Automatic Guitar String Detection Based on the Inharmonicity Coefficient

Alexandros Iliadis*, Chrisoula Alexandraki
Department of Music Technology and Acoustics
School of Music and Optoacoustic Technologies
Hellenic Mediterranean University
*mta56@edu.hmu.gr (corresponding)

ABSTRACT

This paper explores the automatic detection of guitar strings, a topic closely related to the Music Information Retrieval task of automatic music transcription. The proposed methodology draws inspiration from previous research works focusing on the inharmonicity feature of guitar strings, while presenting a straightforward approach to associating isolated note samples with a string-fret pair. The implemented algorithm is based on the computation of the inharmonicity coefficient from monophonic guitar audio recordings via partial frequency tracking and curve fitting. Furthermore, an adaptation process requiring a minimal number of audio samples to adjust the algorithm to different guitars has been implemented. The evaluation process yielded positive results, encouraging the further improvement of the algorithm to handle more complex recordings.

Αυτόματη Ανίχνευση Χορδής Κιθάρας Βάσει του Συντελεστή Αναρμονικότητας

ΠΕΡΙΛΗΨΗ

Η παρούσα εργασία διερευνά την αυτόματη ανίχνευση χορδών κιθάρας, ένα θέμα στενά συνδεδεμένο με το ζήτημα Ανάκτησης Μουσικής Πληροφορίας της αυτόματης μεταγραφής μουσικής. Η προτεινόμενη μεθοδολογία αντλεί έμπνευση από προηγούμενες ερευνητικές εργασίες που επικεντρώνονται στο χαρακτηριστικό της αναρμονικότητας των χορδών κιθάρας, παρουσιάζοντας μία άμεση προσέγγιση για το συσχετισμό δειγμάτων απομονωμένων νοτών με κάποιο ζεύγος χορδής-τάστου. Ο υλοποιημένος αλγόριθμος βασίζεται στον υπολογισμό του συντελεστή αναρμονικότητας από ηχογραφήσεις μονοφωνικής κιθάρας μέσω του εντοπισμού μερικών συχνοτήτων και της προσαρμογής καμπυλών. Επιπλέον, υλοποιήθηκε μια διαδικασία που απαιτεί ελάχιστα ηχητικά δείγματα για την προσαρμογή του αλγορίθμου σε διαφορετικές κιθάρες. Η διαδικασία αξιολόγησης απέφερε θετικά αποτελέσματα, ενθαρρύνοντας την περαιτέρω βελτίωση του αλγορίθμου για την αναγνώριση πιο σύνθετων ηχογραφήσεων.

Introduction

The guitar is undoubtedly one of the most popular musical instruments in contemporary music. Its distinctive sound is produced by a set of strings that vibrate between two fixed points, the bridge and the nut. Most guitars feature six strings that follow the so-called standard tuning (i.e. E2-A2-D3-G3-B3-E4), however the tuning pegs provided on the headstock can be used to alter the pitch of any string by adjusting the corresponding string tension.

During performance, a guitarist changes the pitch of the strings by pressing on them on the fretboard, thus splitting their length and restricting their vibration between the bridge and a set of thin metal strips called frets. The frets are logarithmically spaced along the fretboard, enabling the progressive rise of the pitch by one semitone when moving the finger from the neck towards the bridge of the guitar. A fretboard can typically host between 19 and 24 frets, depending on the guitar type (e.g. classical, acoustic, or electric), while the 12th fret corresponds to a pitch that is an octave higher than the pitch of the associated open string.

Based on the above, it is easy to understand that it is possible to play notes of the same pitch in different positions, which is a property shared among stringed instruments. Therefore, a guitarist with a certain level of expertise can select among different variations of string-fret pairs to play a melody or a chord. This has led to the rising popularity of guitar tablatures, a form of musical notation that indicates the finger positions on the fretboard, as an alternative to traditional sheet music which depicts the pitch and the duration of the notes (Figure 0.1). Due to this popularity of tablature notation, the task of string detection has been introduced as a prerequisite of audio to tablature transcription within the MIR research community.

