Polyphonic Detection System for Electric Guitar

Junha Na
MARTE Lab.
Dongguk University
Seoul, Korea

Baron Kim
MARTE Lab.
Dongguk University
Seoul, Korea

Abstract— **This research proposes the real-time chord detection based on frequency analysis of guitar. Harmonics patterns of the specific notes are analyzed through Max/MSP which is the visual programming language for music and multimedia developed by company 'Cycling '74'. Unnecessary frequency energies for chord detection are subtracted and amplify tracked fundamental frequencies to overcome monophonic detection system, and proposing polyphonic detection system for guitar.**

Keywords-guitar chord detection, polyphonic detection, audio analysis, pitch tracking

I. INTRODUCTION

The pitch detection is a technology of detecting fundamental frequency from an audio signal. In general, the frequency values of played notes are tracked and converted into MIDI values so that the values are controlled by digital/acoustic instrument through audio interfaces. For example, the Max/MSP object called fiddle~, developed by Miller Puckette, is based on FFT which tracks the pitch of played instrument in real-time and converting them into MIDI data. However, tracking single fundamental frequency could be suitable for monophonic instruments such as flute, but not so effective for chord detection on polyphonic instruments like guitar and piano. Also, the commercial chord detectors are based on fundamental frequency method, and not concerned for frequency setting between unique partials of each instruments have, so the result could be unstable detection. These problems could cause the limitation of live performance of pitch detecting on polyphonic instrument and relative computer system such as score following. Score following is the computer process that automatically listens to the live performance and tracks the score in real time which is produced in IRCAM. Different from the performance with recorded backing track, in score following system, notes that performer plays are a trigger of audio actions. It gives 'tempo rubato (free tempo)' to musician so the musician can become a performer and conductor of music at the same time. In other word, performer controls the backing track and This is opposite principle of general musical performance. However, this kind of systems could be quite attractive elements to manufacture computer music performance. Thus, the ultimate purpose of guitar polyphonic detection system is to help guitarists who wants to make interactive music that includes chord playing. Therefore, this paper proposes and describes the polyphonic chord detection system which is specialized for electric guitar through characteristic frequency analysis and effective mapping.

This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Human Resources Research & Development Program 2016

II. REAL-TIME MULTI PITCH TRACKING

A. Feature of Partials

First of all, frequency domain spectrum analyzation has been done to find out partial character of guitar, Gibson Les Paul Standard in this case, and FabFilter Pro-Q was in use to analyze spectrum. [Figure 1] is the frequency spectrum distribution of guitar C3 note.

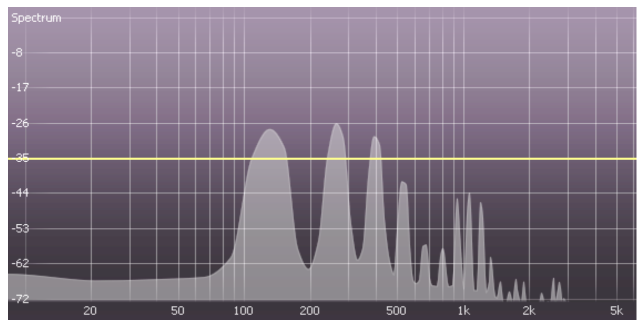

Figure 1. Frequency spectrum of C3

The spectrum shows that there are several frequency groups having obviously strong energy apart from (in addition to) 139.81Hz, the fundamental frequency of C3. In general, the strongest energy can be found on the fundamental frequency, an octave higher and fifth from the higher octave are similar energy of fundamental frequency has. That is to say the strongest partials in guitar are as the following.

- fundamental frequency (P1)
- 1 octave higher frequency (P1 * 2 = P2)
- fifth from 1 octave higher (P2 * 1.49847 = P3)

These partial pattern maintains in overall electric guitar notes. Therefore, it is necessary to differentiate between these 3 main partials and the other unnecessary noisy partials.

