Automatic Guitar String Detection Based on Inharmonicity

Alexandros Iliadis
MSc in Sound & Music Technologies
Hellenic Mediterranean University
mta56@edu.hmu.gr

ABSTRACT

This paper explores the automatic detection of guitar strings, a topic closely related to the Music Information Retrieval (MIR) task of automatic music transcription. The proposed methodology draws inspiration from previous works focusing on the inharmonicity feature of guitar strings, while presenting a simplistic approach to associating note instances with a string-fret pair. Essentially, the implemented algorithm computes the inharmonicity coefficient from monophonic guitar audio recordings by means of partial tracking and curve fitting, while the estimation process requires minimal data for adaptation. The system's evaluation yielded encouraging results in favor of extending and expanding the current research in future work.

1. INTRODUCTION

The guitar is undeniably 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 6 strings that follow the so-called standard tuning (i.e. E2-A2-D3-G3-B3-E4), but the tuning pegs on the headstock can be used to alter the pitch of any string by adjusting the corresponding string tension. During a performance, the guitarist changes the pitch of the strings by pressing on them, thus restricting their vibration between the bridge and a set of thin metal strips called frets, essentially shortening their length. The frets are logarithmically spaced along the fretboard, enabling the consecutive raise of the pitch by one semitone. A fretboard can typically host between 19 and 24 frets, depending on the guitar type (e.g. classical, acoustic, or electric), and 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 can be deduced that on the guitar it is possible to play notes of the same pitch on multiple positions, a fact that holds true for most stringed instruments. Consequently, 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 player's finger positions on the fretboard, in contrast to traditional sheet music which focuses on the pitch and the duration of the notes (Figure 1). As a result, the need for string detection in the Music Information Retrieval (MIR) task of automatic music transcription from guitar recordings has emerged.

Audio Signal Processing ASP-2024, February 7, 2024, Rethymno, Greece.

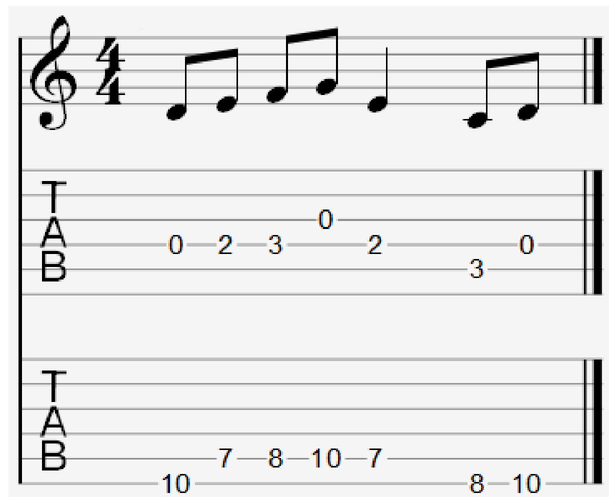

Figure 1. Sheet music and tablature variations of a melody (numbers correspond to frets, 0 indicates an open string)

2. BACKGROUND

2.1 Related Work

Guitar tablature transcription has been a topic of interest for many years, with related works emphasizing on the estimation of note-related parameters such as pitch, onset and offset, and instrument-related parameters such as fretboard position and playing style. Regarding the focus of this paper on string detection, a significant portion of published research adopts various audio-based methodologies which exploit the spectral feature of real strings known as *inharmonicity* [1-8]. Another interesting and unconventional approach introduces the concept of String-Inverse Frequencies (SIFs) for determining the string-fret pair of played notes [9]. 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]. Following the path of machine learning, a few approaches have been based on feature extraction and classification with Support Vector Machines (SVMs) [3], [4], [8]. Furthermore, 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 in the future. Lastly, it's worth mentioning that there have also been multimodal approaches combining audio and video analysis to determine the fretboard position of played notes [14], [15], and even solely video-based approaches by exclusive means of computer vision [16], [17].

