Proceedings of 20th International Congress on Acoustics, ICA 2010

23-27 August 2010, Sydney, Australia

Estimating a musical note of solo performance on the electric bass guitar by using wavelet analysis

Shunsuke Miyake (1) and Masanobu Miura (2)

(1) Graduate School of Science and Technology, Ryukoku University, Japan
(2) Department of Science and Technology, Ryukoku University, Japan

PACS: 43.75.-z Music and musical instruments

Musical scores are used by many musicians because they frequently use them to practice. However, not many scores have been transcribed, so desired scores are often not found. Therefore, those who need the score are required to listen to the performance and to transcribe it into a musical notation. When transcribing a musical note, several musical abilities such as absolute pitch and relative pitch are required, causing difficulty for people with little musical experience. From this background, studies on "automatic transcription" aimed at automatically transcribing the acoustic waveform of the performance has been developed. Therefore, this study aims to establish the automatic transcription system for musical note of solo performances onan electric bass guitar by using the wavelet analysis, a frequency analysis method with the high frequency resolution and time resolution. To achieve this, a method for estimating musical notes was proposed. Here, a musical note is a structure composed of an onset time, offset time, and fundamental frequency. To evaluate the proposed method, the sounds performed in various ways (fingering, picking, slapping, and muting) were analysed, and the results were compared with data taken by hand labeling. When the F-measure was used as a performance indicator, the average F-measures were 0.80 for the onset time, 0.77 for the offset time, and 0.71 for the fundamental frequency.

1. INTRODUCTION

Musical scores are used as a record of music by a lot of musicians. For instance, they are used to practice and to play musical instruments in popular music such as jazz and rock. However, when people practice songs without musical scores, they need to detect performance aspects such as onset time, offset time, fundamental frequency, and tempo from the recorded performance. Therefore, when the tempo of the performance is comparatively fast and/or multiple sounds are performed at the same time, it is difficult for beginners to detect the above aspects.

Therefore, past studies have developed "automatic transcription" aimed at automatically transcribing the acoustic waveform of the performance [1-2]. To automatically transcribe music, it is necessary to estimate the musical note, tempo, and beat time from acoustic signals of the performance. Thus, this study describes a way to estimate musical notes. In this study, the structure of the single sound is composed of three pieces of information: onset time, offset time, and fundamental frequency. To estimate a musical note means to estimate these three things. Musical note estimation is necessary for automatic transcription, because the fundamental frequency and the length of the musical note in the musical score can be obtained by estimating the musical note.

Although many studies have focused on the sound of music from the mid-tone to high-tone ranges for instruments such as the piano and guitar, few studies have focused on low-tone range sounds such as those of an electric bass guitar. This is because the frequency analysis with a high frequency resolution and time resolution for transcribing onto a musical note are required, and it has been difficult to analyse a performed sound at a low-tone range in Fourier analysis used in previous studies.

Therefore, the target of this study is to establish an automatic transcription system for musical notes of solo performances on the electric bass guitar by using the wavelet analysis, a frequency analysis method with the high frequency resolution and time resolution. In this study, as a preparation step for the automatic transcription, we propose the estimation method of musical note.

2. WAVELET ANALYSIS

2.1 The basis of wavelet analysis

Wavelet analysis is a technique for analysing the frequency by convolving the mother wavelet with the acoustic waveform. The short shape of waves called mother wavelets can be expanded, reduced, and/or moved parallel to the time axis. At this time, the coefficient that decides the length of the mother wavelet to be expanded or reduced is called a scale. The mother wavelet with value of a small scale corresponds to the high frequency, while the one with the value of a large scale corresponds to the low frequency. Thus, the value of scale corresponds to the value of frequency [3].

Equation 1 indicates the method for convolving the mother wavelet with the acoustic waveform. In this equation, f is the acoustic waveform of a solo performance played on an electric bass guitar, and $\phi(t)$ is the mother wavelet. In addition, a and b are positive real numbers: a is value of scale, b is centre time of the Gauss window in the mother wavelet, and t is the time of the signal input.

$$WT(b,a) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} f(t) \overline{\psi(\frac{t-b}{a})} dt \qquad (1)$$

In this study, the mother wavelet of Gabor was used on the wavelet analysis because it can take into account the phase gap of the shape of waves [4]. Figure 1 shows the mother wavelet of Gabor.