B. Classification of Partials

As the main concept of multi-pitch detection system is based on guitar frequency partials, P1, P2, P3 of each notes had to be extracted. In this papers(research), the objects which called reson~ and mtof of Max/MSP were used for resonance bandpass filter and converting MIDI note value to frequency value.

The MIDI value of C3(48) is equal to 130.81Hz in frequency value. The function of mtof object is to calculate this scaling process. With the equations of P2 and P3, 261.63Hz and 392.04Hz are calculated and set on reson~, the bandpass filter. As the result, the frequency energy of C3 played through the detection system is getting stronger, and getting weaker when the other note is played. [Figure 2] is a pitch detection patch that extracts 3 strongest partials of C3 note.

The object that looks microphone is ezadc~ object which sends input audio signal to other object. In this case, pure audio signal goes to three bandpass filters then it changes to three processed signals that sound quite close to the sine wave. If any audio signal gets through these filters, they always try to pick assigned signals. It means, even if A3 note has played in the C3 detection patch, filters would leach C3's P1, P2 and P3.

In [Figure 4], it shows an energy distribution in A3 note when played through C3 detection patch. The reason of C3 partials could appear in A2 is because of the guitar noise. As the acoustic instruments has their own unique noise sound, every note of guitar must be containing unassigned partials. However, these noise sounds have much smaller energy than fundamental note.

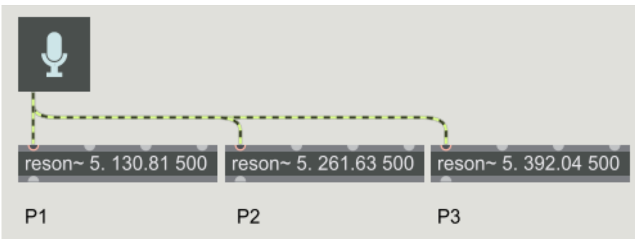

Figure 2. C3 detection patch

By comparison with [Figure 3] and [Figure 4], there is obvious energy difference between C3 and A2 when played through C3 detection system. The partial frequencies react with the setup of C3 and the energy builds up over the line, while the other note's energy level is low. So, the key architecture of this polyphonic detection system detects chords by separating these three partial energies have reached at certain level or not.

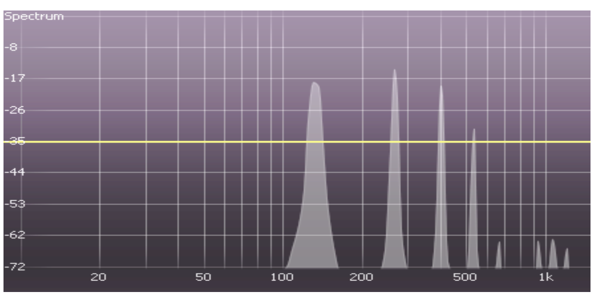

Figure 3. Energy distribution in C3 when played through C3 detection system

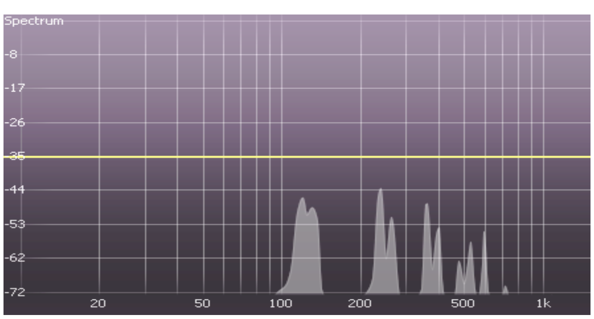

Figure 4. Energy distribution in A3 when played through C3 detection system

C. *Logarithmic Amplitude Mapping*

The detection system is based on separating amplitude value of input (income) frequency. meter~ object could be used to digitize amplitude value, but audio signal from guitar has many irregularities. The signal can be varying with circumstance such as choice of guitar, audio interface and plug. The noise from guitar itself has consistent amplitude level as well. Especially, loudness depends on how hard the performer play is not reliable. For example, digitized amplitude values through meter~ object is around 0.15~0.2 when performer played as moderate, 0.08~0.1 when played as pianissimo. The test result is shows no major difference between two playing style. And while the played notes are being sustained, amplitude value through meter~ object is similar to the value of when it is silence.

