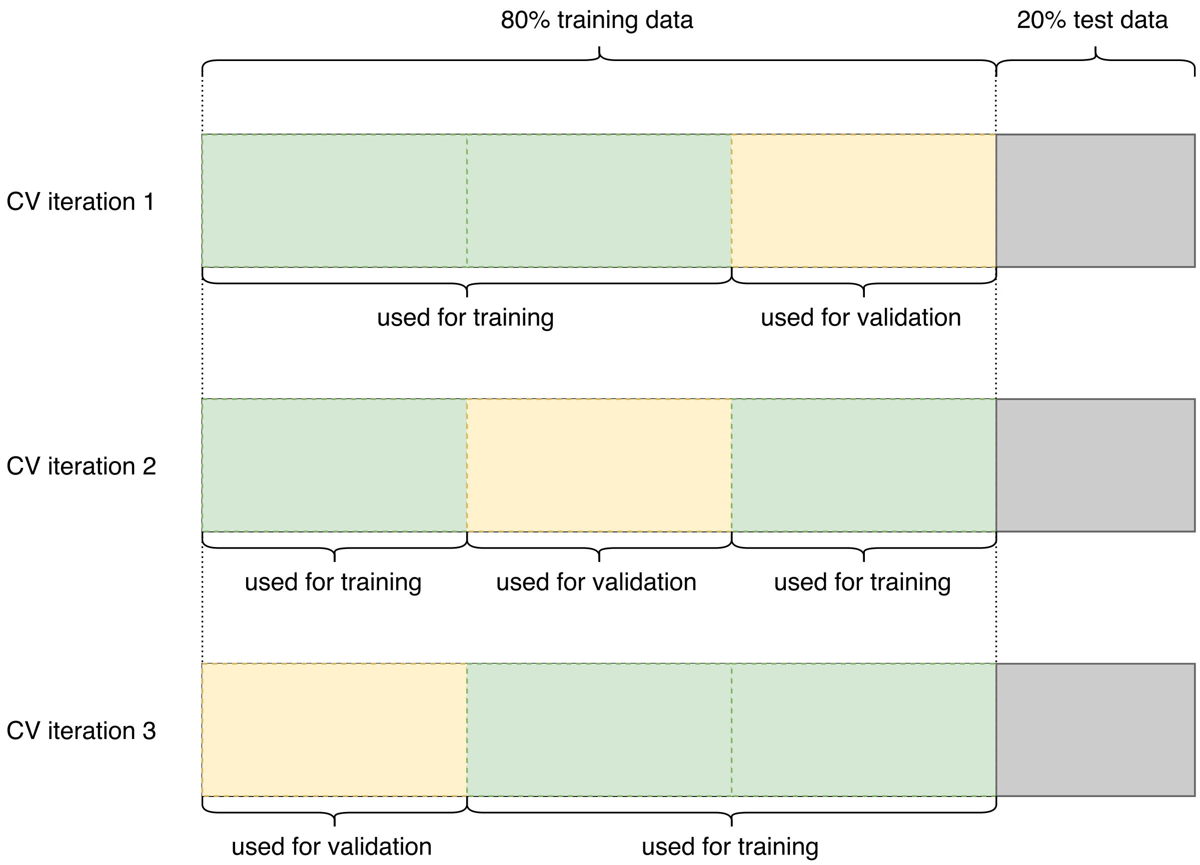


Figure 4.5: Depiction of a three-fold cross validation. The test data is never used throughout all iterations.

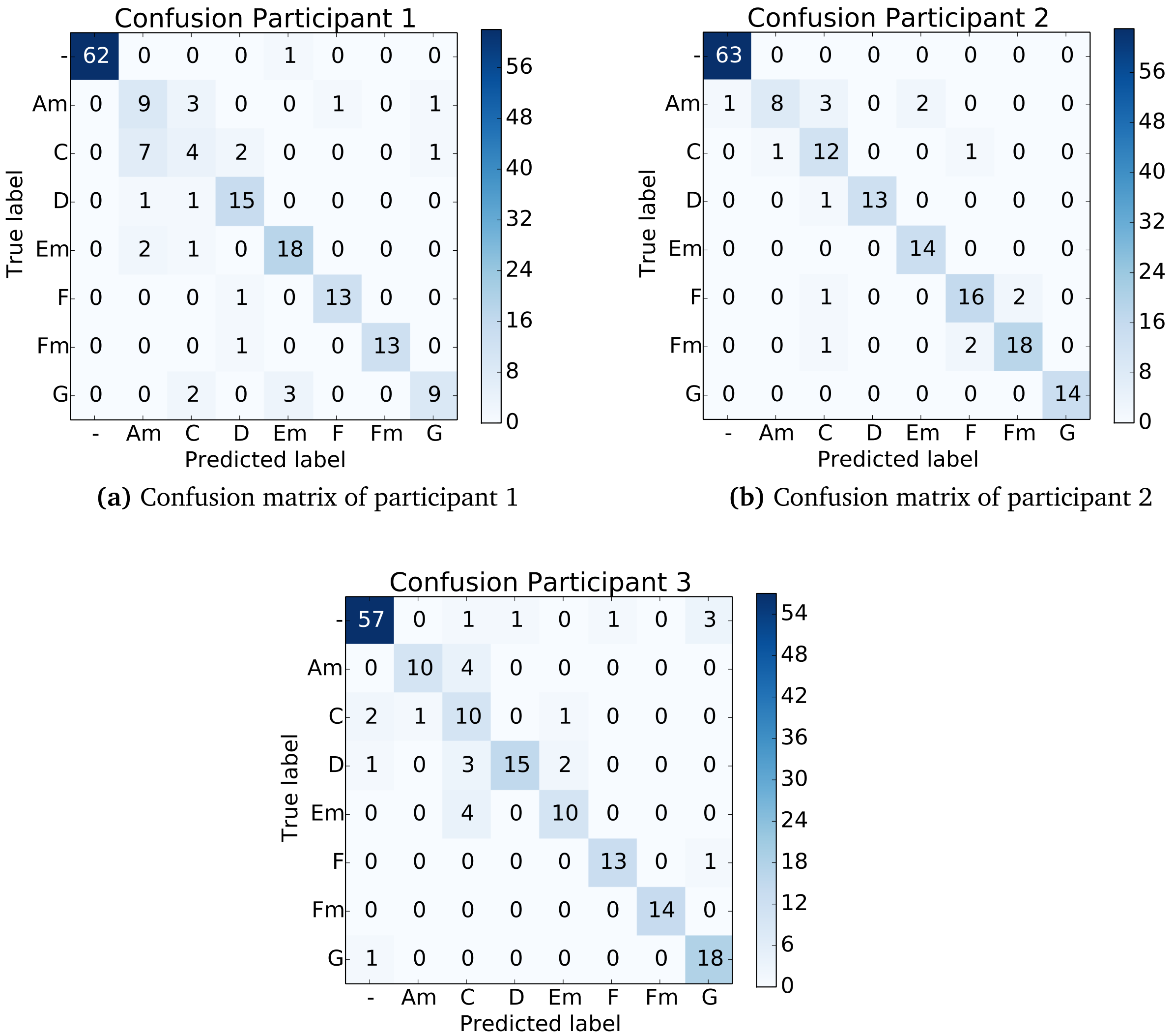
4.3 Results

We collected data of three male participants to gather training data of 15 minutes length in total. We analyzed the classifier's performance with respect to different feature sets. The results are shown in Table 4.1.[9] We tested further combinations of features, all tested combinations could not outperform the combination of RMS and RMS ratios. The respective confusion matrices for the individual subjects using RMS and RMS ratio features are depicted in Figure 4.6.

Feature Group	Accuracy	Precision	Recall	F1
RMS	83 *SD=5*	85 *SD=4*	83 *SD=5*	83 *SD=4*
RMS Ratios	83 *SD=1*	84 *SD=1*	83 *SD=1*	83 *SD=1*
Frequency Bins	47 *SD=6*	41 *SD=13*	47 *SD=6*	41 *SD=11*
Phase Coherences	62 *SD=3*	62 *SD=5*	62 *SD=3*	59 *SD=3*
ARIMA	52 *SD=4*	50 *SD=4*	52 *SD=4*	49 *SD=4*
Wavelet Transform	43 *SD=11*	36 *SD=13*	43 *SD=11*	37 *SD=12*
RMS and RMS Ratios	87 *SD=3*	88 *SD=3*	87 *SD=3*	87 *SD=3*
All Features	73 *SD=5*	75 *SD=5*	73 *SD=5*	73 *SD=5*

Table 4.1: Weighted average accuracy, precision, recall and F1-measure for different feature groups averaged over all participants. All metrics are measured in percent.

[9] We report weighted averages throughout this thesis. The single classes' measures are weighted with their respective instance count in the test data.



(a) Confusion matrix of participant 1

(b) Confusion matrix of participant 2

(c) Confusion matrix of participant 3

Figure 4.6: Confusion matrices with respect to the classifier's performance on the test-data. The total number of instances per class varies over the three participants because of the random starting position in the pattern and the train-test split.

To gain an overview of how well the different features can be used in a real-time system, we measured the processor time that is required to calculate the features within our setup. The results are shown in Table 4.2 together with the feature groups' dimensions.

Feature Group	Dimensionality	Averaged Processor Time
RMS	8	1.32 s
RMS Ratios	28	1.38 s
Frequency Bins	10	1.73 s
Phase Coherences	36	7.20 s
ARIMA	16	582.34 s
Wavelet Transform	1008	0.90 s
RMS and RMS Ratios	36	1.40 s
All Features	1106	597.59 s

Table 4.2: Dimensions and averaged processor times required to compute the respective features for a five minute recording.

Regarding real-time usage, we evaluated the classifier's performance depending on the window size. To keep the results comparable, we discard more windows if the windows get smaller. This ensures that the neglected signal parts (see Figure 4.4) stay constant when the window size is varied. As the combination of RMS and RMS ratios yields the best performance, we used these two sets of features. The detailed results are shown in Table 4.3. To get a more intuitive understanding of those values, we additionally plotted the averaged F1-measure in dependence of the window size in Figure 4.7.

Window Size	Discarded Segments	Accuracy	Precision	Recall	F1
500 ms	1	88 *SD=3*	90 *SD=3*	88 *SD=3*	89 *SD=3*
250 ms	2	87 *SD=3*	88 *SD=3*	87 *SD=3*	87 *SD=3*
166 ms	3	84 *SD=4*	84 *SD=3*	84 *SD=4*	84 *SD=4*
126 ms	4	82 *SD=4*	83 *SD=3*	82 *SD=4*	82 *SD=3*
84 ms	6	79 *SD=4*	79 *SD=4*	79 *SD=4*	79 *SD=4*
64 ms	8	75 *SD=5*	75 *SD=4*	75 *SD=5*	75 *SD=4*
50 ms	10	72 *SD=5*	72 *SD=5*	72 *SD=5*	72 *SD=5*
32 ms	16	68 *SD=4*	67 *SD=5*	68 *SD=4*	67 *SD=5*
20 ms	24	63 *SD=3*	61 *SD=4*	63 *SD=3*	61 *SD=4*

Table 4.3: The performance of the classifier given the window size of the segments used for classification. All values are calculated using RMS and RMS ratio features. Window sizes like 126 ms are used instead of 125 ms in order to keep the neglected signal size constant. All metrics are measured in percent.