2.2 Setting wavelet analysis in this study

In the wavelet analysis, the time shift must be set. When the frequency is analysed by using the wavelet analysis, the value of scale and information of time can be obtained from the acoustic waveform. However, to estimate a musical note, information of frequency is required. Therefore, the scale must be converted into information of frequency.

i) Setting shift time

The shift time must be set when analysing the frequency. Therefore, to obtain an appropriate shift time, a sufficient width was needed to be able to distinguish the two consecutive sounds played on an electric bass guitar. Then, the duration with the shortest two consecutive sounds was examined by observing an actual performance. In the experiment, width was investigated with nine bass players with 2-25 years of experience. Playing style was not specified for players, because favourite style depends on players. They were asked to answer the number of maximum single sounds to be performed per second with a metronome of 60 bpm. In the results, the average number of maximum single sounds that could be performed per second was 12 notes. The minimum duration of a single sound was 83 ms. In other words, the shift time should be less than 83 ms when analysing the frequency of a general solo performance.

Moreover, to detect a single sound for 83 ms precisely, the mother wavelet has to be shifted less than 83 ms, and the frequency must be analysed in detail. However, the duration of the single sound in the bass performance is almost 83 ms. Therefore, the performed sounds do not need to be analysed in detail by shifting the mother wavelet about 1 ms. Thus, in this study, to keep the mother wavelet shifts within the range of less than 83 ms and avoid a short shift, the time shift was set at 14 ms.

ii) Correspondence of frequency and scale

Although the value of scale and information of time are obtained in the wavelet analysis, the value of scale must be converted into information of frequency. In this study, the correspondence table of frequency information and the scale value were made by using each pure tone of the equitempered scale, and the information of frequency was obtained from the value of scale using this correspondence table. Figure 2 shows the correspondence of the value of scale and that of frequency.

The correspondence table was made as follows. First, the value of the mother wavelet scale was expanded from 0.0 to 500.0 at intervals of 0.1. Next, the expanded mother wavelet and the acoustic waveform in each pure tone tuned by equal temperament were convolved. Thus, the value of the scale with the highest value was decided to be a value of the scale on the pure tone. In this study, the estimated musical note is a performed sound in the range from 29.1 to 523.2 Hz, 51 sounds in total. Thus, the correspondence table was made from 51 values of scale corresponding to each pure tone and piece of information of frequency.

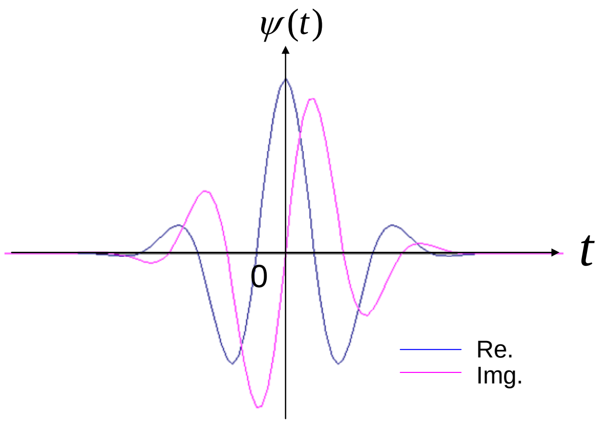

Figure 1. Mother wavelet of Gabor

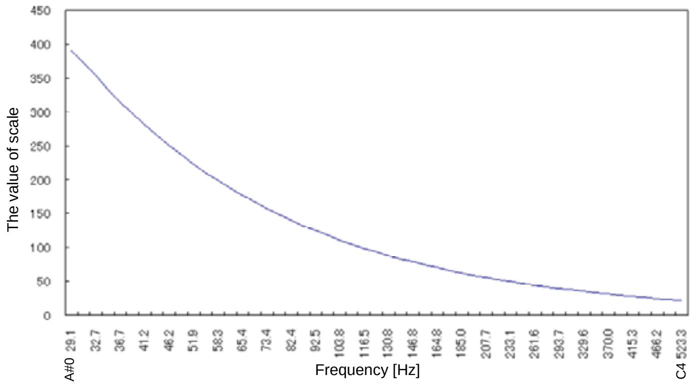

Figure 2. Correspondence of frequency and value of scale

3 PROPOSED METHOD OF MUSICAL NOTE ESTIMATION

3.1 Intended performance

In this study, an electric bass guitar was played as described in chapter 1. The solo performance played on the electric bass guitar is a monophonic performance. Therefore, in the proposed method, all performed sounds have to be completely muffled before the next sound starts. Thus, the performances in which multiple sounds are played at the same time are not targeted. In addition, the intended sound is performed in various ways such as fingering, picking, slapping (including thumbing and pop slapping), and muting.