Mapping the pure amplitude value while these phenomenons still happening, accurate detection is can never be achieved by several variables mentioned above. This is why the detection system should be made in logarithmic scale, similar to how human detects difference of sound level. Therefore, this research is using 3dB per step standard.

As it shows in [Figure 5], by applying atodb object on meter~, amplitude data converted into dB value. But dB value in silent state is -inf, which not useful to calculate. live.gain~ is connected next to atodb to receive and convert change values in range of -70dB to 0dB. Then, /3 divides dB value from third outlet of live.-gain~ by 3 in order to distinguish by 3dB. Lastly, round object rounds off the outcome to integer to simplify further calculation.

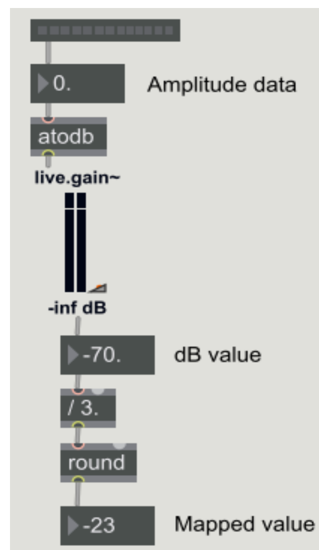

Figure 5. 3dB step mapping

Due to this process, computer can detect the change of audio amplitude like human ear. It means, data value's change is not subtle change like meter~ anymore but it is quite intuitive change.

D. *Simultaneity of Partial Energy*

Making of proper pitch detection system for guitar is now possible, as the setup of effective amplitude value mapping is finished. The biggest problem in polyphonic detection system for guitar is the P3, the third partial. P3 is fifth from the 1 octave higher, while other partials are related as octave. 392.04Hz for example, the third partial frequency of C3 can be misdetected as G4, due to high amplitude value. To avoid this detection error, the detection system should have activated only when three audio signals satisfy certain amount of amplitude value at same time.

[Figure 6] shows the solution to prevent misdetection. The three integers on the top of each patcher is calculated amplitude value of partials from [Figure 5]. bondo object synchronizes and outputs a set of inputs when any input is received. Each outputs from bondo are sent to >-10 to be checked. The checked value will be converted to 1 if the outputs are over -10, 0 if it is lower than -10. The outputs are packed and added up through pack and zl.sum, and the range of the result will be 0 to 3 in integer. 3 acts when amplitude values of all three partials satisfy the condition. As the result, the detection system will be activated only when the data through the sel object is 1. The P1, P2, P3 in the patcher on the middle of the [Figure 6] are -7, -4, -3, satisfy the condition to make 3 to toggle on the detection system.

Depending on the environment such as audio interface, electricity power, performing place, the power of audio signal could be quite different. Therefore, the standard number '-10' should be set to the right standard number according to the situation.

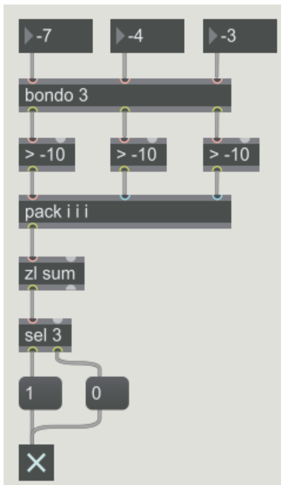

Figure 6. Simultaneity of partial energy

E. *Short Latency Filtering for Attack*

The filtering of basic partial frequency is now available through the method mentioned above. However, picking and stroking guitar causes short bursts of sounds in wide frequency range, resulting unnecessary peak by resonance between noises. Basic pitch detection is still possible with this factor, but with possibility of many unwanted random pitches to be detected. [Figure 7] shows the distribution of frequency when the 'attack' moment of C3 note. As it shows, it contains P1, P2, P3 and powerful wide range of noises at the same time.