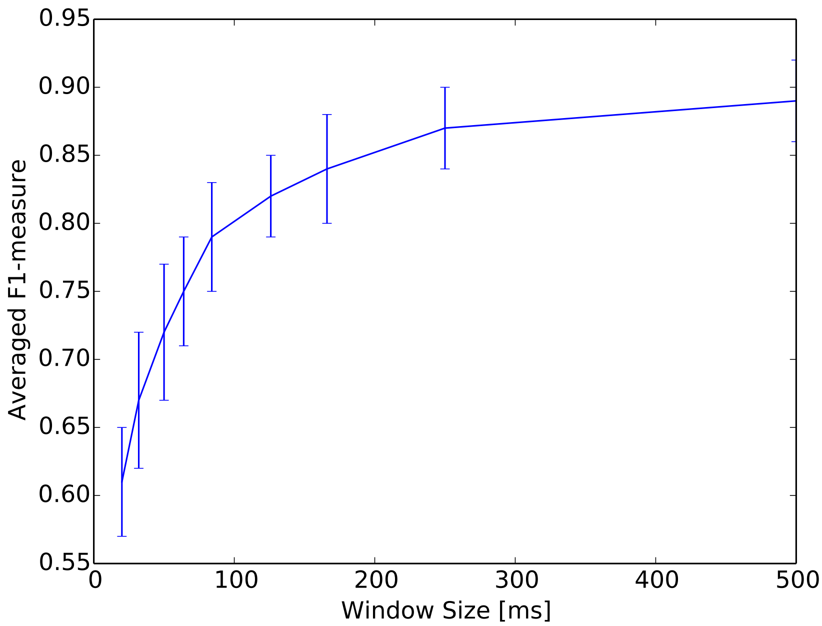


Figure 4.7: Averaged F1-measures of the classifier in dependence of the used window size. The error bars indicate one standard deviation. All values are calculated using RMS and RMS ratio features.

In addition to the separate analysis of the subjects' EMG recordings, we also evaluated the inter-person generalizability. Therefore, we trained a SVM using a leave-one-subject-out approach. This means that the SVM was trained on the joint instances of two subjects and tested on the remaining third subject's data. The results are shown in Table 4.4.

Test Set	Training Set	Accuracy	Precision	Recall	F1
Subject 1	Subject 2 ∪ Subject 3	23	14	23	17
Subject 2	Subject 1 ∪ Subject 3	18	39	18	12
Subject 3	Subject 1 ∪ Subject 2	39	40	39	36

Table 4.4: Inter-person performances using RMS and RMS ratio features. All metrics are measured in percent.

4.4 Discussion

The results show that the combination of RMS and RMS ratios yields the best performance whereas using RMS and RMS ratios alone results in slightly lower values. The measures for frequency bins, phase coherences, ARIMA coefficients and the wavelet transform show an evidently lower performance. The wavelet transform gives the lowest performance, however, we believe this can be attributed to the high number of dimensions of this feature group (1008 dimensions) and the resulting sparsity problem[10]. We did not evaluate this in more detail, but suggest to reconsider less detailed coefficients for our application in future work. We believe that the mentioned sparsity problem also causes the decrease in performance from RMS and RMS ratios to using all features. At first sight, it is counter-intuitive to gain a lower performance even though the classifier can use the same and more information. However, it's more difficult for the classifier to distinguish between the classes if the number of dimensions increases but the number of instances stays the same as this results in an undersampling and the so called *curse of dimensionality* [Bis06]. This problem can be addressed by dimensionality reduction methods. Overall, our results show that different features should be tested prior to adapting existing work in the field of EMG to new gestures.
Considering the computation times, we want to highlight two values. Firstly, the ARIMA coefficients time exceeds all other features calculation time by more than factor 80. This very high time consumption has to be attributed to the implementation we used. We do not investigate this in detail, as the coefficients do not yield a beneficial performance

[10]For a high number of dimensions, one needs exponentially many training instances.

regarding our application. Secondly, the computation time of the joint RMS and RMS ratio is faster than all other feature groups except the RMS and RMS ratios alone and the wavelet transform. This suggests that RMS and RMS ratio are suitable features for a real-time setting.
Altogether, the combination of RMS and RMS ratios is a promising set of features regarding both, classification performance and computation time. A possible trade-off between discriminative power and computation-time could not be observed in our results.
Besides performance and computation time, the results show an interesting phenomenon regarding the used window sizes. Starting with an F1-measure of 89% for a window size of 500 ms, the performance drops to 87% for a 250 ms window. Further shrinking the window size monotonously decreases all reported measures and finally drops to a F1-measure of 61% for a 20 ms window. Smaller windows intuitively carry less information. However, it is important to note that with a fixed calibration time, the number of training instances increases with a decreasing window size. In theory, this effect might counteract the information loss. Yet the reported performance of small window sizes suggests that this advantage can not compensate the information loss.
Overall, our results show that there is a trade-off between real-time feasibility and classification performance. On one hand, a short window size is desirable as the system's latency depends on it. On the other, large window sizes decrease the number of misclassifications. This should be considered when designing new interactive system that rely on EMG to detect hand- or finger-gestures.
The three reported confusion matrices indicate that classification on a binary level (a chord is played versus no chord is played) seems to be very accurate throughout all participants. However, the confusion matrices show that the distinction between Am and C seems to be more difficult than classifying the other labels. Especially in the confusion matrix of participant 1, there seems to be a misconception of the C-chord. However, this phenomenon cannot be observed clearly in the other two confusion matrices. This leaves the question whether the distinction between Am and C is a general challenge or just a result of the participant's playstyle or physiology. We will further investigate this in Chapter 5.
In addition to the training and testing on each subject's data separately (intra-person performances), we also reported the inter-person performances. With a highest F1-measure of 36% and a lowest F1-measure of 17%, these results show that the models aren't able to generalize to data of a subject that was not observed yet. This is due to two reasons. Firstly, the slight changes in the electrodes position and the different physiology of the participants' arms result in a change of intensity the electrodes capture from each muscle. PCA and other dimensionality reductions are a promising approach to tackle this problem. We will address this within Chapter 5. Secondly, our training database of only two participants does not contain enough variance to enable the classifier to learn robust concepts of chords. This problem can be addressed by recording data from more subjects. We will reconsider this question in Chapter 5 using a larger data set.

5 Guitar Tutor System

In Chapter 4 we showed that we can classify guitar chords using the EMG signal of the left forearm with an average F1-measure of 87% on the data of our three subjects. Nevertheless, these results raised new questions:

- Can we reach a higher inter-person[1] classification performance using training data of more participants?
- Is the distinction between specific chords (such as Am and C) a general problem?

We decided to conduct another study to address these questions. Furthermore, we investigate a real-world scenario. We implemented an EMG guitar tutor system to achieve this.

5.1 Scenario

Systems, which support an aspiring musician in learning to play an instrument, already exist in the consumer market. However, the majority of these systems use a microphone to evaluate the musicians' play by the sound they produce. This is a reasonable way to evaluate playing skills, as the produced sound is, more or less, what matters most to the musician. Nevertheless, the audio-based tutor systems fail to detect *how* the musician plays. Interfaces with EMG capabilities are able to close this gap. We can measure at which chord changes the musician is slow. The system then provides advise which chords and chord changes should be practiced. Additionally, the EMG signal provides a notion of the applied pressure exerted by the player, which can be used to detect bad hand postures. These advices cannot be given using the audio signal.
In the first version of our system we focus on the adjustment of the tempo depending on the playing accuracy. To make lasting progress in increasing one's playing proficiency, it is vital to start off slow. If the player is comfortable with the current tempo, one can increase the speed gradually until the desired speed is reached. Hence, we want to

[1]This means that we train on the other subject's data and predict labels of a subject that classifier has not seen.

determine whether the user is comfortable at a given speed by estimating the playing accuracy. If the accuracy is low, we slow down the tempo for the next iteration. If the accuracy is high, we speed it up.
This system is complex enough to evaluate how the musicians interact with an EMG tutor system and which requirements and problems arise when they use the system.

5.2 Method

In Chapter 4 we evaluated seven chords and the rest state. In this study, we want to focus on a specific set of chords: *C, G, Am, F*. We do not include the rest-state, as our evaluation already showed good classifier performances regarding it. The chosen chords form the *I-V-vi-IV-progression*. This progression is one of the most used in pop music[2]. We do not want to limit the chord changes to this progression. Therefore, we shuffled all possible permutations of these four chords and divided the resulting set into two patterns. Inside each pattern, we repeated each chord as often as the others. We distributed the number of successive repetitions of one chord arbitrarily. The resulting patterns can be seen in Figure 5.1 and Figure 5.2.

As we changed the set of chords, we also changed the calibration pattern. In order to record a sufficient amount of samples, we tasked the participants to play two bars per chord. The resulting calibration pattern is depicted in Figure 5.3. The participants were asked to repeat the pattern two times during each calibration.

[2]An incomplete list of songs using this progression can be found at https://en.wikipedia.org/w/index.php?title=List_of_songs_containing_the_I-V-vi-IV_progression&oldid=784377090 (followed last on June 26, 2017).
A video that shows a compilation of popular songs using this progression can be found at https://youtu.be/oOlDewpCfZQ (followed last on June 26, 2017).