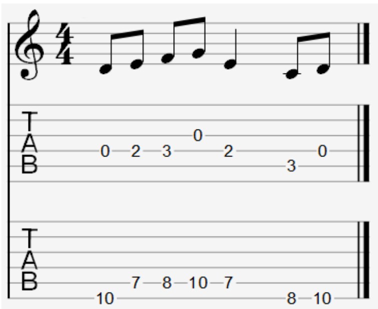

Figure 0.1 Conventional score vs tablature variations of the same melody

1. Background

1.1 Related Work

Guitar tablature transcription has been a topic of interest for many years. Relevant research initiatives generally emphasize on the estimation of note-related properties (i.e. pitch, onset and offset) and instrument-related parameters (e.g. fretboard position and performing style). A significant body of these works approach the task of string detection by computing a spectral feature which correlates to the physical properties of each string and is known as *inharmonicity* [1]-[8].

A different approach introduces the concept of String-Inverse Frequencies (SIFs) for determining the string-fret pair of played notes [9], while some works have applied playability constraints by utilizing probabilistic and optimization techniques such as Hidden Markov Models (HMMs) [10], Genetic Algorithms (GAs) [1], [11] and dynamic programming [12].

Regarding machine learning approaches, a few works employ feature extraction and classification techniques such as Support Vector Machines (SVMs) [3], [4], and [8]. Additionally, a deep learning approach deploying Convolutional Neural Networks (CNNs) has also been presented recently [13], paving the road for neural networks to possibly overtake traditional signal processing and statistical methods.

Besides these initiatives, multimodal approaches combining audio and video analysis to determine the fretboard position of played notes present an alternative perspective [14], [15]. Moreover, there are similar works which focus exclusively on video information and computer vision to achieve the same goal [16], [17].

1.2 Inharmonicity

Following the main direction of the relevant literature, this paper also adopts inharmonicity as a basis for detecting guitar strings. In general, the inharmonicity coefficient of a real string is mainly a result of its stiffness and it is defined by Equation 1.1, where Q is a material-specific property known as Young's Modulus, d is the diameter of the string, l its length, and T its tension [18].

$$\beta = \frac{\pi^3 Q d^4}{64 l^2 T} \tag{1.1}$$

The effect of inharmonicity on a vibrating string is the shift of the partial frequencies from harmonics to non-integer multiples of the fundamental frequency, as expressed by Equation 1.2, where f_k is the k^{th} partial and f_0 is the fundamental frequency of the ideal string without stiffness [19].

$$f_k = k f_0 \sqrt{1 + \beta k^2}\,, k \geq 1 \tag{1.2}$$

Specifically for guitar strings, it has been shown that the inharmonicity coefficient $\beta(s,n)$ of a string s pressed down on fret n can be derived from the inharmonicity coefficient $\beta(s,0)$ of the corresponding open string according to Equation 1.3 [6]. Thus, this equation allows the estimation of the inharmonicity coefficient across the entire fretboard, resulting in an inharmonicity curve per string, by computing only the coefficient of each open string.

$$\beta(s,n) = \beta(s,0) \cdot 2^{\frac{n}{6}} \tag{1.3}$$

2. Methodology

The methodology presented in this paper applies to monophonic guitar audio recordings (i.e. recordings without overlapping notes or chords) and assumes that the pitch, as well as the onset and offset timestamps of the played notes, are known in advance. From this point onward, the circumflex symbol (^) will denote computed values, while the tilde symbol (~) will indicate estimated values.

2.1 Inharmonicity Computation

The process of computing the inharmonicity coefficient consists of four stages, namely *preprocessing*, *partial tracking*, *curve fitting* and *postprocessing*.

The initial step of the *preprocessing* stage is to obtain an audio spectrogram *X(t,f)* from the audio waveform *x(t)* of a guitar recording. To achieve this, the Short-Time Fourier Transform (STFT) using a Hann window is applied. The window length and hop size are set to 2048 and 1024 samples respectively, which correspond to 46.4ms and 23.2ms of audio at a sample rate of 44100Hz. As the frequency deviations of the partials due to the inharmonicity effect are expected to be relatively small, the windowed signal for each audio block is zero-padded to 2^{18} samples to achieve a very high frequency resolution [6]. The next step is to iteratively retrieve the computed FFT audio spectrum *X(f)* for each audio block between the onset and offset timestamps of the target note and proceed to the next two stages, after normalizing *X(f)* to [0,1] and applying a threshold so that all values below 0.01 (1%) are zeroed [6].