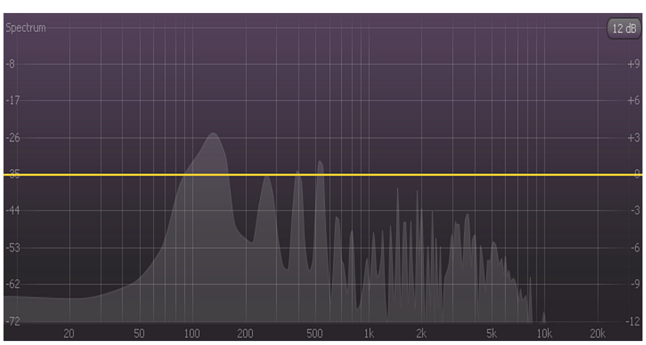

Figure 7. Attack moment of C3 note

Because of these noises, it is risky to detect when the guitar is strummed. Thus, it is necessary to capture the moment when the noise sound gets weaker.

In [Figure 8], it shows the spectrum aspect after 50ms of 'attack' moment. In this case, it could be called as 'sustain' moment. By comparison with attack moment and sustain moment, it is possible to see the noisy partial has been reduced rapidly but the main partials which are P1, P2 and P3 remains stably.

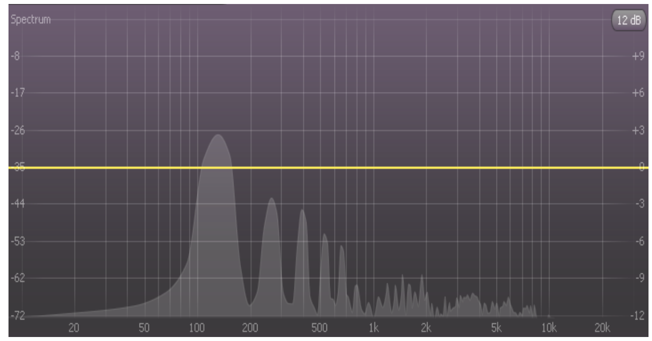

Figure 8. 50ms later of attack moment

To reduce the risk of detection, using a sustain sound is better idea rather than noisy attack sound. Therefore, detection system needs to have 50ms latency. The patcher on [Figure 9] shows the short latency noise filtering design. This timer patcher is added to the patcher from [Figure 6] to send activation signal after 50 milliseconds from received, in order to avoid detection on the plosive sounds occurred by picking and stroking. the length of the timer is carefully tested to minimize signal error from noise of picking and stroking. The role of metro object is similar to metronome and it works when the toggle is on. This time the cycle of metro is 1ms. Counter object keeps track the received message. So every 1ms, inlet number under the counter would increase. When the inlet number reaches to upper 50, if object will react and send 1 to last inlet.

A single pitch detection patch is through several processes of filtering above. By making this pitch detection patcher from E2, the open note of sixth string to E5, the twelfth fret of first string of guitar to achieve multiple pitch detection.

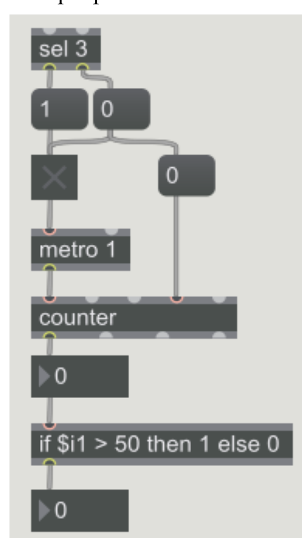

Figure 9. Short latency noise filtering

III. CHROMAGRAM

Final task after enabling polyphonic detection is making chromagram to visualize detected chord. Chromagram is a graph that visualizes 12 chromatic notes from C to B in real-time by input energy value changes. [Table 1] shows the number of notes to E5 on first string of guitar. Chromagram does not require separation of octaves as the graph consists of only the 12 chromatic scale notes. Therefore, octave for each notes were combined into a single patch to move the graph responsible for the note. The highest note on the detection system is E5 and is because the chord plays of using higher than E5 is unusual.