2.2 Inharmonicity

Following the direction of many related works, this paper also adopts the 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:

$$\beta = \frac{\pi^3 Q d^4}{64 l^2 T} \quad (1)$$

where Q is a material-specific property known as Young's Modulus, d is the string's diameter, l its length, and T its tension [18]. The effect of the 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:

$$f_k = k f_0 \sqrt{1 + \beta k^2}, k \geq 1 \quad (2)$$

where f_k is the k^{th} partial and f_0 is the fundamental frequency of the string without stiffness [19]. In the context of guitar strings, it has been deduced that the coefficient $\beta(s,n)$ of a string s pressed down on fret n is related to the coefficient $\beta(s,0)$ of the corresponding open string [6], according to:

$$\beta(s,n) = \beta(s,0) \cdot 2^{\frac{n}{6}} \quad (3)$$

This equation enables the estimation of the inharmonicity across the entire fretboard, resulting in an inharmonicity curve per string, by computing only the coefficients of the open strings.

3. METHODOLOGY

The proposed methodology is applicable on monophonic guitar audio recordings and assumes that the pitch, as well as the onset and offset timestamps of the played notes, are known.

3.1 Inharmonicity Computation

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

3.1.1 Preprocessing

The initial step of this stage is to obtain an audio spectrogram $X(t,f)$ from the associated audio waveform $x(t)$. To achieve this, the Short-Time Fourier Transform (STFT) using a Hann window is applied. The window and overlap lengths are set to 2048 and 1024 samples respectively, which correspond to 46.4ms and 23.2ms of audio at a sample rate of 44100kHz. Moreover, since the frequency deviations of the partials due to the inharmonicity effect are relatively small, the FFT length is set to 2^{18} samples to achieve a very high frequency resolution [6]. The next step is to obtain the associated audio spectrum $X(f)$ for every audio block between the onset and offset timestamps and proceed to the next two stages, after normalizing each spectrum to [0,1] and applying a threshold so that all values below 0.01 (1%) are zeroed [6].

3.1.2 Partial Tracking

In this stage, the locations of up to a maximum of 15 partials on every previously obtained spectrum are determined. The process is carried out by centering search windows of length $\widehat{f}_0/2$ at integer multiples of the computed fundamental frequency $\widehat{f}_0$ [6]. The location of the highest spectral peak inside the k^{th} window's range corresponds to the detected partial $\widehat{f}_k$. Regarding $\widehat{f}_0$, it is considered that a pitch detection algorithm can provide the system with the required information. If the actual fundamental frequency f_0 of the played note is known instead, $\widehat{f}_0$ is computed around it in a similar fashion to the rest of the partials.

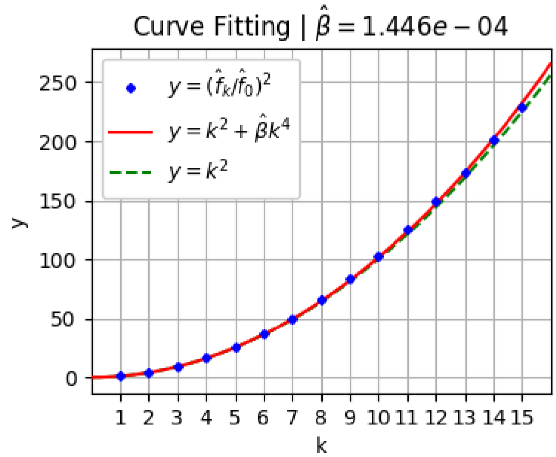

Figure 2. Visualization of the curve fitting process

3.1.3 Curve Fitting

The purpose of this stage is to compute the inharmonicity coefficient by utilizing the previously detected sets of partials. To achieve this, Equation (2) shall first be rewritten as:

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

The inharmonicity coefficient can then be computed through a non-linear least squares regression [3], as described in:

$$\left(\frac{\hat{f}_k}{\hat{f}_0}\right)^2 \approx k^2 + \hat{\beta} k^4, k \geq 1 \quad (5)$$

This curve fitting process is visualized in Figure 2 where the ideal curve $y = k^2$ is also shown.

3.1.4 Postprocessing

After computing the inharmonicity coefficient for every audio block between the onset and offset timestamps, the last stage of the methodology entails initially removing any outliers that fall outside the range $[10^{-7}, 10^{-2}]$. Subsequently, the median of the remaining values is calculated, providing a robust representation of the inharmonicity coefficient over the note duration.