In the *partial tracking* stage, the frequency values of up to a maximum of 15 partials on each previously obtained spectrum are determined. This process is carried out by centering search windows of length $\hat{f}_0/2$ at integer multiples of the computed fundamental frequency $\hat{f}_0$ of the target note [6]. The location of the highest spectral peak inside the range of the k^{th} window corresponds to the detected partial $\hat{f}_k$. Here, it is assumed that the fundamental frequency $\hat{f}_0$ is provided by a pitch detection algorithm. If the actual fundamental frequency f_0 of the played note is known instead (e.g. from dataset annotations), $\hat{f}_0$ is computed around it similarly to the rest of the partials, essentially being identical to $\hat{f}_1$.

The purpose of the *curve fitting* stage is to compute the inharmonicity coefficient by utilizing the previously detected sets of partials. To achieve this, the formula of Equation 1.2 is initially rearranged to obtain the equivalent expression of Equation 2.1. The inharmonicity coefficient β can then be computed through non-linear least squares [3], as described in Equation 2.2. This curve fitting process is visualized in Figure 2.1 where the ideal curve $y = k^2$ is also displayed.

$$\left(\frac{f_k}{f_0}\right)^2 = k^2 + \beta k^4, k \geq 1 \tag{2.1}$$

$$\left(\frac{\hat{f}_k}{\hat{f}_0}\right)^2 \approx k^2 + \hat{\beta} k^4, k \geq 1 \tag{2.2}$$

After computing a provisional inharmonicity coefficient value from every audio block between the onset and offset timestamps of the target note, the *postprocessing* stage of the methodology involves initially removing those coefficients that fall outside the value range of $[10^{-7}, 10^{-2}]$, considering them as outliers. Subsequently, the median of the remaining values is calculated, ultimately providing a robust representation of the inharmonicity coefficient over the note duration.

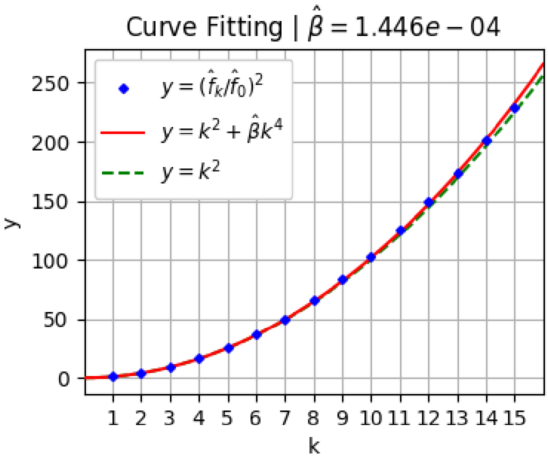

Figure 2.1 Visualization of the curve fitting process

2.2 Inharmonicity Estimation

The previous subsection covered the computation of the inharmonicity coefficient for a single note based on Equation 1.2. In contrast, the current subsection describes the estimation of the inharmonicity coefficient across the guitar fretboard, relying on Equation 1.3, in order to establish a basis for identifying notes as string-fret pairs. A more generalized form of this equation is presented in Equation 2.3.

$$\tilde{\beta}(s,n) = \hat{\beta}(s,0) \cdot 2^{\frac{an+b}{6}} \tag{2.3}$$

The introduced coefficients a and b allow the adaptation of the algorithm to different guitars. This is achieved through the computation of the inharmonicity coefficient on additional frets, by utilizing audio samples of multiple notes per string. Specifically, four adaptation schemes which are shown in Table 2.1 have been proposed [1]. In these schemes, the unknown values x, x_1, and x_2 of the coefficients a and b are calculated by initially computing the inharmonicity coefficient on the corresponding frets and then solving Equation 2.3.