3.2 Feature and frequency characteristics of sounds in different playing styles

Figures 3-7 show the frequency characteristic when the open string of the second string (73.4 Hz) was played with each playing style. In each figure, a horizontal axis shows the time, vertical axis the value of frequency, and magnitude of the sound pressure level. In Figs.3, 5, and 7, when the sound was played, the sound pressure level was amplified in various bandwidths. The sound pressure level of F0 was, however, the highest of all bandwidths, and F0 declined smoothly. In Fig.4, when the sound was played by picking, the sound pressure level of the harmonic tone was amplified more than F0, and the harmonic tone declined quickly and F0 declined smoothly. In Fig.6, when the sound is played, the sound pressure level was amplified in various bandwidths, and F0 was steady. Thus, each playing style has different features.

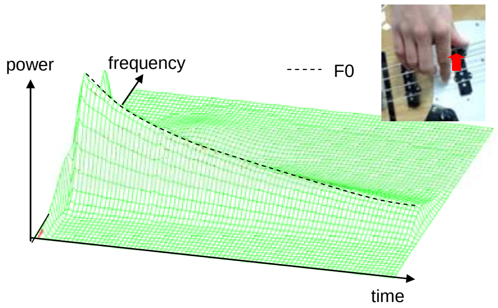

Figure 3. Sound of "fingering"

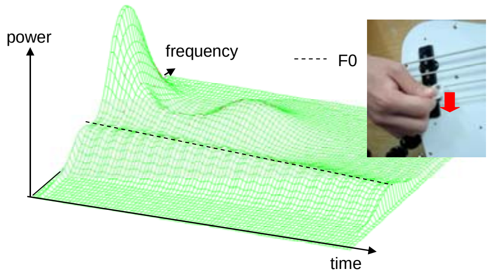

Figure 4. Sound of "picking"

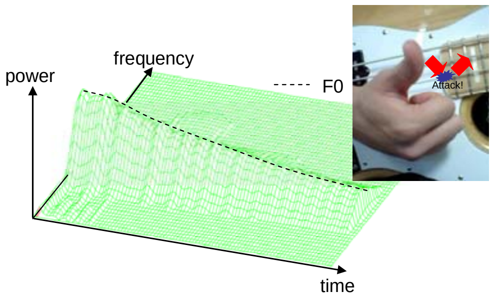

Figure 5. Sound of thumb slapping"

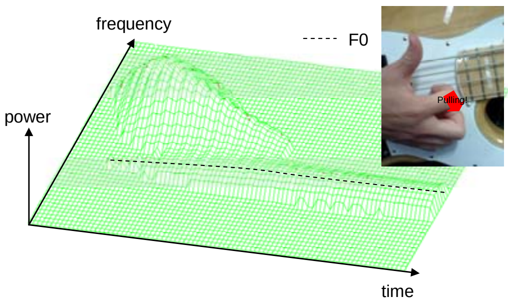

Figure 6. Sound of "pop slapping"

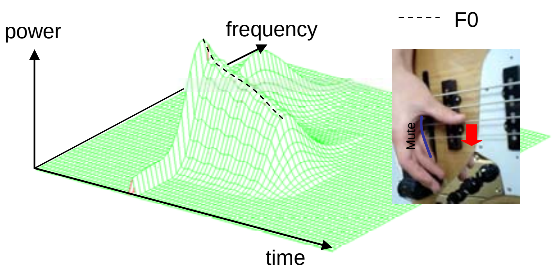

Figure 7. Sound of "Muting"

3.3 Feature of performed sound in the acoustic waveform and result of wavelet analysis

Chapter 3.2 described how the sound pressure level increases rapidly at the onset time of the sound of the electric bass guitar. Moreover, the sound pressure level decreases rapidly at the offset time. Here, the term "offset time" is the time that the sound pressure level decreases and people can no longer hear the sound. Furthermore, the rising sound pressure level in various bandwidths at the onset time of the sound confirms that the sound pressure level of the performed sound rises continuously during a certain time. Thus, in this study, the phenomenon of the sound pressure level rising during the constant time is called "continuation phenomenon of the performed sound" in areas of short time frequency. In addition, the sound pressure level is thought to rapidly decrease at almost the same time as the continuation phenomenon of the performed sound finished.

3.4 Problem of noise generated by electromagnetic induction

When a monophonic performance on an electric bass guitar is recorded, noise is generated. This noise is the tiny power generated by the electromagnetic induction of the cable that connects the electrical bass guitar with the recording device. Tiny power is generated by the influence of the noise at F0 of the sound and by the frequencies other than the harmonic tone element. The noise is sounds other than those the bassist plays. Thus, this noise should be removed to catch the sound played, for example, at the result of the wavelet analysis.