[Table 1] is the full list of applied notes on the detection system in this research. Digitized value of detected note will be represented as 1, or else will be 0. The chromagram is made in piano style with use of multislider object, and visualize detected note as blue key on the piano shape. For example, the chord components of Cmaj7 is [C, E, G, B] is [1,0,0,0,1,0,0,1,0,0,0,1] in the chromagram.

TABLE I. NUMBER OF NOTES THAT CONSIST OF CHROMAGRAM

Note	Number of notes	
C	3	(C3, C4, C5)
C#	2	(C#3, C#4)
D	3	(D3, D4, D5)
D#	2	(D#3, D#4, D#5)
E	3	(E2, E3, E4, E5)
F	3	(F2, F3, F4)
F#	3	(F#2, F#3, F#4)
G	3	(G2, G3, G4)
G#	3	(G#2, G#3, G#4)
A	3	(A2, A3, A4)
A#	3	(A#2, A#3, A#4)
B	3	(B2, B3, B4)

IV. DATA SIGNAL FLOW

The methodology of this polyphonic detection system is based on the energy division of audio signal. First, the original signal goes through the resonance filter then it becomes to filtered signal which is similar to sine wave. After that, the amplitude values divided though the 3dB step mapping technique which is the logarithmic mapping system. Especially, the main point of this 3dB mapping system is to make a 'humanized computer ear'. Next, there is a short latency filter system and computer distinguishes assigned signal and unassigned noises in 50ms. Therefore, short latency system is able to pick only sustained three strongest partials and abandon unnecessary noises. These process is basic form of multiple pitch detection in this paper. When the guitarist plays the chord such as C maj chord, this mechanism applies to each configure notes which are C, E, G. Finally, these detected notes appear in choromagram visually.

[Figure 10] shows the overall flow-chart of the polyphonic detection system.

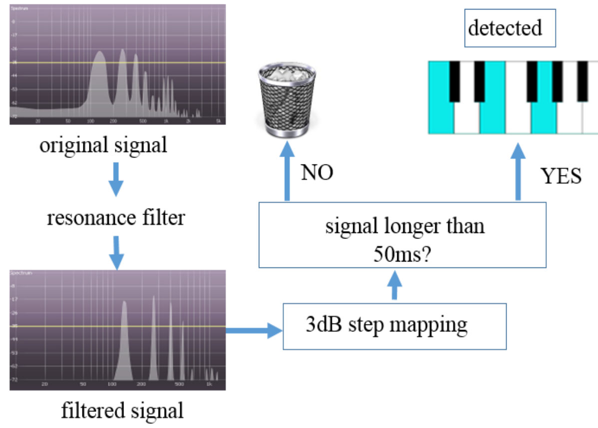

Figure 10. Polyphonic detection flow-chart

V. CONCLUSION

This paper is about chord detection system for polyphonic instrument which overcomes the limit of existing monophonic pitch detection system based on FFT analysis. The range of detection is from E2 to E5, with high accuracy for every chord played.

However, the further research is required in order to solve following problems.

First, the energy of the pitch is varied by strings as the nature of guitar. When it comes to A2 note as open string of fifth string and fifth fret of sixth string, the patterns of noise are different, especially around 1000Hz and 2000Hz. It could be a cause of misdetection in chromagram so it is necessary to unify the energy difference between two strings. According to the guitar type, it must be contained with different partial patterns and noise pattern so the generalization of detection system would be critical issue in the future.