Table 2.1 Proposed Adaptation Schemes

Scheme	**Frets**	**Coefficient *a***	**Coefficient *b***
1Fret	{0}	1	0
2FretA	{0,12}	x	0
2FretB	{0,12}	1	x
3Fret	{0,3,12}	x_1	x_2

2.3 String-Fret Identification

After estimating the inharmonicity coefficient across a guitar fretboard using Equation 2.3, the system is ready to perform automatic string detection of incoming guitar notes. Assuming that a pitch detection algorithm provides the fundamental frequency of a note, a set *C* of candidate string-fret pairs that produce the same pitch is then defined and the inharmonicity coefficient of the note is computed as outlined in Subsection 2.1. An identification process then determines the candidate pair that minimizes the absolute distance between the computed and the corresponding estimated inharmonicity coefficient, as described in Equation 2.4.

$$(\hat{s}, \hat{n}) = \arg \min_{(s,n) \in C} \left(\left| \hat{\beta} - \tilde{\beta}(s,n) \right| \right) \tag{2.4}$$

3. Evaluation

3.1 Dataset

For the algorithm evaluation, three distinct subsets from the IDMT-SMT-Guitar dataset were utilized [4]. These subsets correspond to separate raw audio recordings acquired from the bridge, middle, and neck pickups of a standard-tuned Ibanez Power Strat electric guitar. Each subset comprises one isolated note recording for every string-fret pair of the guitar up to the 12th fret, resulting in a total of 78 samples per subset. The dataset includes ground-truth annotations that specify the string, fret, pitch, onset, and offset for each note sample, and is therefore appropriate for evaluating the presented methodology.

3.2 Results

Note samples that were utilized during the adaptation phase (Subsection 2.2) or have only one candidate string-fret pair were excluded from the evaluation process, as their correct identification is essentially predetermined. The computed and estimated inharmonicity curves for four pickup-scheme configurations are displayed in Figure 3.1. The identification accuracy for all configurations is presented in Table 3.1. Notably, the neck pickup exhibits an intriguing drop in accuracy, which may be attributed to the lack of higher overtones in the corresponding audio signals compared to the middle and bridge pickups. However, the algorithm appears to be fairly accurate overall, with schemes *3Fret* and *2FretA* performing slightly better than *1Fret* and *2FretB*.

Table 3.1 Identification Accuracy

Scheme	**Bridge**	**Middle**	**Neck**	**Overall**
1Fret	98.41%	92.06%	<u>92.06%</u>	94.17%
2FretA	98.28%	<u>100%</u>	91.38%	96.55%
2FretB	94.83%	98.28%	87.93%	93.68%
3Fret	<u>100%</u>	<u>100%</u>	90.57%	<u>96.85%</u>

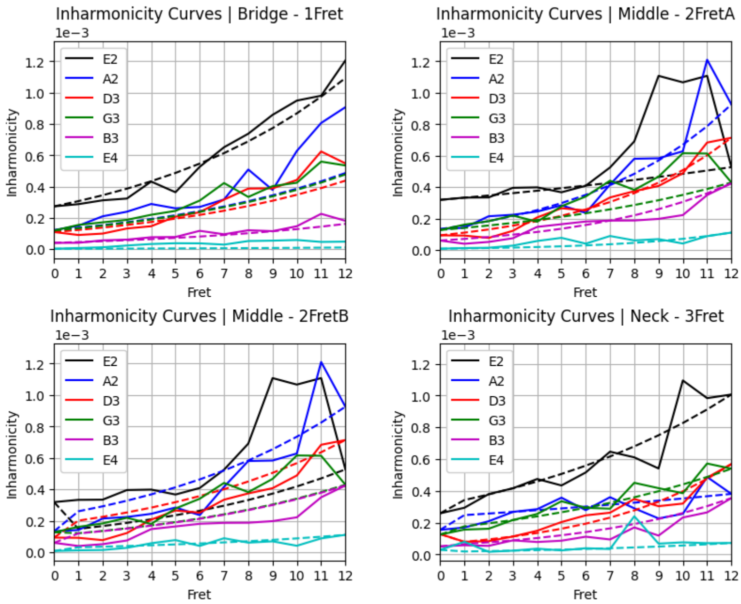

Figure 3.1 Computed (solid) and estimated (dashed) inharmonicity curves

4. Conclusions

This paper presents an algorithm for guitar string detection based on the computation of the inharmonicity coefficient and applicable to monophonic guitar audio recordings. The proposed methodology consolidates elements of the relevant literature, while pursuing simplicity and comprehensibility. The evaluation results on a well-known dataset of guitar recordings were deemed satisfactory and inspire future investigations for further improving and evaluating the effectiveness of the algorithm on additional datasets, such as the *Guitarset* [20], as well as for assessing its performance for real-time tablature transcriptions.