3.5 Principle of proposed musical note estimation

Chapter 3.3 described how the sound pressure level of the acoustic waveform of performed sound is relatively high at the onset time, enabling the continuation phenomenon of the performed sound to be observed. Moreover, the continuation phenomenon of the performed sound ended at the offset time in the results of the wavelet analysis. Chapter 3.4 also detailed why the noise should be removed because the noise of tiny power generated by the electromagnetic induction was included in the performance. When this noise is estimated, the musical note becomes easier to estimate.

Thus, at first, the noise is removed in the proposed method. Next, onset time is when the sound pressure level in the acoustic waveform rapidly increases just before the continuation phenomenon of the performed sound, so it is necessary to select the peak of the onset time from a lot of peaks in an amplitude envelope. Then the correlation coefficient for spectrum between the time R_t and R_{t-1} is calculated, and the time function of the correlation coefficient an be obtained. After that, the peaks at high correlation coefficient times are ignored when estimating the onset time using the amplitude envelope, because in such ranges spectrum at R_t is similar to that at R_{t-1}. On the other hand, offset time is when the continuation phenomenon of performed sound ends. Moreover, the fundamental frequency is estimated on the basis of frequency elements between the onset and offset times.

3.6 Flow of proposed musical note estimation

Figure 8 shows the flow of the musical note estimation. Musical note estimation here means to estimate the onset time, the offset time, and the fundamental frequency as described in chapter 1. First of all, the input acoustic waveform of the solo performance of an electric bass guitar was analysed by the wavelet analysis, and the time function of each frequency corresponding to a musical note was obtained (Fig.8(a)). Next, amplitude envelope was obtained from the results of the wavelet analysis (Fig.8(b)). At the same time, the time function of the correlation coefficient between frequency elements of the time R_t and those of the time R_{t-1} is obtained from the results of the wavelet analysis (Fig.8 (c)). The peaks at high correlation coefficients are ignored when estimating the onset time (Fig.8 (d)). Then, on the basis of the maximum value of the sound pressure level, tiny power less than *N* dB from the maximum is ignored in the results of obtained amplitude envelope and of wavelet analysis (Fig.8(e)). Thereafter, the continuation phenomenon of the performed sound was detected (Fig.8(f)). Finally, on the basis of the amplitude envelope in Fig.8(e) and the continuation phenomenon of the performed sound in Fig.8(f), both the onset time and the offset time were estimated (Fig.8(g)), and the fundamental frequency is estimated by observing the sound pressure level of each frequency element belonging to the chroma obtained from pitch class profiles (PCP)[6] during the onset time and the offset time (Fig.8(h)). Details of each process are explained in chapters 3.3-3.11.

