A REAL-TIME SYSTEM FOR HANDS-FREE GUITAR LOOP RECORDING AND DETECTION OF CHORDS AND RHYTHM STYLES

Arndt Eppler[1] **Andreas Männchen**[1] **Jakob Abeßer**[1]
Sascha Grollmisch[1] **Christof Weiß**[2]
[1] Semantic Music Technologies, Fraunhofer IDMT, Ilmenau, Germany
[2] International Audio Laboratories, Erlangen, Germany
jakob.abesser@idmt.fraunhofer.de

ABSTRACT

Guitar players commonly practice in a loop-based fashion, where a given chord sequence is played in a specific rhythmic style. While offline methods for extracting rhythmic and harmonic features from guitar recordings exist, transferring them to a real-time analysis in a live recording scenario is challenging. We present a hybrid system that simultaneously detects chords as well as the rhythmic style of a live guitar performance. Furthermore, we present an intuitive approach to interpret audio gestures on the guitar as control signal to predefine the tempo and meter of the guitar loop. In addition, we publish a newly created dataset of annotated guitar recordings to the research community as a public benchmark. The dataset covers various guitar models and recording conditions and can be used to evaluate the different analysis tasks covered here.

1. INTRODUCTION

Many music learning applications such as Songs2See [1], Yousician [2], and Rocksmith [3] support the guitar due to its high popularity. Recent technological advances in music education software were driven by Music Information Retrieval (MIR) research. A number of studies dealt with the automatic estimation of musical parameters from guitar recordings. This involves various estimation tasks such as polyphonic music transcription [4, 6, 9], chord estimation [2, 5], classification of playing styles [3, 6, 7], and the estimation of the applied fretboard position on the instrument (string and fret number) [1,2]. However, no real-time system for a simultaneous analysis of both harmonic and rhythmic properties from live guitar performances was presented so far. The ability to recognize tempo, meter, chord sequences, and rhythm styles in real-time allows to quickly select accompaniment tracks that match the recorded guitar loop both rhythmically and harmonically. The main challenge of such an analysis system is the large variety of sound characteristics among different guitar models, affecting for instance the magnitude relationships among different overtones. Also, the processing constraints of a real-time application limit the complexity of the underlying analysis algorithms. We focus on the recording scenario, where an isolated guitar track is recorded either directly from the instrument output or with a microphone.

In the following sections, we first describe a hybrid system for real-time detection of chord sequences and the rhythmic style in loop-like guitar recordings. The evaluation results of the individual processing steps will be presented and discussed in the final poster in detail. Secondly, we present a novel approach that allows the musician to remotely control music applications in an intuitive "hands-free" way by using acoustic playing gestures on the instrument instead of defining parameters in the software or hardware beforehand. Thirdly, we describe an extended version of the published IDMT-SMT-Guitar dataset, which includes monophonic and polyphonic guitar recordings from various guitar models and recording conditions as well as extensive annotations such as score information, rhythm style, chords, and playing techniques.

[1] http://www.songs2see.com/
[2] http://yousician.com/
[3] http://rocksmith.ubi.com/

2. PROPOSED ANALYSIS SYSTEM

The proposed real-time system for the simultaneous detection of chords and rhythm style in guitar recordings follows a modular design philosophy. Using a parallelized signal-flow, we combine four separate systems: parsing of the control signal, chord detection, onset detection, and rhythm style detection (see Figure 1).

At first, the guitarist plays a control signal to predefine the tempo and meter of the intended musical idea as explained in Section 3. These parameters define the underlying metrical structure for all further processing steps. After the control signal has been analyzed, the input signal is stored in two parallel buffers of a size of which is directly determined by the tempo. Each time a buffer is filled, the samples are passed on to the corresponding detection algorithm, as seen in Figure 1. The two processing paths are continuously running in parallel execution threads.

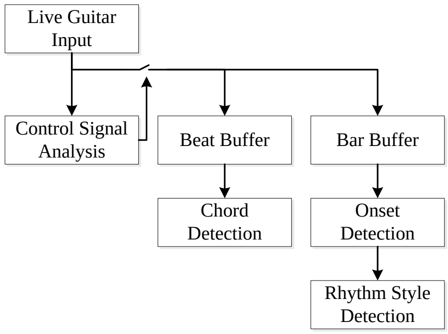

Figure 1. Outline of the signal path in the proposed system

2.1 Chord Detection

The chord detection algorithm uses tempo and beat position information to synchronize the audio processing with the underlying beat grid. The tuning frequency is estimated and beat-synchronous chroma features are computed while focussing on a guitar-relevant frequency range of 82.4 Hz to 1318.5 Hz. Silent frames and frames with a high flatness of the chroma vectors are discarded.

The final chord detection is performed by computing the Euclidean distances of the chroma vector to various binary chord templates, which range from from so-called power chords (fifth interval) and simple major and minor chords to complex jazz harmonies with upper-structure chord notes. In total, we use 24 types of chords, which leads to 288 separate templates due to 12 possible root notes for each chord. We assign a temporary chord label by choosing the chord that is represented by the template with the smallest distance value. This label is then refined by comparing it with the chord label assigned to the previous input block.