3.2 Inharmonicity Estimation

The previous subsection covered the computation of the inharmonicity coefficient for a single note played on a string-fret pair based on Equation (2). In contrast, the current subsection shifts the focus on the estimation of the inharmonicity coefficients across the guitar fretboard relying on Equation (3). To begin, a more generalized form of this equation will be presented:

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

The introduced coefficients a and b enable a potentially better adaptation of the system to a specific guitar string by computing the inharmonicity coefficients on additional frets [1]. Specifically, four adaptation schemes are proposed:

- **1Fret:** {0} ($a = 1, b = 0$)
- **2FretA:** {0,12} ($\boldsymbol{a = x}, b = 0$)
- **2FretB:** {0,12} ($a = 1, \boldsymbol{b = x}$)
- **3Fret:** {0,3,12} ($\boldsymbol{a = x_1}, \boldsymbol{b = x_2}$)

In the above schemes, the unknown coefficients a or/and b are calculated by initially computing the inharmonicity coefficient for the corresponding extra fret(s) and then solving Equation (6).

3.3 String-Fret Classification

After going through the adaptation phase, the system is ready to perform automatic string detection of incoming played notes. Assuming, once again, 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. Subsequently, the inharmonicity coefficient is computed as outlined in Subsection 3.1. The classification process then identifies the candidate pair that minimizes the absolute distance between the computed and the corresponding estimated inharmonicity coefficient, as notated in:

$$(\hat{s}, \hat{n}) = \arg \min_{(s,n) \in C} \left(\left| \hat{\beta} - \tilde{\beta}(s, n) \right| \right) \tag{7}$$

4. EVALUATION

4.1 Dataset

For the system 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. Furthermore, the dataset includes annotations that specify the string, fret, pitch, onset, and offset for each note sample, thus ensuring the fulfillment of all prerequisites outlined in Section 3.

4.2 Results

For unbiased results, any note samples that were utilized during the adaptation phase or have only one candidate string-fret pair are excluded from the evaluation process, since their correct classification is essentially predetermined. The computed and estimated inharmonicity curves for three pickup-scheme configurations are displayed in Figures 3-5. The accuracy results for all configurations are presented in Table 1. Notably, the neck pickup exhibits an intriguing drop in accuracy. However, the system appears to be fairly accurate overall, with schemes *2FretA* and *3Fret* performing slightly better than *1Fret* and *2FretB*.

Table 1. Classification Accuracy Results

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

5. CONCLUSIONS

In this paper, an inharmonicity-based pipeline for the automatic detection of guitar strings from monophonic audio recordings was presented. The proposed methodology combines elements from related works, while pursuing simplicity and comprehensibility. The achieved accuracy results are deemed to be satisfactory, prompting the exploration of the system's effectiveness over more datasets, such as the *Guitarset* [20] or other [5], in future work.

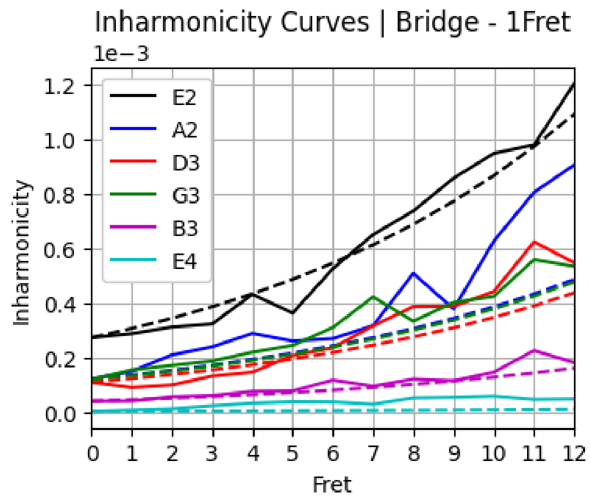

Figure 3. Computed (solid) and estimated (dashed) inharmonicity curves for the *Bridge - 1Fret* configuration

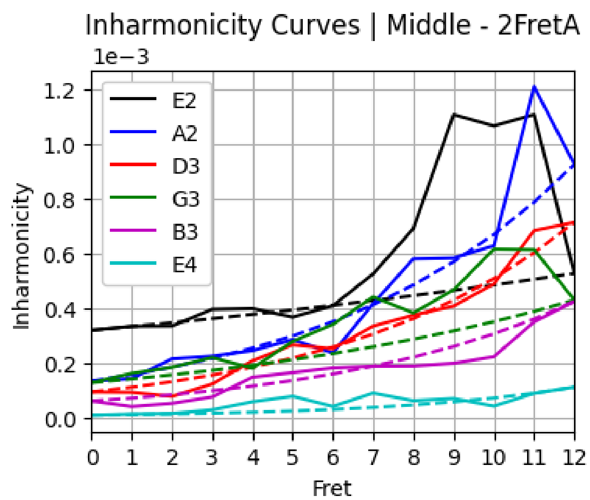

Figure 4. Computed (solid) and estimated (dashed) inharmonicity curves for the *Middle - 2FretA* configuration