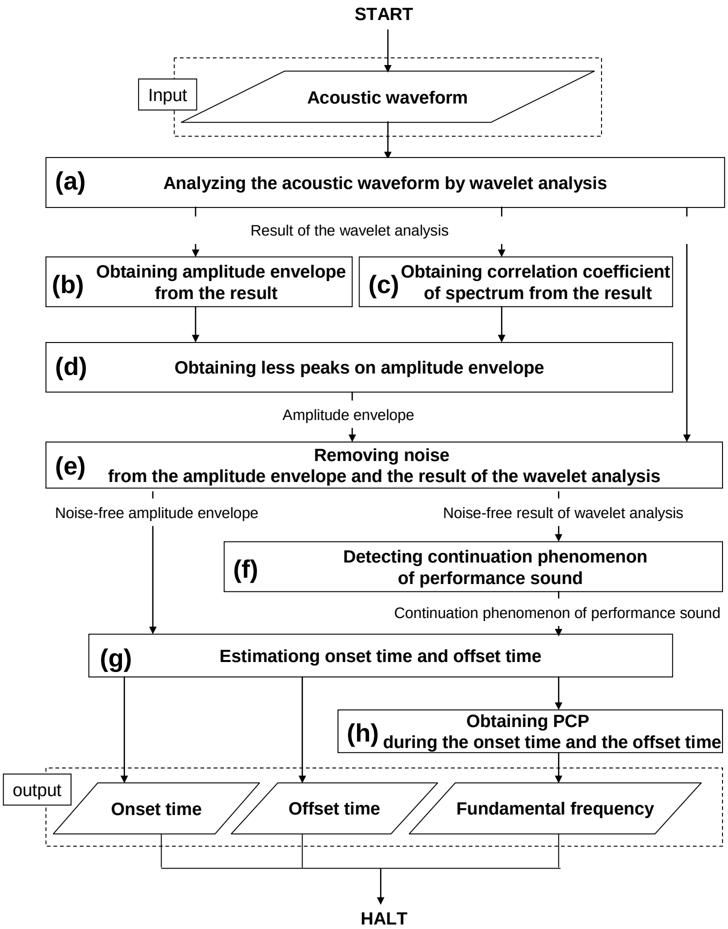

Figure 8. Flow of proposed musical note estimation

3.7 Wavelet analysis

As described in chapter 1, a frequency analysis was conducted by the wavelet analysis with the high frequency resolution and time resolution. Then, as described in chapter 2.2, the time shift for the wavelet analysis was set at 14 ms. The result of wavelet analysis is described as a time function of the power at each scale corresponding to 51 notes, representing a musical note the frequency of which is between 29.1 and 523.2 Hz.

3.8 Calculation of amplitude envelope

In this chapter, amplitude envelope is calculated from the results of the wavelet analysis to estimate onset, because frequency-based onset detection is thought to be robust in the case of dynamically changing consecutive notes. In chapter 3.3, the sound pressure level tended to increase rapidly in each bandwidth at the onset time of the sound of the electric bass guitar. Here, the onset time is estimated by observing the summation of sound pressure level for all bandwidths at each time. Then, from the results of wavelet analysis, the sound pressure levels of every bandwidths at each time are summed up to obtain the amplitude envelope. To estimate the onset time from this amplitude envelope, it is necessary to determine the onset time from a lot of peaks in the amplitude envelope. Then the correlation coefficient of frequency elements between R_t and R_{t-1} is calculated from the results of wavelet analysis, and the time function of the correlation coefficient is obtained. Next, the peaks at high correlation coefficient are ignored when estimating onsets from the amplitude envelope. This amplitude envelope was also used to estimate the onset time. The estimation of the onset time is described in chapter 3.11.

3.9 Removal of noise

Chapter 3.4 explained why the noise should be removed. To do so, two flows were considered. In the first, the noise is removed from the input acoustic waveform, and in the second, the noise is removed from the results of wavelet analysis as well as the amplitude envelope for the inputted acoustic waveform. Therefore, in this study, the noise was removed by the latter flow on the basis of a heuristic method, because the studies along the flow of these processing have produced similar results. Specifically, the noise was removed from the results of wavelet analysis as well as the amplitude envelope for the inputted acoustic waveform. First of all, tiny power of amplitude envelope less than *N* dB (in this study *N* is 20) from the maximum was removed. Next, in the results of wavelet analysis, the power when noise was removed from amplitude envelope was removed. Moreover, the power less than *N* dB from the maximum in the results of wavelet analysis was removed at each time. The noise was removed from the performed sound by sequentially performing all these three noise removal methods.

3.10 Detection of continuation phenomenon of the performed sound

This chapter describes the method for detecting the continuation phenomenon of the performed sound from the noise-free results of wavelet analysis. The performed sound of the electric bass guitar is comprised of F0 and harmonics. Therefore, it was expected that at least these frequencies could be obtained when the frequency of the sound was analysed by the wavelet analysis. Considering the frequency obtained from the performed sound, a frequency element lower than F0 is thought to be rarely generated. Therefore, according to the results of wavelet analysis, F0 becomes the peak of the frequency with the lowest frequency corresponding to the value of scale [5]. However, when the peak of the lowest frequency corresponding to the value of scale at each time is extracted, it is possible to extract not only the frequency of correct F0 but also other frequencies slightly different from the correct F0, because of the reasons for obtaining the scalar product of the frequency in the wavelet analysis. For this reason, when the continuation phenomenon of the performed sound is detected, sometimes a frequency slightly different from F0 is detected.

Thus, in this study, in the results of wavelet analysis, first of all, the peak of the lowest frequency corresponding to the value of scale is extracted at each time. Then, because a minimum duration of the single sound is 83 ms as discussed in chapter 2.2, when the peak continues during 83 ms or more, the peak is deemed to generate the continuation phenomenon of the performed sound and is then detected as the onset of the current note. In addition, to deal with the problem of extracting a slightly different frequency as a continuation phenomenon of the performed sound, once the continuation phenomenon of the performed sound was detected, the slightly different F0 was unfortunately extracted as correct F0, but only if the difference was within an acceptable range. As stated previously, the continuation phenomenon of the performed sound extracted by this technique might be slightly different from F0. The continuation phenomenon of performed sound is therefore not used to estimate the fundamental frequency, while the other technique described in chapter 3.11 is.