2.2 Onset Detection & Pattern Retrieval

We detect musical onsets in the recording based on the modified onset detection approach described in [6], which we use in a frame-wise manner to account for the real-time analysis. With the predefined tempo and meter information, we normalize the onset times to the beat duration and obtain a tempo-independent rhythmic representation of all played notes. Mapping the onsets on a consistent beat grid highly facilitates the detection of repeated note patterns. We use 16th notes as the smallest metrical division extended by triple 8th notes to account for triplets and dotted rhythms. In addition to the quantized onset times, we identified three parameters that shape the listeners perception of the rhythmic feeling: note energy, note duration, and harmonic texture. The harmonic texture indicates if one or more notes are being played at a particular onset time. We classify the harmonic texture using several low-level spectral features in combination with a Support Vector Machine (SVM) classifier.

In order to detect repeated patterns, we consider different hypothesis of patterns of one, two, or four bar length, which appear most often in music practice. In particular, we compute the Earth-Mover's Distance (EMD) [8] between different patterns and derive a final decision from a distance-based criterion.

2.3 Rhythm Style Detection

For the detection of the rhythmic style of a recorded guitar pattern, we use a classification approach based on various groups of rhythmic features in order to cope with patterns of different length. The first feature group contains the rhythmic information such as beat length, beat note value, number of beats, tempo, number of onsets, and beat division (duple or triple). In the second group we derive features from the onset position in the patterns such as dominance of different rhythmic levels and degree of syncopation. Furthermore, we add common vector statistics (min, max, mean, median, and variance) and the ratio between polyphonic and monophonic onsets. Again, an SVM classifier is used.

3. AUDIO CONTROL GESTURES

In a lot of state-of-the-art rhythm analysis algorithms, tempo octave errors and ambiguities in meter deduction remain unsolved problems. In the proposed system, we follow a patented semi-automatic approach, where the user predefines the tempo and meter using a specific control signal, which is played before the actual music recording starts and which is automatically interpreted by the system.

The control signal is played by muting the strings with the fretting hand (so called *dead notes*), while playing the beats of the desired meter for two bars in the desired tempo. With an accentuation on the first beats, the start of each bar is marked. The tempo is derived from the average beat distances.

4. IDMT-SMT-GUITAR DATASET

The IDMT-SMT-GUITAR dataset is a large database for automatic guitar transcription. [4] Seven different guitars in standard tuning were used with varying pick-up settings and different string measures to ensure a sufficient diversification in the field of electric and acoustic guitars.

The dataset contains several subsets of guitar recording with extensive annotations for different evaluation tasks. For instance, a set of 4700 note events in monophonic and polyphonic recordings cover various guitar playing techniques (3 plucking styles and 6 expression styles). Another set was created for the evaluation of the discussed chord recognition and rhythm style classification. This set contains recordings of 64 short musical pieces grouped by genre. Each piece has been recorded at two different tempi with three different guitars and is provided with a bit depth of 16 Bit. Annotations regarding onset positions, chords, rhythmic pattern length, and texture (monophony/polyphony) are included in various file formats.

[4] The dataset can be accessed on `http://www.idmt.fraunhofer.de/en/business_units/m2d/smt/guitar.html`

5. REFERENCES

[1] J. Abeßer. *From Sounds to Music and Emotions*, volume 7900, chapter Automatic String Detection for Bass Guitar and Electric Guitar, pages 333–352. Springer, London, UK, 2012.

[2] A. M. Barbancho, A. P. Klapuri, L. J. Tardon, and I. Barbancho. Automatic transcription of guitar chords and fingering from audio. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(3):915–921, 2012.

[3] Y.-P. Chen, L. Su, and Y.-H. Yang. Electric guitar playing technique detection in real-world recordings based on f0 sequence pattern recognition. In *Proceedings of the 16th International Society for Music Information Retrieval Conference (ISMIR)*, pages 708–714, Málaga, Spain, 2015.

[4] Xander Fiss and Andres Kwasinski. Automatic Real-Time Electric Guitar Audio Transcription. In *Proc. of the IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 373–376, Praha, Czech Republic, 2011.

[5] E.J. Humphrey and J.P. Bello. From music audio to chord tablature: Teaching deep convolutional networks toplay guitar. In *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6974–6978, May 2014.

[6] C. Kehling, J. Abeßer, C. Dittmar, and G. Schuller. Automatic tablature transcription of electric guitar recordings by estimation of score- and instrument-related parameters. In *Proceedings of the 17th International Conference an on Digital Audio Effects (DAFx)*, Erlangen, Germany, 2014.

[7] Tan Hakan Özaslan, Enric Guaus, Eric Palacios, and Josep Lluis Arcos. Attack Based Articulation Analysis of Nylon String Guitar. In *Proc. of the* 7^{th} *International Symposium on Computer Music Modeling and Retrieval (CMMR)*, pages 285–298, Málaga, Spain, 2010.

[8] Y. Rubner, C. Tomasi, and L. J. Guibas. The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2):99–121, 2000.

[9] Kazuki Yazawa, Daichi Sakaue, Kohei Nagira, Katsutoshi Itoyama, and Hiroshi G. Okuno. Audio-based Guitar Tablature Transcription using Multipitch Analysis and Playability Constraints. In *Proc. of the* 38^{th} *IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 196–200, Vancouver, Canada, 2013.