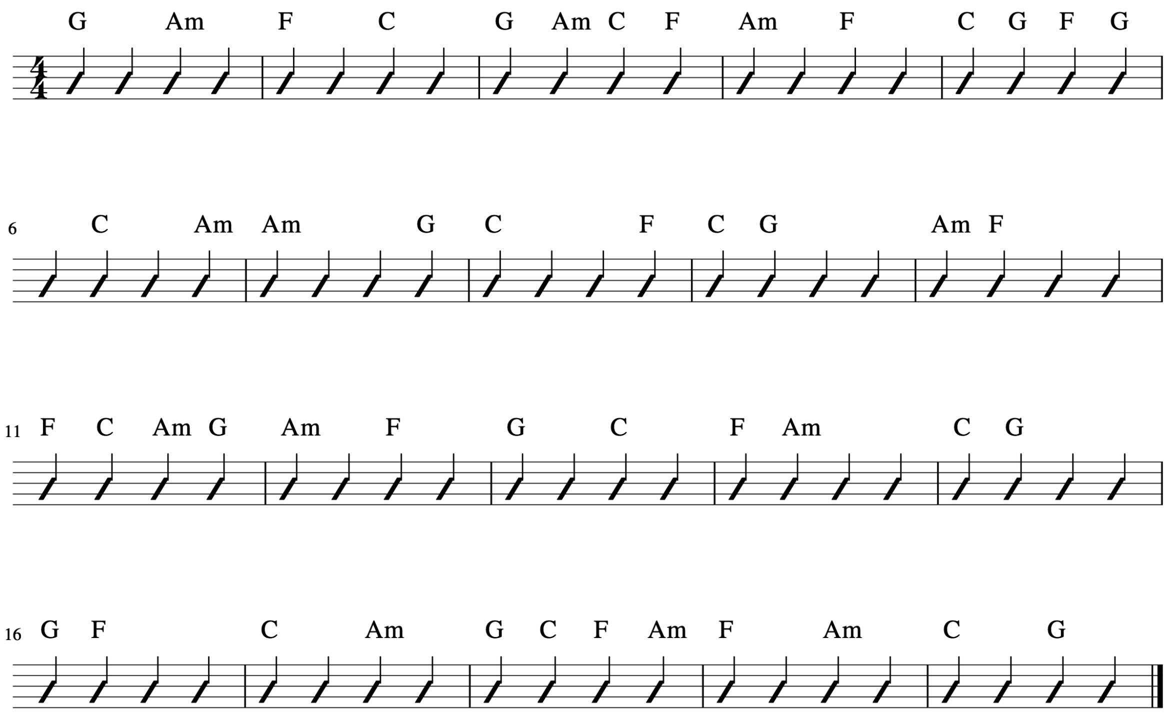


Figure 5.1: Pattern A

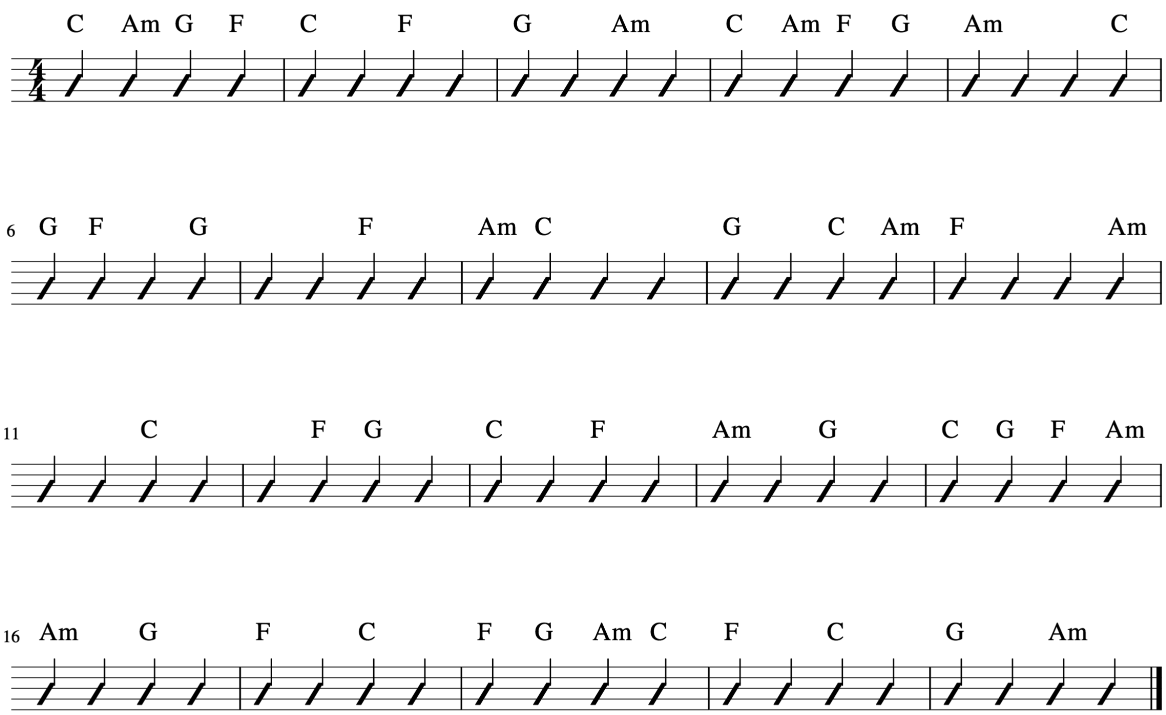


Figure 5.2: Pattern B

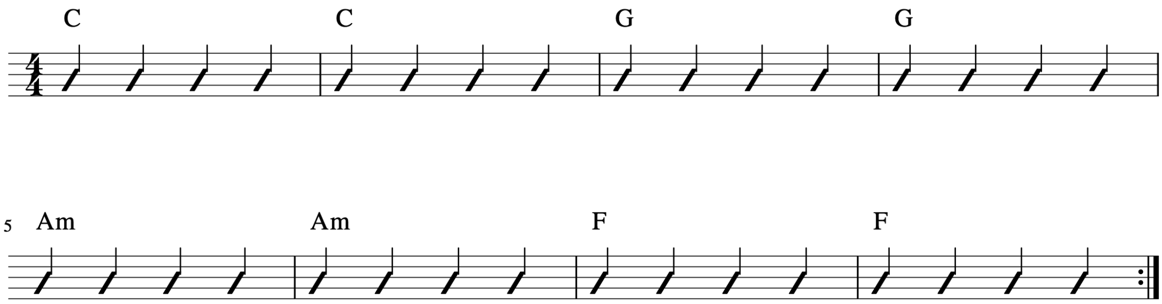


Figure 5.3: Calibration pattern

The procedure of our study is depicted in Figure 5.4, a picture of the setup is shown in Figure 5.5. After we welcomed the participants, they were given some minutes to familiarize themselves with the guitar. Subsequently, we conducted the first calibration with 40 BPM. After this, the subjects were asked to play one of the patterns A or B for five times. After each repetition, the playback stopped and the tempo was adjusted. This adjustment was either made *manually* (the participants chose the next tempo freely) or *automatically* (the system showed the estimated accuracy to the subject and adjusted the tempo respectively). In both cases the tempo had to be between 40 BPM and 90 BPM as a tempo below 40 BPM is hardly playable and our choices of window size and the number of neglected windows do not allow classification of chords played faster than 90 BPM. Despite this range seems to be quite slow, quarter note chord changes at a speed of 90 BPM are thoroughly challenging. We chose the new tempo as

$$\min(90, \max(40, p + (a - 0.5) \cdot 100))$$

with the previous repetition's tempo p and the estimated accuracy a. We empirically evaluated this mapping during test runs. Throughout all five repetitions in each run, the system type (automatic or manual) was kept consistent. In either case, the tempo started with 40 BPM. After the fourth round, the tempo was set to 90 BPM. This ensures the same starting and end tempo for each subject. Following the five rounds, another calibration was conducted. Then, the subject was asked to play another run of five repetitions, but this time using the other pattern and the other system (manual or automatic). The combinations of patterns and systems were assigned to the participants using a Latin Square design.

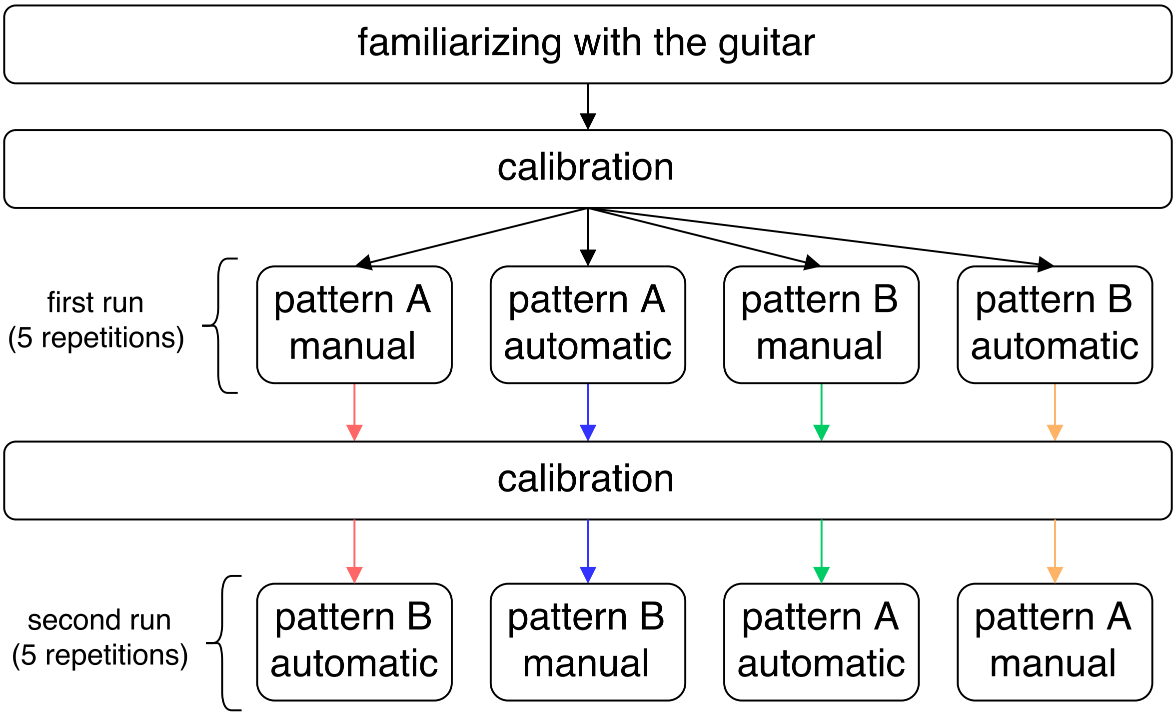


Figure 5.4: Study procedure. The colored arrows indicate a participants path through the study. If, for example, pattern A is played manually in the first run, the subject is assigned pattern B with the automatic system in the second run.