3.11 Estimation of musical note

As described in chapters 3.7-3.10, to estimate a musical note, the onset and offset times are estimated from the amplitude envelope and the noise-free continuation phenomenon of the performed sound. Afterwards, the fundamental frequency is estimated from PCP [6] for the interval of the onset and offset time. PCP is the intensity set of the 12 chroma. The results of wavelet analysis indicate that F0 becomes the peak of the lowest frequency corresponding to the value of scale [5]. The fundamental frequency is estimated from the lowest frequency in each frequency element of the chroma obtained by PCP. In this study, PCP is obtained at each time between the onset to offset times according to the results of wavelet analysis. PCP is obtained by summing up all sound pressure levels of frequency elements corresponding to each chroma. Therefore, PCP corresponding to 12 chroma in the equitempered scale is obtained. Next, summations of the sound pressure level of the results of wavelet analysis obtained in order to construct the PCP at each time are used to compare the sound pressure level of each chroma among them, so that the chroma with the highest summation is obtained. Then, the number of cases of each chroma evaluated as sounding is obtained from the onset to offset times, so that the chroma for the note is then estimated. Finally, the sound pressure level of each frequency elements belonging to a note is evaluated from the lowest frequency, so the strength of the chroma is determined by observing the sound pressure level. Figure 9 shows an example of estimating the onset time and the offset time. In Fig.9, the left-hand vertical axis shows the power of the amplitude envelope, and the right-hand vertical axis shows the frequency corresponding to the value of scale of the continuation phenomenon of performed sound.

The flow for the musical note estimation is as follows: at first, the peak time h_p of the amplitude envelope is extracted. When the continuation phenomenon of performed sound is detected within 83 ms, i.e., a minimum duration of the performed sound described in chapter 2.4, h_p is estimated as the onset time represented as t_i. If the continuation phenomenon of performed sound is not detected within 83 ms, h_p is not recognized as the onset time of the performed sound. Therefore, h_p is updated to h_{p+1} , and the continuation phenomenon of performed sound is again investigated. When the continuation phenomenon of performed sound was detected within 83 ms, the time when it disappeared is estimated as offset time S_i. The fundamental frequency is estimated from the PCP corresponding to the duration between the onset time t_i, the sound pressure level of each scale, and the offset time S_i. After estimating the onset time, the offset time, and the fundamental frequency, a peak of the amplitude envelope after the offset time h_p is place of h_{p+1}. Therefore, the continuation phenomenon of performed sound is again detected.

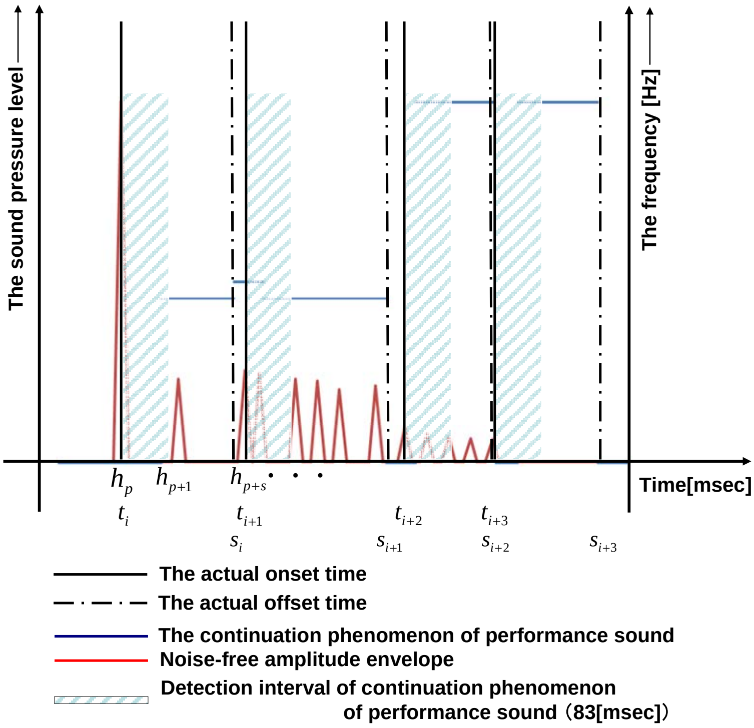

Figure 9. Example of estimating onset time, offset time, and fundamental frequency

4 EXPERIMENT

Amateur bassists played repertoires on an electric bass guitar. From acoustic waveform of performed sounds, the onset time, the offset time, and the fundamental frequency were confirmed by hearing (hand labeled data) and were also analyzing by the method for musical note estimation, and both sets of results were compared. In addition, the accuracy of the musical note estimation method is verified with F-measure. The outline is discussed below.