5. References

[1] G. Bastas, S. Koutoupis, M. Kaliakatsos-Papakostas, V. Katsouros, and P. Maragos "A Few-Sample Strategy for Guitar Tablature Transcription Based on Inharmonicity Analysis and Playability Constraints" in *International Conference on Acoustics, Speech and Signal Processing* (2022)

[2] J. J. Michelson, T. M. Sullivan, and R. M. Stern "Automatic Guitar Tablature Transcription from Audio using Inharmonicity Regression and Bayesian Classification" in *145th Audio Engineering Society International Convention* (2018)

[3] J. Abeßer, "Automatic String Detection for Bass Guitar and Electric Guitar" in *Lecture Notes in Computer Science* (2013)

[4] 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 *International Conference on Digital Audio Effects* (2014)

[5] J. M. Hjerrild and M. Grasboll Christensen "Estimation of Guitar String, Fret and Plucking Position Using Parametric Pitch Estimation" in *International Conference on Acoustics, Speech and Signal Processing* (2019)

[6] I. Barbancho, L. J. Tardon, S. Sammartino, and A. M. Barbancho "Inharmonicity-Based Method for the Automatic Generation of Guitar Tablature" IEEE Transactions on Audio, Speech and Language Processing, **20** (2012)

[7] J. M. Hjerrild, S. Willemsen, and M. G. Christensen "Physical Models for Fast Estimation of Guitar String, Fret and Plucking Position" in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (2019)

[8] C. Dittmar, A. Mannchen, and J. Abeßer "Real-Time Guitar String Detection for Music Education Software" in *International Workshop on Image Analysis for Multimedia Interactive Services* (2013)

[9] T. Geib, M. Schmitt, and B. Schuller "Automatic Guitar String Detection by String-Inverse Frequency Estimation" in *Lecture Notes in Informatics* (2017)

[10] A. M. Barbancho, L. J. Tardón, I. Barbancho, and A. Klapuri "Automatic Transcription of Guitar Chords and Fingering from Audio", IEEE Transactions on Audio, Speech and Language Processing, **20** (2012)

[11] D. R. Tuohy and W. D. Potter "A Genetic Algorithm for the Automatic Generation of Playable Guitar Tablature" in *International Computer Music Conference* (2005)

[12] K. Yazawa, D. Sakaue, K. Nagira, K. Itoyama, and H. G. Okuno "Audio-Based Guitar Tablature Transcription using Multipitch Analysis and Playability Constraints" in *International Conference on Acoustics, Speech and Signal Processing* (2013)

[13] A. Wiggins and Y. Kim "Guitar Tablature Estimation with a Convolutional Neural Network" in *20th International Society for Music Information Retrieval Conference* (2019)

[14] M. Paleari, B. Huet, A. Schutz, and D. Slock "A Multimodal Approach to Music Transcription" in *International Conference on Image Processing* (2008)

[15] A. Hrybyk and Y. Kim "Combined Audio and Video Analysis for Guitar Chord Identification" in *11th International Society for Music Information Retrieval Conference* (2010)

[16] C. Kerdvibulvech and H. Saito "Vision-Based Guitarist Fingering Tracking using a Bayesian Classifier and Particle Filters" in *Lecture Notes in Computer Science* (2007)

[17] A. M. Burns and M. M. Wanderley "Visual Methods for the Retrieval of Guitarist Fingering" in *International Conference on New Interfaces for Musical Expression* (2006)

[18] H. Fletcher, E. D. Blackham, and R. Stratton "Quality of Piano Tones" The Journal of the Acoustical Society of America, **34** (1962)

[19] H. Järveläinen, T. Verma, and V. Välimäki "The Effect of Inharmonicity on Pitch in String Instrument Sounds" in *International Computer Music Conference* (2000)

[20] Q. Xi, R. M. Bittner, J. Pauwels, X. Ye, and J. P. Bello, "Guitarset: A Dataset for Guitar Transcription" in 19th International Society for Music Information Retrieval Conference, pp. 453–460 (2018)