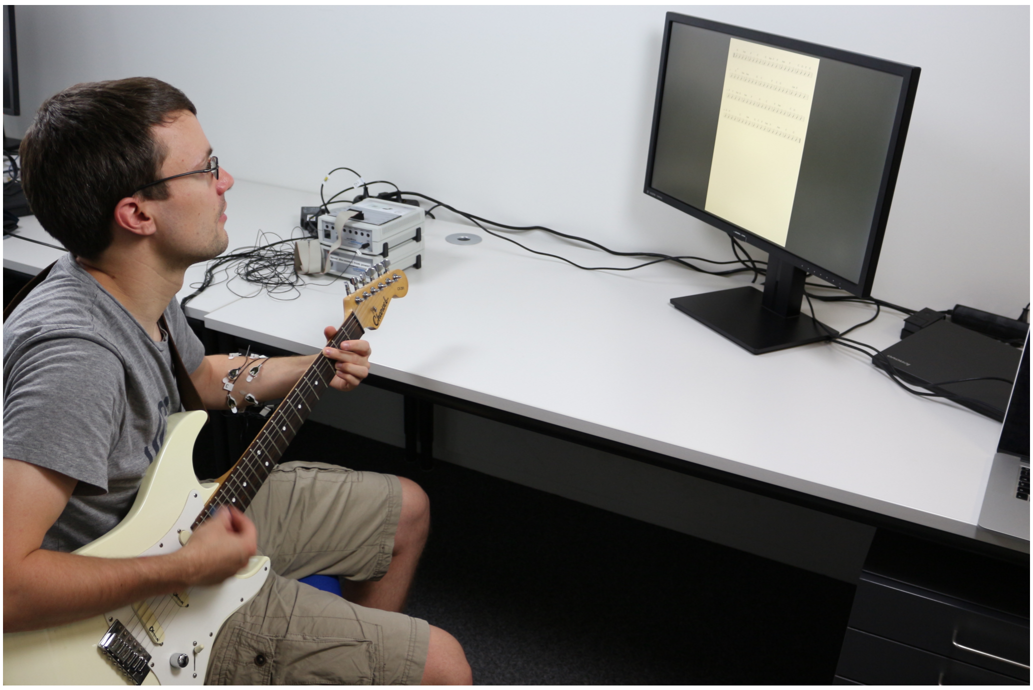


Figure 5.5: A picture showing the setup of the EMG guitar tutor system study

At the end of the study, we asked the subjects to fill out three questionnaires. The first and the second questionnaire both contained a 7-point Likert scale and the questions:

- "How well did you feel mastering the piece?" (*not well at all* to *very well*)
- "How correct did you feel you played?" (*not correct at all* to *very correct*)
- "How easy was it to learn the piece?" (*very hard* to *very easy*)
- "How enjoyable was it to learn the piece?" (*not enjoyable at all* to *very enjoyable*)

These questions are based on the work of Yuksel et al. [Yuk+16] who explored the usage of EEG signals within a piano tutor system. The first questionnaire referred to the system that was used first by the subject, the second questionnaire respectively referred to the second system. The third questionnaire asked questions regarding the ease of use and the realistic usage of the system. It included the questions:

- "Please rate your perceived accuracy of the system in evaluating your performance." (5-point Likert item from *Very bad* to *Perfect*)
- "How did you perceive the tempo adjustments the system made?" (5-point Likert item from *Did not suit my needs at all* to *Did perfectly suit my needs*)
- "How challenging were the system's tempo choices?" (5-point Likert item from *Too challenging* to *Too easy*)
- "Did the electrodes limit you in playing the guitar?" (5-point Likert item from *No* to *Could not play at all*)
- "Did the system help you to learn the chords? Why, or why not?" (free text)
- "Would you use the system to learn to play guitar? Why or why not?" (free text)
- "Imagine the system/electrodes could be integrated into your garments. How does this change your perceived usability of the system?" (free text)
- Further comments (free text)

Throughout the study we used the same preprocessing and feature extraction steps as outlined in Section 4.2. Our evaluation had shown that RMS and RMS ratios had yielded the best classifier performance. Therefore, we chose this feature set and used a window size of 250ms[3]. In the previous evaluation we neglected two windows. Contrary, we only neglect one window in this experiment as we now also have chord changes after one and not only after four beats. With the highest tempo of 90 BPM one segment has a length of approximately 667 ms ($\frac{60\,\text{s}}{90} = 0,\bar{6}$ s ≈ 667 ms). Thus, we can neglect

[3]We used the same window size as we did in Section 4.3.

at most one full window to keep at least one window for classification. If we apply these changes to our previous evaluation (the dataset from Chapter 4), the average F1 measure decreases from 87% (*SD=3%*) to 84% (*SD=3%*) regarding all chords. If we only measure the performance on the four chords we selected, the performance also decreases to 84% but with a standard deviation of 9%.[4] To counteract the introduced noise (due to chord-change artifacts), we set up a majority vote post-processing for the accuracy estimation. This improves the performance for slower tempos, that is when multiple windows can be used to predict a chord.

5.3 Results

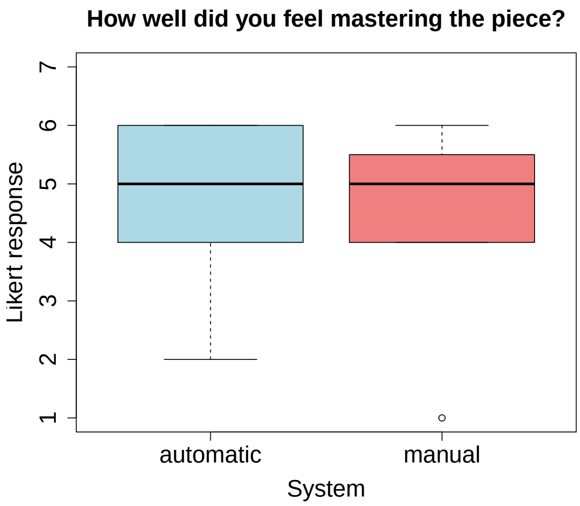
We collected data of 8 participants (2 female, 6 male) with a mean age of 22.6 (*SD=2.3*). Due to technical problems, we discarded one participant's data. All subjects reported normal sight and hearing. They all played the guitar before and knew how to finger the four chords.
We compared the automatic system answers with the manual system answers regarding the four questions. The respective boxplots are shown in Figure 5.6. A Wilcoxon Signed-ranks test showed that the reported Likert scores for all four questions were not significantly affected by the system the subjects were using. The detailed results of the test are shown in Table 5.1.

Question	$\bar{x}$	SD	V	p (apx.)
How well did you feel mastering the piece?	4.43	1.72	2.5	0.42
How correct did you feel you played?	4.14	1.77	5.0	0.28
How easy was it to learn the piece?	4.00	1.73	10.0	1.00
How enjoyable was it to learn the piece?	5.00	1.63	12.5	0.20

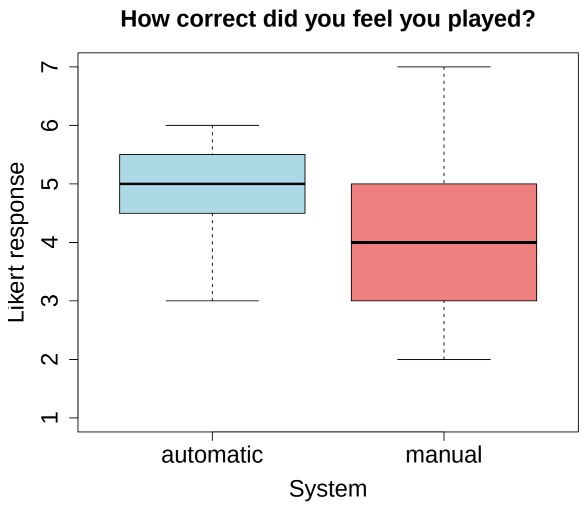
Table 5.1: Statistics and results of the Wilcoxon Signed-ranks test for the four questions including the mean, the standard deviation, the test statistic and the approximate p-value.

Figure 5.7 shows the boxplots of the Likert-items in the third questionnaire.

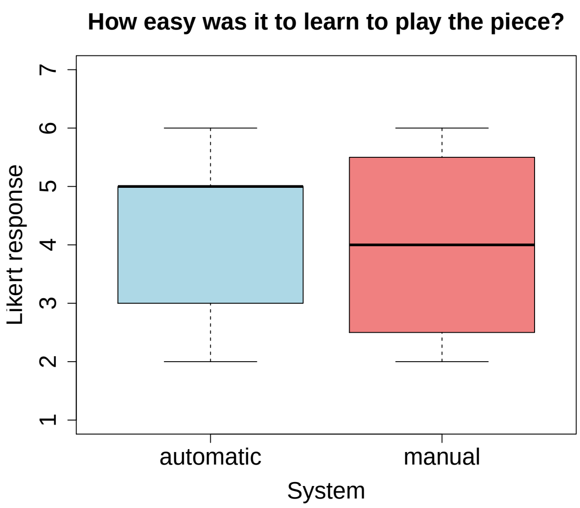
[4] This high variance can be explained by the classifier's problems regarding Am and C.



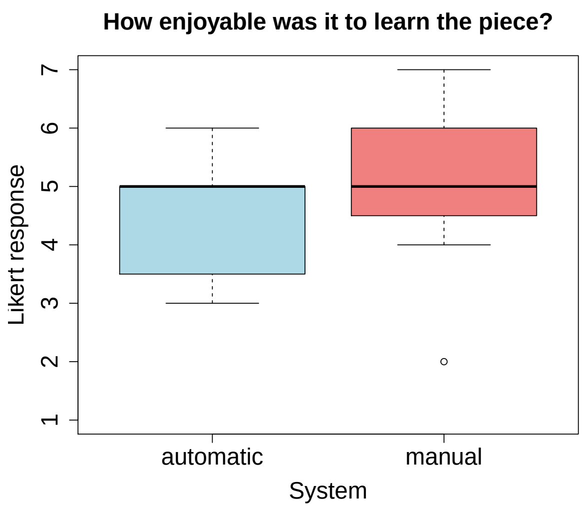
(a) From *not well at all* (1) to *very well* (7).



(b) From *not correct at all* (1) to *very correct* (7).

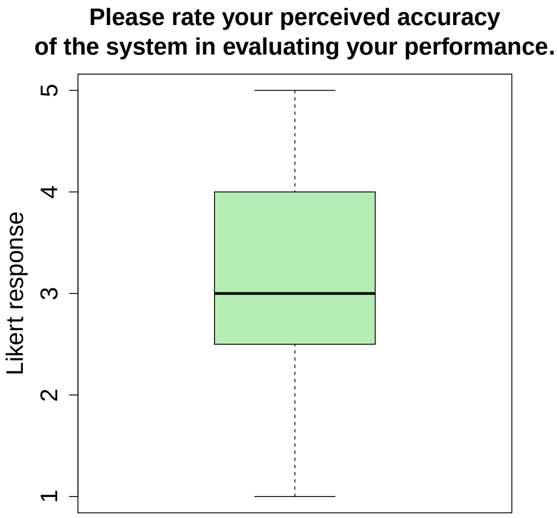


(c) From *very hard* (1) to *very easy* (7).