4.1) Bassist

The bassist had played the bass for 8 years.

4.2) Repertoire

The bassist performed the repertoire (in total 155 sounds of 20 bars) at 80,120 and 160 bpm. Repertoire from Bass Bible [7], which has a lot of frequently played scores, were played. The music used as repertoires were 266, 267, 268, 269, 270, 272, 273, 274, 275, and 278 in Bass Bible [7]. Figure 10 shows one example of theses musical scores.

4.3) Various ways of playing

The sound was performed in various ways: fingering, picking, slapping, and muting.

4.4) Recording environment

The equipment used in the experiment were a Fender Jazz Bass (four strings bass), audio interface (EDIROL UA-4FX), FUJITSU ESPRIMO (CPU: 2.8 GHz, the main memory: 2.24 GB), and metronome (KORG MA-30). When tuning the strings, the fourth, the third, the second and the first string were assigned to E0, A0, D1 and G1, respectively. To record the performed sound, the software of "XGworks ST" was used. The sampling frequency of 22050 Hz and 16 bit was adopted.

4.5) Evaluation method

The onset time, the offset time, and the fundamental frequency of the performed sound by the hand labeling were compared with the results obtained by the proposed method, and F-measure, the preciseness, the reproducibility were calculated. Equation 2 shows the formula for computing F-measure. Here $P_{recision}$ and R_{ecall} represent the preciseness and reproducibility, respectively, defined as follows.

$$\mathrm{F-measure} = \frac{2P_{recision}R_{ecall}}{P_{recision} + R_{ecall}} \tag{2}$$

The onset time, the offset time, and fundamental frequency obtained by the hand labeling data and the proposed method were compared. The number of correct answers is denoted as H, and the total of the onset time, the offset time, and fundamental frequency is denoted as G. Also, $P_{recision}$ is calculated in Eq.(3).

$$P_{recision} = \frac{H}{G} \tag{3}$$

Reproducibility is derived from the total hand labeling data and aforementioned H as in Eq.(4).

$$R_{ecall} = \frac{H}{T} \tag{4}$$

When the F-measure, the preciseness, and the reproducibility are evaluated, a time error margin allowable for the correct answer is needed. This allowable margin should be set less than 83 ms because the minimum duration of the single sound was set less than 83 ms. Thus, we consider an onset to be correctly matched if a detected onset is reported within 50 ms of the onset time. Under these limitations, the F-measure, the preciseness, and the reproducibility were calculated. When the fundamental frequency corresponding to each onset time is estimated, the fundamental frequency at which the onset time could not be estimated correctly cannot be regarded as correct. Therefore, only the fundamental frequency at which the onset time was estimated correctly was evaluated.

4.6) Procedure for the hand labeling data

Since information on the accurate onset time and offset time is not described in the recorded sound, it is necessary to manually estimate the onset time and the offset time from the acoustic waveform. In this study, the reading was done by seeing the shape of waves and by listening to the sound performed. The time when power rapidly rose was defined as the onset time. The time just before the onset time or the time when power reduced to 0 in accordance with the amplitude envelope reduction is the offset time. The fundamental frequency was read from the note that had been described in the score.

4.7) Limitation

To avoid mistakes in recording the sounds, when there were mistakes in the performance, the sound was performed again.

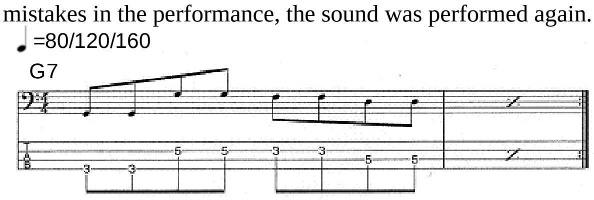

Figure 10. Example of repertoire used

5 RESULTS AND CONSIDERATIONS

Tables 1-4 detail the results of the experiment. The average F-measure for all the performances was calculated. The average F-measures were 0.80 for the onset time, 0.77 for the offset time, and 0.71 for the fundamental frequency. When onset time was not estimated, frequency element near F0 of performed sound was swinging. Therefore, the continuation phenomenon of performed sound could not be recognized.