Second problem to be solved is about building computer electronic action system through detection. For example, if there are two interactive preset commands set as 'proceed action A when Cmaj7(C, E, G, B) is played' and 'proceed action B when Emin(E, G, B) is played' having overlap notes of E, G, B. Playing Cmaj chord will proceed both action A and action B because of these duplicated notes. Additional mapping is in need to correct detection by separating duplicated chord composition. This issue would be very import when the performer uses an interactive live performance system like score following. If the computer misunderstands the chord Cmaj7 as Emin, unintended electronic action could be happened and it would be a trouble in the live performance progression.

Third, when it comes to 50ms latency of detection system, this latency is to filter noises when the player strums the guitar, but the detection delay would not be ignored in performance. For example, if the guitarist plays 16 beats phrase in fast music like 160BPM, the interval of each note would be 93.75ms so it would be fine to make a collaboration with the computer action. However, it he plays 32beats phrase in 160BPM, the interval becomes

46.8ms then computer cannot react properly. Making electronic actions in short time like 50ms is not quite common however it would be great to make a faster mechanism. Therefore, future research aims reduce the latency under 30ms to minimize the delay time for better performance.

The performance of the chord detection system will be significantly improved by solving these problems mentioned above, and will be applied on wide range of creative multimedia performance and interactive computer music for the various polyphonic instruments.

REFERENCES

[1] T. Fujishima.1999. *Real-time chord recognition of musical sound: A system using common Lisp music.* in Proc ICMC. 464–467.

[2] Stark, A.M. and Plumbley, M.D.2009. *Real-Time Chord Recognition for Live Performance.* In Proc ICMC.

[3] Laurent Oudre., Yves Grenier, Cédric Févotte.2011. C*hord Recognition by Fitting Rescaled Chroma Vectors to Chord Templates.* IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, vol. 19, no. 7, 2222-2233

[4] Arshia Cont.2006. *Realtime multiple pitch observation using sparse non-negative constraints.* in 7th International Symposium on Music Information Retrieval (ISMIR)

[5] Laurel S. Purdue., Christopher Harte & Andrew P. McPherson. 2015. *A Low-Cost Real-Time Tracking System for Violin.* Journal of New Music Research

[6] B. Pardo and W. Birmingham.2002. *Algorithms for chordal analysis.* Comput. Music J., vol. 26, no. 2, 27–49

[7] C. Harte and M. Sandler.2005 *Automatic chord identification using a quantised chromagram,* in Proc. of the Audio Engineering Society, Barcelona, Spain.

[8] Nicolas Boulanger-Lewandowski., Yoshua Bengio and Pascal Vincent.2013. *Audio chord recognition with recurrent neural networks.* In Proceedings of the International Conference on Music Information Retrieval (ISMIR), 335–340.

[9] M. Mauch.2010. *Automatic chord transcription from audio using computational models of musical context*. PhD thesis, University of London.

[10] McLeod, P. H. & Wyvill, G.2005. *A Smarter Way To Find Pitch.* Proc ICMC. Barcelona, Spain.

[11] Philip McLeod.2008. *Fast, Accurate Pitch Detection Tools for Music Analysis.* PhD thesis, the University of Otago, Dunedin, New Zealand.

[12] A. Pertusa and J. M. Inesta.2008. *Multiple fundamental frequency estimation using Gaussian smoothness.* IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 105-108.

[13] Patricio de la Cuadra., Aaron Master., Craig Sapp. 2001. *Efficient Pitch Detection Techniques for Interactive Music.*Center for Computer Research in Music and Acoustics(CCRMA), Stanford University

[14] N Orio. 2001. *Score following using spectral analysis and hidden markov models*. In Proc ICMC,.

[15] M Puckette. 1995. *Score following using the sung voice*. In Proc ICMC.

[16] M. Puckette. 1990. *Explode: a user interface for sequencing and score following.* In Proc ICMC. 259–261.

[17] Arshia Cont. 2008. *Antescofo: Anticipatory Synchronization and Control of Interactive Parameters in Computer Music*. In Proc ICMC. 33–40.