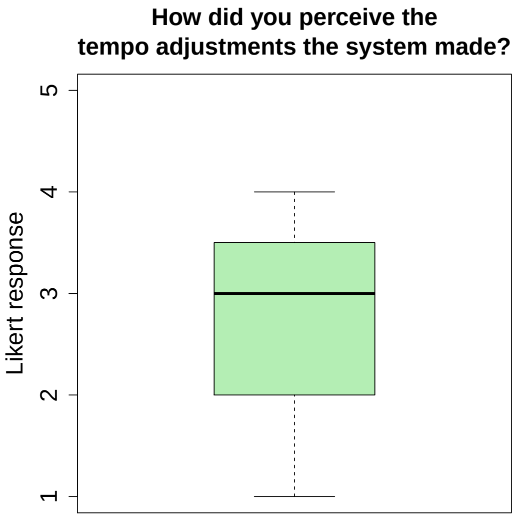


(d) From *not enjoyable at all* (1) to *very enjoyable* (7).

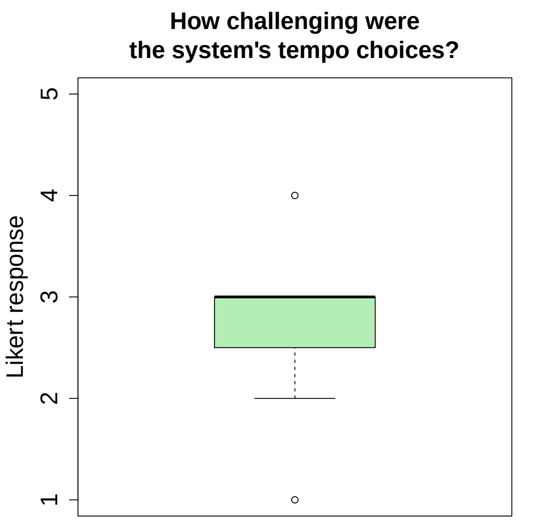
Figure 5.6: Boxplots of the compared first and second questionnaire's responses. The boxplots' reach from the lower to the upper quartile. The whiskers mark the furthest observation within a distance of 1.5 times the box size from the box. The small circles correspond to outliers.



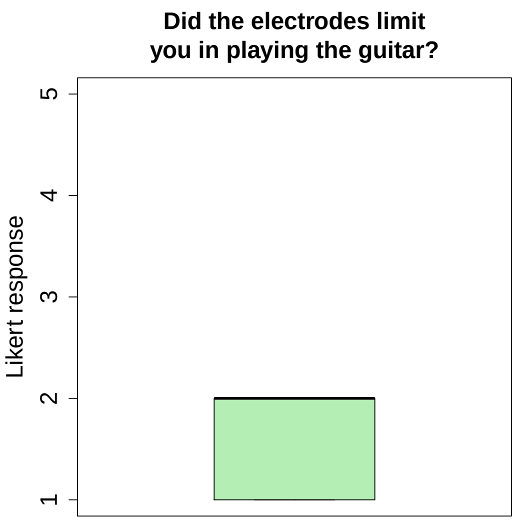
(a) From *Very Bad* (1) to *Perfect* (5).



(b) From *Did not suit my needs* (1) to *Did perfectly suit my needs* (5).



(c) From *Too challenging* (1) to *Too easy* (5).



(d) From *No* (1) to *Could not play at all* (5).

Figure 5.7: Boxplots of the third questionnaire's responses. The boxplots' reach from the lower to the upper quartile. The whiskers mark the furthest observation within a distance of 1.5 times the box size from the box. The small circles correspond to outliers.

Regarding the textual feedback for the question "Did the system help you to learn the chords? Why, or why not?" we received mixed answers. Those include:

> "*Yes it did. The given tempi helped to challenge me in play more quickly but still accurate.*" [P7[5]]

> "*Yes, because my own adjustment was probably too high, the system probably is more realistic.*" [P2]

> "*No, I think the system's suggested accuracy of playing the chords is not so precise.*" [P4]

> "*No direct feedback what chord was wrong, which chords where wrong or which transitions where wrong.*" [P6]

Among the answers for the question "Would you use the system to learn to play guitar? Why or why not?" we received:

> "*Yes because it's fun and it has the self learning aspect into it without having a teacher to keep telling me what to do.*" [P3]

> "*No, I did not have the feeling that the system supported me in my learning-process.*" [P4]

> "*Why not. The right hand is not checked. Yes, if it has more functions.*" [P6]

> "*Maybe at the beginning to learn the riffs and train to change the riffs. But you need much time to fix the electrodes, so it wouldn't be this often. Furthermore, it would be more joyful if there were known songs to play along.*" [P7]

For the question "Imagine the system/electrodes could be integrated into your garments. How does this change your perceived usability of the system?" we received answers including:

> "*That would be much easier and cooler.*" [P3]

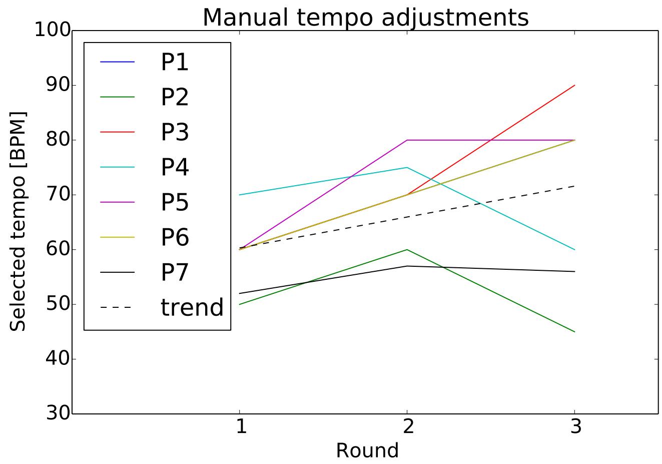
> "*It wouldnt change much.*" [P1]

> "*If it would not be at the upper forearm it would be better.*" [P6]

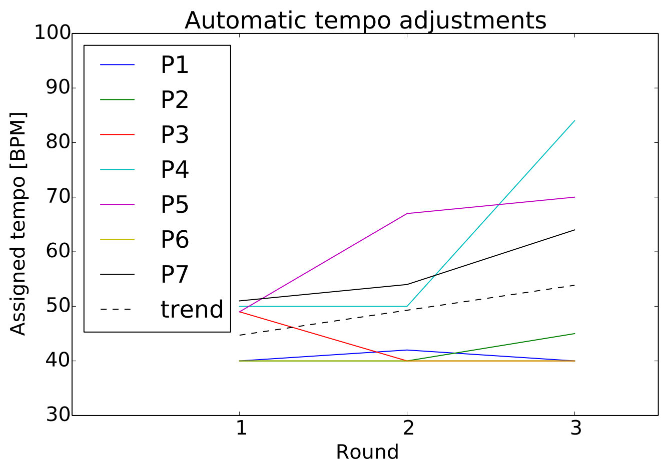
> "*I would probably use this system more often and would learn to play the guitar faster.*" [P7]

[5] We refer to the participants with P1 to P7.

We compared the tempo adjustments of the manual with the automatic system. One plot for each of the systems is shown in Figure 5.8. The respective accuracy estimates regarding the automatic system can be seen in Figure 5.9.



(a) Manual tempo adjustments after the first three rounds. The lines corresponding to participant 1 and participant 7 are identical. The dashed line visualizes a linear regression.



(b) Automatic tempo adjustments after the first three rounds. The dashed line visualizes a linear regression.

Figure 5.8: The plots show the results of the manual and automatic tempo adjustments per participant together with the plots of the respective linear regressions.

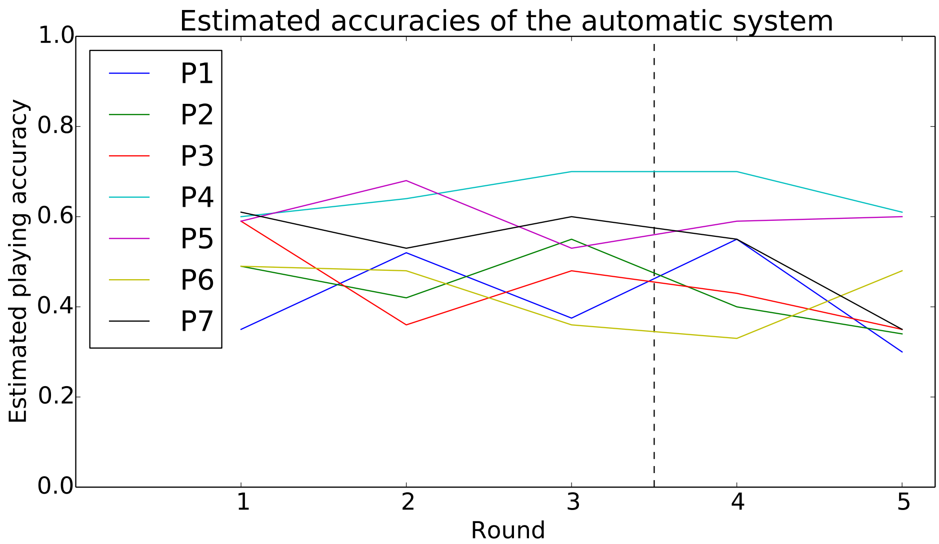


Figure 5.9: The playing accuracies estimated by the automatic system. The dashed vertical line highlights that after the third estimate the estimates did not influence the tempo adjustment any more as the fifth round was always played with 90 BPM.

To assess the classifier's performance and detect system failure, such as not correctly recognizing the played chords, we test the SVM on the second calibration while training it on the first one. The second calibration was conducted approximately 20 minutes after the first one. To interpret these values regarding the performance that the classifier reaches on single sets, we trained and tested the classifier on the first and the second set separately. The results are shown in Table 5.2. Additionally, we extracted the per class and per person F1-measures for the training on the first calibration and the test on the second one. The respective values are plotted in Figure 5.10.