Table 1. F-measure for musical note played by "fingering"

Fingering	The onset time	The offset time	The fundamental frequency
80bpm	**0.88**	**0.83**	**0.84**
120bpm	**0.72**	**0.65**	**0.64**
160bpm	**0.79**	**0.73**	**0.72**

Table 2. F-measure for musical note played by "picking"

Picking	The onset time	The offset time	The fundamental frequency
80bpm	**0.79**	**0.73**	**0.71**
120bpm	**0.80**	**0.78**	**0.70**
160bpm	**0.71**	**0.71**	**0.64**

Table 3. F-measure for musical note played by "slapping"

Slapping	The onset time	The offset time	The fundamental frequency
80bpm	**0.80**	**0.82**	**0.73**
120bpm	**0.87**	**0.87**	**0.69**
160bpm	**0.74**	**0.77**	**0.60**

Table 4. F-measure for musical note played by "muting"

Muting	The onset time	The offset time	The fundamental frequency
80bpm	**0.83**	**0.78**	**0.80**
120bpm	**0.83**	**0.82**	**0.75**
160bpm	**0.80**	**0.79**	**0.67**

6. CONCLUSIONS

We proposed a method to estimate musical notes, the structure of which was composed of onset time, offset time, and fundamental frequency. The acoustic waveform played on the electric bass guitar was analysed by the wavelet analysis. Firstly, the time function of the sound pressure level was obtained for each value of scale corresponding to the 51 sounds. From the results obtained from the frequency analysis, the amplitude envelope was obtained by summing up the sound pressure level of each bandwidth at each time. Pickup of the electric bass guitar and the cable to connect it to equipment contained noise caused by electromagnetic induction. The noise was removed from the results of wavelet analysis and the amplitude envelope. The continuation phenomenon of performed sound was then detected from the results of the wavelet analysis, and the onset and offset times of the musical note were estimated by using the continuation phenomenon of performed sounds and amplitude envelope. Moreover, PCP was calculated from the frequency elements during the onset time and the offset time, and the fundamental frequency was estimated by confirming the sound pressure level of each frequency element belonging to the chroma obtained from PCP. The evaluation of the proposed method was as follows. Hand labeling data of the onset time, the offset time, and the fundamental frequency were corrected. In this case, the fundamental frequency was used as it was in musical score information in the repertoire. The onset time, the offset time, and the fundamental frequency obtained by hand labeling and those by the proposed method were compared. When the allowable range of time error for the onset time, the offset time, and the fundamental frequency was limited to 50 ms, the average F-measures were 0.80 for the onset time, 0.77 for the offset time, and 0.71 for the fundamental frequency. Therefore, the results confirmed the proposed method using the wavelet analysis to be effective.

Future works are to more accurately estimate musical notes, to apply the method to a variety of performance techniques, to make the estimated musical notes into a musical score, and to estimate the tempo. The beat times are also to be studied. To improve the accuracy of the musical note estimation, the peak of the value of scale corresponding to the lowest frequency in the results of wavelet analysis at each time was used to extract the continuation phenomenon of performed sound. It was difficult to estimate the musical note because the continuation phenomenon of performed sound could not be confirmed when F0 swang. Thus, it is necessary to consider a novel extraction method with a continuation phenomenon of performed sound using information on the harmonic tone obtained from the results of wavelet analysis. The proposed method should be applicable to a variety of performance techniques, for example, the solo performance using general performance techniques such as hammering on, pulling off and sliding, and we will consider adapting the method to them. The final target is to estimate the musical score from the solo performance played on electric bass guitar, so tempo and beat time must also be estimated. The above-mentioned subjects to make the solo performance played on an electric bass guitar into a musical score are now being studied and will be discussed in the near future.

REFERENCES

1 Anssi Klapuri, “Sound Onset Detection by Applying Psychoacoustic Knowledge” Proc. ICASSP '99. pp. 3089-3092, 2003.
2 Stephen W. Hainsworth, Malcolm D. Macleod, “Automatic Bass Line Transcription from Polyphonic Music” Proc. ICMC 2001. pp.431-434, 2001.
3 H. Toda, H. Kawabata, Z. Zhang, “The latest wavelet practice course, introduction and application”, Softbank creative Co.Ltd, Tokyo, pp.63-78, 2005 (in Japanese).
4 S. Sakakibara, “Wavelet beginner's guide”, Tokyo denki university press, Tokyo, pp.16-23, 1995 (in Japanese).
5 K. Nakamura, “Mechanism of sound”, Natsumesha Co.Ltd, Tokyo, pp.86-117, 1999 (in Japanese).
6 T. Fujishima, “Realtime chord recognition of musical sound: A system using Common Lisp Music,” Proceeding of the International Computer Music Conference, Beijing, China, pp.464-467, Oct. 1999.
7 Paul Westwood, “Bass Bible”, AMA Verlag, pp.101-103, 2001.