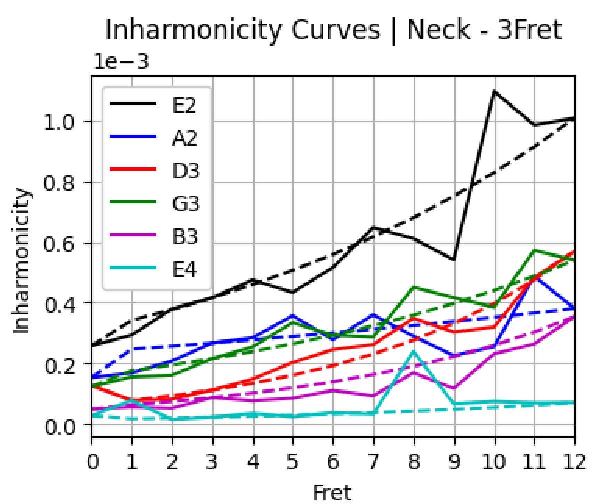

Figure 5. Computed (solid) and estimated (dashed) inharmonicity curves for the *Neck - 3Fret* configuration

6. 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 ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2022. doi: 10.1109/ICASSP43922.2022.9747169.

[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, AES 2018, 2018.

[3] J. Abeßer, “Automatic string detection for bass guitar and electric guitar,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2013. doi: 10.1007/978-3-642-41248-6_18.

[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 Proceedings of the International Conference on Digital Audio Effects, DAFx, 2014.

[5] J. M. Hjerrild and M. Grasboll Christensen, “Estimation of Guitar String, Fret and Plucking Position Using Parametric Pitch Estimation,” in ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2019. doi: 10.1109/ICASSP.2019.8683408.

[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, vol. 20, no. 6, 2012, doi: 10.1109/TASL.2012.2191281.

[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. doi: 10.1109/WASPAA.2019.8937157.

[8] C. Dittmar, A. Mannchen, and J. Abeber, “Real-time guitar string detection for music education software,” in International Workshop on Image Analysis for Multimedia Interactive Services, 2013. doi: 10.1109/WIAMIS.2013.6616120.

[9] T. Geib, M. Schmitt, and B. Schuller, “Automatic guitar string detection by string-inverse frequency estimation,” in Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI), 2017. doi: 10.18420/in2017_08.

[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, vol. 20, no. 3, 2012, doi: 10.1109/TASL.2011.2174227.

[11] D. R. Tuohy and W. D. Potter, “A genetic algorithm for the automatic generation of playable guitar tablature,” in International Computer Music Conference, ICMC 2005, 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 ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2013. doi: 10.1109/ICASSP.2013.6637636.

[13] A. Wiggins and Y. Kim, “Guitar tablature estimation with a convolutional neural network,” in Proceedings of the 20th International Society for Music Information Retrieval Conference, ISMIR 2019, 2019.

[14] M. Paleari, B. Huet, A. Schutz, and D. Slock, “A multimodal approach to music transcription,” in Proceedings - International Conference on Image Processing, ICIP, 2008. doi: 10.1109/ICIP.2008.4711699.

[15] A. Hrybyk and Y. Kim, “Combined audio and video analysis for guitar chord identification,” in Proceedings of the 11th International Society for Music Information Retrieval Conference, ISMIR 2010, 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 (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2007. doi: 10.1007/978-3-540-77129-6_54.

[17] A. M. Burns and M. M. Wanderley, “Visual methods for the retrieval of guitarist fingering,” in Proceedings of the 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, vol. 34, no. 6, 1962, doi: 10.1121/1.1918192.

[19] H. Järveläinen, T. Verma, and V. Välimäki, “The effect of inharmonicity on pitch in string instrument sounds,” in International Computer Music Conference, ICMC Proceedings, 2000.

[20] Q. Xi, R. M. Bittner, J. Pauwels, X. Ye, and J. P. Bello, “Guitarset: A dataset for guitar transcription,” in Proceedings of the 19th International Society for Music Information Retrieval Conference, ISMIR 2018, International Society for Music Information Retrieval, 2018, pp. 453–460.