Subject	F1 trained first, tested first	F1 trained second, tested second	F1 trained first, tested second
1	96	93	71
2	79	74	36
3	89	83	65
4	82	86	72
5	96	100	70
6	92	84	35
7	94	96	64

Table 5.2: Average F1-measure in percent for different train/test settings on the first and the second calibration

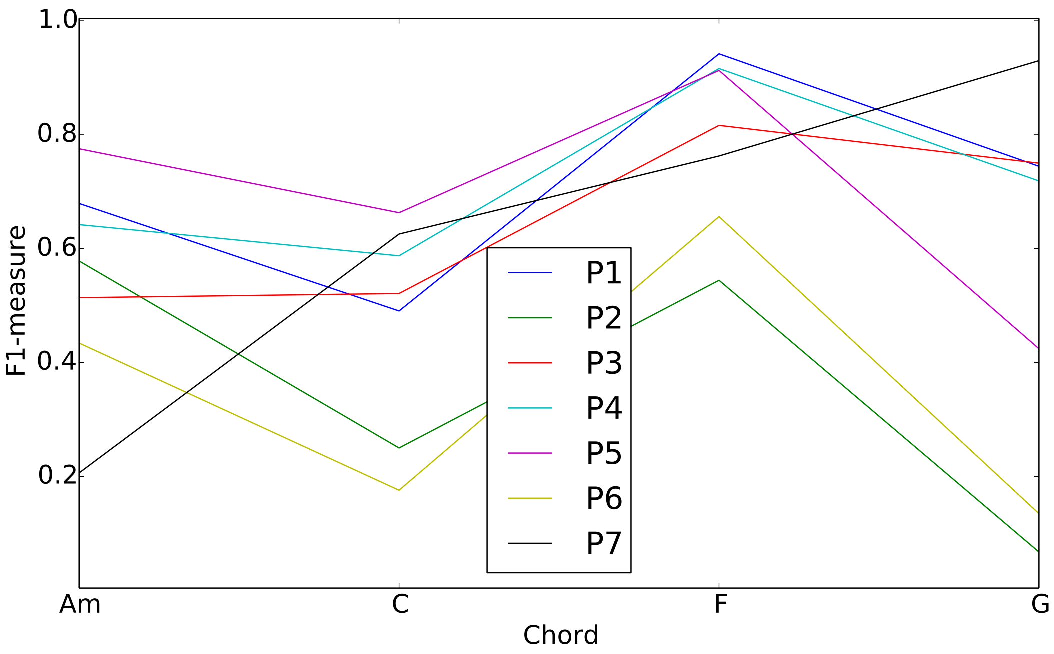
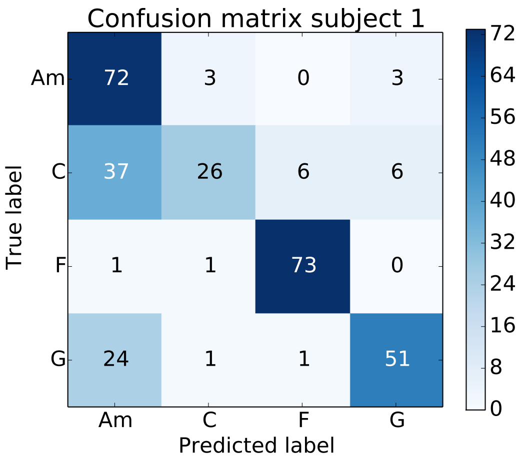
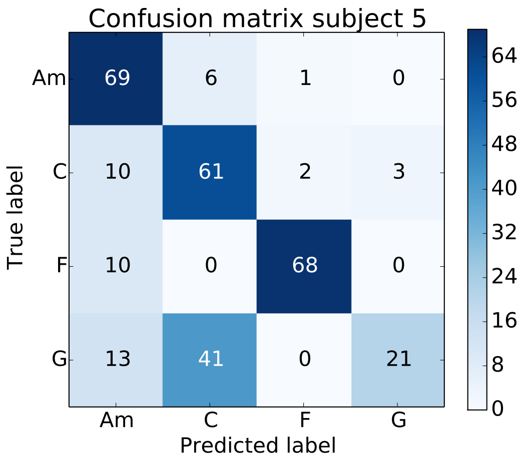


Figure 5.10: F1-measures per chord, plotted for all subjects separately. Trained on first calibration, tested on second calibration.

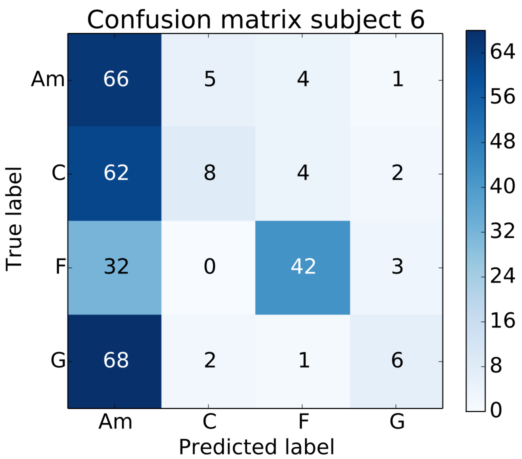
In order to see if there is a specific structure in the classifier's mistakes, we also plotted the confusion matrix for each subject. Figure 5.11 shows the matrices of Subject 1, 3, 5 and 6. We selected these subjects as their matrices show typical classes of mistakes.



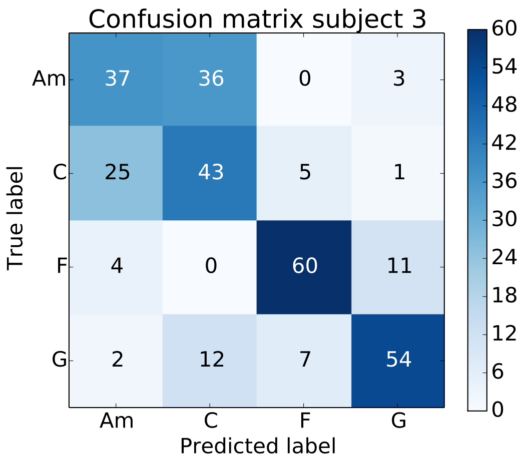
(a) Subject 1



(b) Subject 5



(c) Subject 6



(d) Subject 3

Figure 5.11: Exemplary confusion matrices of predictions of the second calibration using a classifier trained on the first calibration. The shown matrices correspond to typical classes of mistakes.

To evaluate the generalizability of our approach, we analyzed the inter-person classification performance[6] using the same leave-one-subject-out approach as in our first evaluation. The results for each subject are shown in Table 5.3.

Subject	Accuracy	Precision	Recall	F1
1	41	51	41	42
2	25	6	25	10
3	38	32	38	33
4	49	43	49	43
5	48	51	48	46
6	35	45	35	33
7	43	43	43	42

Table 5.3: Inter-person weighted average accuracy, precision, recall and F1-measure. For each subject a SVM was trained on all other subject's calibrations. All metrics are measured in percent.

In addition, we repeated the same analysis with either a principal component analysis (PCA) or a kernel PCA (using an RBF kernel) before the SVM. The PCA reduces the feature-vectors' dimensions while preserving the maximal possible variance. We introduced the number of dimensions of the PCAs as another dimension in our CV grid (see Section 4.2), covering the range from 1 to 9. In this CV, we neglected the polynomial kernel for computational efficiency as all CVs on this data set suggested either a linear or a RBF kernel[7]. The results are shown in Table 5.4. We also tested the PCA with an additional whitening transform which ensures uncorrelated components. However, this did not increase the performance for any of the participants compared with the classifier without any PCA.

[6]This means that the classifier is not trained and evaluated on one subject's data, but trained on all subjects except one and tested on the remaining subject.
[7]This refers to the kernel the SVM uses, not the kernel the kernel PCA uses.

	PCA		Kernel PCA (RBF)	
Subject	Optimal number of components	F1	Optimal number of components	F1
1	8	44 (+2)	9	53 (+11)
2	9	10 (+0)	9	10 (+0)
3	8	33 (+0)	9	38 (+5)
4	8	43 (+0)	9	40 (-3)
5	9	44 (-2)	8	47 (+1)
6	9	33 (+0)	9	34 (+1)
7	9	42 (+0)	9	29 (-13)

Table 5.4: Inter-person weighted average F1-measure in percent for an additional PCA or an additional Kernel PCA using an RBF kernel. For each subject, a SVM was trained on all other subject's calibrations. The differences to the SVM without dimensionality reduction are indicated by the values in parantheses.

5.4 Discussion

Questionnaires

The comparison of the manual and the automatic system with the aid of the two questionnaires reveals an unexpected behavior. Although the medians regarding the question "How well did you feel mastering the piece?" are equal, the boxplot related to "How correct did you feel you played?" suggests higher values for the automatic system. This is interesting because the automatic system underestimates the playing accuracy due to unavoidable misclassifications, but still gives the subjects the feeling of playing correctly. We suggest that the automatic system conveys a feeling of security by relieving the users of the need to correctly assess themselves. Additionally, we believe, that the subjects feel more confident about playing with the tutor system as it assigned lower tempos (with a mean of 57 ($SD=18$)) than the subjects chose themselves (mean of 66 ($SD=12$)). In contrast, the boxplot regarding "How enjoyable was it to learn the piece" suggests that the manual system was more enjoyable to the subjects. We believe, that this can be explained by the low reported accuracies that might irritate the subjects. Altogether the Wilcoxon Signed-ranks test did not reveal a significant effect of the system type on any of the questions. This can be attributed to the low sample size of seven participants and should be investigated in more detail in future work.
The free text answers of the third questionnaire give clear guidance what problems and what benefits the automatic system has. First of all, the accuracy of the system has to be improved. We think that an alternative electrode placement or a more dense electrode grid could advance the information that is contained within the EMG signal. Rotation- and shift-invariant features regarding the electrode positions could further improve the performance and facilitate the setup. If a person uses the system over a longer period of time the classifier's performance would implicitly improve by gathering more and more training data.
In addition to the improvement of accuracy, the subjects asked for more sophisticated features like chord- or chord-change-specific feedback. This matches with our vision of the system and encourages further development. Figure 5.12 and Figure 5.13 show our vision of a chord-specific and a chord-change-specific feedback.

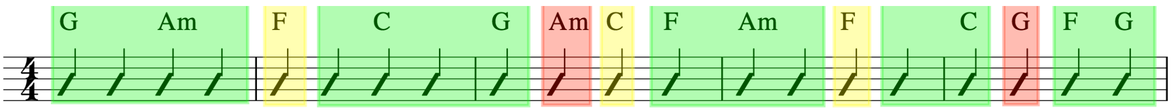


Figure 5.12: The figure shows our vision of an UI that provides chord specific feedback to the user. We map a faultless play of the chord to green, a slightly wrong play to yellow and a wrong play to red. The required data to generate the coloring can already be accessed in our system.

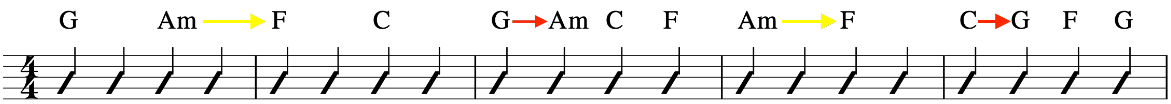


Figure 5.13: The figure shows a visualization of problematic chord changes colored in yellow (slightly problematic) and red (problematic). The required data can be obtained by analyzing the instances between a pair of chords.

Further, the answers regarding the question about electrodes integrated into garments suggest that such a system would not satisfy all users but still constitute a promising enhancement.

The analysis of the third questionnaire's Likert items showed that for the questions "Please rate your perceived accuracy of the system in evaluating your performance" and "How did you perceive the tempo adjustments the system made?" the subject's answers cover the whole range for the first question and the whole range except 5 for the second question. This indicates that the perception of the accuracy and the tempo adjustments was strongly dependent on the subject. The question "How challenging were the system's tempo choices?" was answered with 3 four times and with the values 1, 2 and 4 one time each. This suggests that on average, the tempo choices were neither too easy nor too challenging. However, the answers also show that this does not hold for all subjects. Regarding the question "Did the electrodes limit you in playing the guitar?" the answers solely included the values 1 and 2. This is something we did not expect, as we thought of the wires to be quite obtrusive. As the free text answers suggest that integrating the electrodes into garments would improve the user experience, the Likert responses indicate that the participants were not limited in playing the guitar, but wish for a system that can be set up faster.

Tempo adjustment

The comparison of tempo adjustments shows that the automatic system made slower tempo increments but only decreased the tempo in two rounds throughout all participants. With the manual system, four participants decreased their tempo. This indicates that the automatic system prevents tempo increases that are too demanding for the musician.
Our results in Chapter 4 showed that classifier performance is limited at approximately 87% accuracy given 8 classes. This indicates that an EMG-based tutoring system – as deployed here – needs to consider misclassifications and incorporate those into the tutoring assistance. We addressed this in the selection of our tempo mapping function. Nevertheless, more sophisticated feedback such as shown in Figure 5.12 requires a more fine-grained estimation-correction. We suggest to formulate this correction in form of a hidden Markov model (HMM) (see Figure 5.14) or a more complex dynamic Bayesian network. The respective probability distributions can be learned given a training set of observations of classifier decisions and ground truth labels.

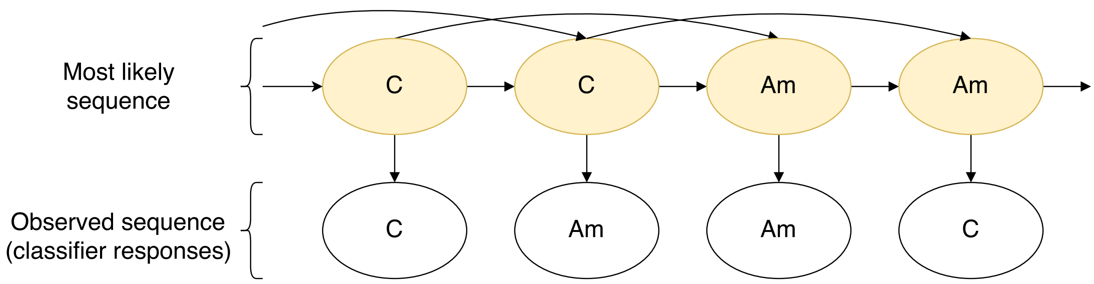


Figure 5.14: This figure shows a schematic depiction of the application of a second-order HMM to the correction of chord estimates. The most-likely sequence is calculated based on the estimated transition and observation probabilities.

The estimated playing accuracies we measured during this experiment were between 30% and 70%. One can observe predominantly high or low accuracies for some subjects (P6 has a throughout low accuracy, P4 a respectively high one). This can be attributed to subject-specific physiology and guitar playing skills. The accuracy of 70% shows that it is possible to classify the chord gestures, even if the chords are changed quickly. Vice versa, it is important to note that low accuracy estimates do not necessarily correspond to a low classifier performance. A 30% accuracy estimate may correspond to a 100% classifier accuracy, if the subject played 70% of the chords incorrectly. To evaluate the pure classifier performance, we have to evaluate it on the calibration data as it is free from playing errors.

Intra-Person Performance

Our tests on the first and the second calibration show that on the training data we reach high F1-scores for most of the participants (four participants over 90% and two over 80%). However, high scores on the training data do not correspond to high scores on the test set (the second calibration). This suggests that the classifier does not generalize well enough using our feature set. We analyzed the per class F1-measure for each subject to find out, for which classes the performance is low. Figure 5.10 shows that the chords Am and C have low scores with a mean of 55% and 47%. With a mean score of 54% the G chord's performance is even lower than the Am one's. Only the F chord reaches a relatively high score with a mean of 79%. To investigate how the mistakes are structured we have to consider the confusion matrices (see Figure 5.11). These let us identify three types of problems:

Complete misconception (see Figure 5.11(c)) The classifier is nearly unable to distinguish between classes (except F as it is a barre chord[8]). This indicates that the features drifted too much, e.g. due to a shifted electrode or a drastic change in conductivity. This type of error can be addressed on the hardware side or with a dynamic drift correction during performance.

Confusion of Am and C (see Figure 5.11(d)) The classifier makes most of its mistakes when distinguishing between Am and C. This can be explained by the similar hand posture of both chords. The presence of this problem in this study confirms that this problem is not an individual case but a systematic phenomenon (it also applies to Subject 4). We suggest that future work analysis the underlying physiology in detail to find a remedy to this problem.

Noise (see Figure 5.11(a) and (b)) In this case the classifier yields respectable performance. It still has misconceptions about certain chords, but these do not seem to follow an obvious explanation. We believe that these mistakes are caused by noise in the training and the test data. This does not refer to noise on the signal, but to the idea that every class has an underlying feature vector and every instance of that class is a noisy version of it. This problem can be addressed by extending the size of the training data.

[8] A barre chord is a chord where one or more fingers are used to press down multiple strings over a single fret. This produces a distinct EMG signal.

Inter-Person Performance

The leave-one-subject-out classification showed that the inter-person performance increased to a mean F1-score of 36% during our guitar tutoring study, contrary to our first evaluation featuring only three participants with a mean of 22%. This confirms our expectation that the inter-person performance increases with the number of subjects and rises the question for how many more subjects the score increases and to which value it converges.
Our analysis of the classifier with an additional PCA (or kernel PCA) showed that we cannot reach as high performance increases as the ones reported in literature [Zha+16][9]. For the PCA, the F1-measure increased by 2% for Subject 1, but also decreased by 2% for Subject 5. Thus, the average F1-measure stays the same. For the kernel PCA, we achieve a 11% higher F1 measure on the first participant, 5% more for the second, but also 13% less on the seventh. In total, this increases the F1-score by two 2%. Thus, we cannot observe a strong improvement by using a (kernel) PCA in general, but still an evident increase for single subjects and a slight increase on average using the kernel PCA. It is already remarkable that the average performance of the PCA stays the same if we train on an affine subspace. This encourages the analysis of further dimensionality reduction techniques and should be investigated in future work.

[9]They reported an accuracy improvement of approximately 10%.

6 Future Work

Within this thesis we identified several challenges that need to be tackled in order to build a real-world system. Our findings propose that to this end, three topics should be addressed in future work:

Wearable electrodes The electrode integration into garments could facilitate the setup and enhance the user experience. This could be realized in form of an EMG-sleeve. Existing work on EMG electrodes explores electrodes that are embedded in garments ([Fin+07], [Tae+07], [Kar+08]). Farina et al. [Far+10] already tested a sleeve covering the upper and lower arm for the control of prostheses. Thus, we believe that such garments will be available within near future and should be evaluated with respect to the requirements of our system.

Machine learning techniques From a machine learning perspective, our work offers multiple starting-points. One is the exploration of different features and dimensionality reductions methods on our data. Further, deep learning methods have been applied to EMG signals regarding various research avenues [MLY16]. Du et al. [Du+17] showed that deep domain adaption yields state of the art intra-session[1] and increased inter-session and inter-person performances. We believe that similar methods will increase the performance of our system. We further suggest to introduce classes that represent chord changes as we believe that fast gesture changes pose a challenge on the classifier. Another option is the reformulation of the problem as a regression which predicts how well a gesture is performed with respect to an optimal gesture.

Expansion of system functions Regarding the current functionality of our system we suggest an extension of the feedback to chord-changes and chord-specific feedback as shown in Section 5.4. This matches the participant's feedback. Further, the tempo adjustment should be revised in a thorough empirical validation to offer the users the best learning experience.

[1] This is the performance within one session, meaning that training and testing is done on a dataset that is recorded during one session.

We are confident that these points can push our system towards a real-world application. The effort in tackling these challenges also pays with respect to closely related scenarios. Possible applications include artistic usages such as an automatic chord playback. As an example, one could realize an accompanying bass-line that adapts to the fingered chord, even if only single or no notes of the chord are picked with the right hand. A different scenario is the usage within rehabilitation systems that encourage patients to practice specific movements in a playful way. Further, one could extend our prototype to a system that allows users with hand or finger injuries – or even missing limbs – to play guitar chords.

7 Conclusion

In this thesis, we explored the usage of electromyography (EMG) to enable new musician-instrument interaction. We focused on recognizing guitar chords by leveraging nerve signals of the upper forearm. Although EMG has been discussed extensively for a variety of hand-gestures, our approach uses domain-specific movements within a musician-instrument interaction system, which is novel to the field.

We implemented a signal processing chain and evaluated a broad range of analysis techniques that are proposed within related work. Our first study showed, that the best set of features – for our application – is the combination of the root mean squares and their respective ratios, reaching an average F1-score of 87%. As the other features did not perform as good as we expected based upon literature, we advise to reconsider the choice of features for each new application of EMG if maximum accuracy is vital. Furthermore, our study revealed a trade-off between real-time feasibility and classification performance that has to be considered when designing a real-time system.

We applied our findings to the development of an EMG guitar tutor system as one specific scenario which requires fine-grained manual interaction. The tutor system estimates the musicians' playing accuracies based on their nerve signals. Depending on their performance, it increases or decreases the playback tempo for the next repetition. We compared the automatic tempo adjustments with the adjustments the participants chose themselves in a study. The results showed that the automatic system did not reduce the ease of use and the electrodes did not limit the participants in playing the guitar. The recorded data allowed us to evaluate the underlying classifier. We found that the classifier performance is dependent on the subject and attribute this to the different physiology. We identified typical classes of mistakes (such as the confusion of the chords Am and C) and proposed respective solutions. Moreover, we showed that dimensionality reduction techniques can slightly enhance inter-person generalizability.

Altogether, this thesis showed that the classification of fine-grained, natural gestures is feasible using EMG. Our findings can be applied beyond the scope of musician-instrument-interfaces such as in supporting manual assembly tasks [Kos+16], helping children to learn how to write, or monitoring the exertion of a sportive motion. We believe that EMG interfaces are an emerging technology that will continue to grow into the consumer market and open new horizons in HCI.

Bibliography

[AIK11] M. R. Ahsan, M. I. Ibrahimy, O. O. Khalifa. “Electromygraphy (EMG) signal based hand gesture recognition using artificial neural network (ANN).” In: *Mechatronics (ICOM), 2011 4th International Conference On*. IEEE. 2011, pp. 1–6 (cit. on p. 23).

[AOY96] K. Akazawa, R. Okuno, M. Yoshida. “Biomimetic EMG-prosthesis-hand.” In: *Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE*. Vol. 2. IEEE. 1996, pp. 535–536 (cit. on p. 12).

[Abr+16] J. G. Abreu, J. M. Teixeira, L. S. Figueiredo, V. Teichrieb. “Evaluating Sign Language Recognition Using the Myo Armband.” In: *Virtual and Augmented Reality (SVR), 2016 XVIII Symposium on*. IEEE. 2016, pp. 64–70 (cit. on pp. 7, 12, 23).

[Amm+15] C. Amma, T. Krings, J. Böer, T. Schultz. “Advancing Muscle-Computer Interfaces with High-Density Electromyography.” In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI ’15. Seoul, Republic of Korea: ACM, 2015, pp. 929–938. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702501. URL: http://doi.acm.org/10.1145/2702123.2702501 (cit. on p. 21).

[Ars+05] B. Arslan, A. Brouse, J. Castet, J.-J. Filatriau, R. Lehembre, Q. Noirhomme, C. Simon. “Biologically-driven musical instrument.” In: *Proceedings of the Summer Workshop on Multimodal Interfaces (eNTERFACE’05)*. Faculté Polytechnique de Mons Mons, BL. 2005 (cit. on p. 13).

[Ars+06] B. Arslan, A. Brouse, J. Castet, R. Léhembre, C. Simon, J.-J. Filatriau, Q. Noirhomme. “A real time music synthesis environment driven with biological signals.” In: *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*. Vol. 2. IEEE. 2006, pp. II–II (cit. on pp. 7, 13).

[BVDS06] S. Bitzer, P. Van Der Smagt. “Learning EMG control of a robotic hand: towards active prostheses.” In: *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*. IEEE. 2006, pp. 2819–2823 (cit. on p. 12).

[Bis06] C. M. Bishop. *Pattern recognition and machine learning*. Springer, 2006. Chap. 1, pp. 33–38 (cit. on p. 35).

[But30] S. Butterworth. "On the theory of filter amplifiers." In: *Wireless Engineer* 7.6 (1930), pp. 536–541 (cit. on p. 18).

[CS09] C. Castellini, P. van der Smagt. "Surface EMG in advanced hand prosthetics." In: *Biological cybernetics* 100.1 (2009), pp. 35–47 (cit. on p. 12).

[DCT+13] M. Donnarumma, B. Caramiaux, A. Tanaka, et al. "Muscular Interactions Combining EMG and MMG sensing for musical practice." In: (2013) (cit. on pp. 7, 11, 13).

[DCT11] T. De Clercq, D. Temperley. "A corpus analysis of rock harmony." In: *Popular Music* 30.01 (2011), pp. 47–70 (cit. on p. 26).

[Du+17] Y. Du, W. Jin, W. Wei, Y. Hu, W. Geng. "Surface EMG-Based Inter-Session Gesture Recognition Enhanced by Deep Domain Adaptation." In: *Sensors* 17.3 (2017), p. 458 (cit. on p. 59).

[Far+10] D. Farina, T. Lorrain, F. Negro, N. Jiang. "High-density EMG E-textile systems for the control of active prostheses." In: *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE. 2010, pp. 3591–3593 (cit. on p. 59).

[Fin+07] T Finni, M Hu, P Kettunen, T Vilavuo, S Cheng. "Measurement of EMG activity with textile electrodes embedded into clothing." In: *Physiological measurement* 28.11 (2007), p. 1405 (cit. on p. 59).

[Guo+15] S. Guo, M. Pang, B. Gao, H. Hirata, H. Ishihara. "Comparison of sEMG-based feature extraction and motion classification methods for upper-limb movement." In: *Sensors* 15.4 (2015), pp. 9022–9038 (cit. on p. 23).

[Hir+16] D. Hiraoka, S.-i. Ito, M. Ito, M. Fukumi. "Japanese Janken recognition by support vector machine based on electromyogram of wrist." In: *Knowledge and Smart Technology (KST), 2016 8th International Conference on*. IEEE. 2016, pp. 114–119 (cit. on p. 12).

[Hof84] A. L. Hof. "EMG and muscle force: an introduction." In: *Human Movement Science* 3.1 (1984), pp. 119–153 (cit. on pp. 9–11).

[KMA08] J. Kim, S. Mastnik, E. André. "EMG-based Hand Gesture Recognition for Realtime Biosignal Interfacing." In: *Proceedings of the 13th International Conference on Intelligent User Interfaces*. IUI '08. Gran Canaria, Spain: ACM, 2008, pp. 30–39. ISBN: 978-1-59593-987-6. DOI: 10.1145/1378773.1378778. URL: http://doi.acm.org/10.1145/1378773.1378778 (cit. on p. 21).

[Kar+08] J. S. Karlsson, U. Wiklund, L. Berglin, N. Östlund, M. Karlsson, T. Bäcklund, K. Lindecrantz, L. Sandsjö. "Wireless monitoring of heart rate and electromyographic signals using a smart T-shirt." In: *Proceedings of International Workshop on Wearable Micro and Nanosystems for Personalised Health*. 2008 (cit. on p. 59).

[Kar14] B. Karlik. "Machine learning algorithms for characterization of EMG signals." In: *International Journal of Information and Electronics Engineering* 4.3 (2014), p. 189 (cit. on p. 23).

[Kos+16] T. Kosch, R. Kettner, M. Funk, A. Schmidt. "Comparing Tactile, Auditory, and Visual Assembly Error-Feedback for Workers with Cognitive Impairments." In: *Proceedings of the 18th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM. 2016, pp. 53–60 (cit. on p. 61).

[MLY16] S. Min, B. Lee, S. Yoon. "Deep learning in bioinformatics." In: *Briefings in Bioinformatics* (2016), bbw068 (cit. on p. 59).

[Moo+03] I. Moon, M. Lee, J. Ryu, M. Mun. "Intelligent robotic wheelchair with EMG-, gesture-, and voice-based interfaces." In: *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*. Vol. 4. IEEE. 2003, pp. 3453–3458 (cit. on pp. 7, 12).

[NHJ15] K. Nymoen, M. R. Haugen, A. R. Jensenius. "Mumyo–evaluating and exploring the myo armband for musical interaction." In: (2015) (cit. on p. 14).

[Ort12] M. Ortiz. "A Brief History of Biosignal-Driven Art From biofeedback to biophysical performance. eContact." In: *Biotechnological Performance Practice/Pratiques de performance biotechnologique* (2012) (cit. on p. 11).

[PJT16] G. Pomboza-Junez, J. H. Terriza. "Hand gesture recognition based on sEMG signals using Support Vector Machines." In: *Consumer Electronics-Berlin (ICCE-Berlin), 2016 IEEE 6th International Conference on*. IEEE. 2016, pp. 174–178 (cit. on p. 23).

[PK93] W. Putnam, R. B. Knapp. "Real-time computer control using pattern recognition of the electromyogram." In: *Engineering in Medicine and Biology Society, 1993. Proceedings of the 15th Annual International Conference of the IEEE*. IEEE. 1993, pp. 1236–1237 (cit. on p. 23).

[RHMY06] M. B. Reaz, M. Hussain, F. Mohd-Yasin. "Techniques of EMG signal analysis: detection, processing, classification and applications." In: *Biological procedures online* 8.1 (2006), pp. 11–35 (cit. on p. 23).

[SG82] G. N. Saridis, T. P. Gootee. “EMG pattern analysis and classification for a prosthetic arm.” In: *IEEE Transactions on Biomedical Engineering* 6 (1982), pp. 403–412 (cit. on p. 12).

[Sap+08] T. S. Saponas, D. S. Tan, D. Morris, R. Balakrishnan. “Demonstrating the Feasibility of Using Forearm Electromyography for Muscle-computer Interfaces.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’08. Florence, Italy: ACM, 2008, pp. 515–524. ISBN: 978-1-60558-011-1. DOI: 10.1145/1357054.1357138. URL: http://doi.acm.org/10.1145/1357054.1357138 (cit. on pp. 9, 16, 18, 20–23, 28).

[Sap+09] T. S. Saponas, D. S. Tan, D. Morris, R. Balakrishnan, J. Turner, J. A. Landay. “Enabling Always-available Input with Muscle-computer Interfaces.” In: *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*. UIST ’09. Victoria, BC, Canada: ACM, 2009, pp. 167–176. ISBN: 978-1-60558-745-5. DOI: 10.1145/1622176.1622208. URL: http://doi.acm.org/10.1145/1622176.1622208 (cit. on pp. 12, 13).

[Sch15] A. Schmidt. “Biosignals in Human-computer Interaction.” In: *interactions* 23.1 (Dec. 2015), pp. 76–79. ISSN: 1072-5520. DOI: 10.1145/2851072. URL: http://doi.acm.org/10.1145/2851072 (cit. on p. 12).

[Sil09] D. U. Silverthorn. “Physiologie. 4., aktualisierte Aufl.” In: *Müchen: Pearson* (2009) (cit. on p. 10).

[TK02] A. Tanaka, R. B. Knapp. “Multimodal interaction in music using the electromyogram and relative position sensing.” In: *Proceedings of the 2002 conference on New interfaces for musical expression*. National University of Singapore. 2002, pp. 1–6 (cit. on pp. 7, 9, 13).

[Tae+07] J. Taelman, T. Adriaensen, C. van der Horst, T. Linz, A. Spaepen. “Textile integrated contactless EMG sensing for stress analysis.” In: *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE. 2007, pp. 3966–3969 (cit. on p. 59).

[Tan] A. Tanaka. “The Use of Electromyogram Signals (EMG) in Musical Performance.” In: *CEC eContact* 14 () (cit. on p. 10).

[Ten+07] F. Tenore, A. Ramos, A. Fahmy, S. Acharya, R. Etienne-Cummings, N. V. Thakor. “Towards the control of individual fingers of a prosthetic hand using surface EMG signals.” In: *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE. 2007, pp. 6145–6148 (cit. on p. 12).

[Vra93] S. R. Vrana. “The psychophysiology of disgust: Differentiating negative emotional contexts with facial EMG.” In: *Psychophysiology* 30.3 (1993), pp. 279–286 (cit. on pp. 7, 12).

[Yuk+16] B. F. Yuksel, K. B. Oleson, L. Harrison, E. M. Peck, D. Afergan, R. Chang, R. J. Jacob. “Learn Piano with BACh: An Adaptive Learning Interface That Adjusts Task Difficulty Based on Brain State.” In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. Santa Clara, California, USA: ACM, 2016, pp. 5372–5384. ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858388. URL: http://doi.acm.org/10.1145/2858036.2858388 (cit. on p. 42).

[Zha+16] X. Zhai, B. Jelfs, R. H. Chan, C. Tin. “Short latency hand movement classification based on surface EMG spectrogram with PCA.” In: *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE. 2016, pp. 327–330 (cit. on p. 57).

Declaration

I hereby declare that the work presented in this thesis is entirely my own and that I did not use any other sources and references than the listed ones. I have marked all direct or indirect statements from other sources contained therein as quotations. Neither this work nor significant parts of it were part of another examination procedure. I have not published this work in whole or in part before. The electronic copy is consistent with all submitted copies.

place, date, signature