Presented at the DLSU Research Congress 2015
De La Salle University, Manila, Philippines
March 2-4, 2015

Automatic Guitar Music Transcription

Lance Alcabasa[1*] and Nelson Marcos[2*]
1 Software Technology Department, College of Computer Studies
2 Software Technology Department, College of Computer Studies
**Corresponding Authors: lance.alcabasa@dlsu.edu.ph, nelson.marcos@dlsu.edu.ph*

Abstract: Various technologies continue to emerge to improve people's experience with music. These technologies include media programs that focus on helping people learn music. One of the most common ways how people learn is by practicing through studying and playing music transcriptions (or written music). Music transcriptions exist in many different forms, one of which is the Guitar Tablature, which are generally easy to read and understand even by beginners. However, writing tablatures can become a tedious process. There exist already technologies which aid in music transcription, however, there is a lack of focus for the guitarist community, specifically in helping generate guitar tablatures. The general objective of the paper was to analyze audio from a guitar and generate the corresponding music transcription for it, specifically guitar tablature. This covers studying key music elements and tablature structure analysis, digital signal processing and systems development, accuracy testing and user survey. The first step of the research was to analyze key music concepts and tablature structure for consistency. Written tablatures have common characteristics, however, they generally do not follow strict format and are reliant on the style of the creator, thus the need to generate a formal structure for use in the research. The next step of the research was to analyze existing techniques for digital signal processing and utilize these to create the system. Finally, the system was tested through a controlled accuracy testing and a survey among a sample group of users. Results were generally acceptable, with the sample users finding the system useful, both in helping aid guitarists create their own tablature as well as helping the guitarist community as a whole. For future work, improvements to the system can be done, particularly when working with more complex types of music and increasing accuracy overall for the system.

Key Words: guitar ; music ; transcription ; tablature ; score

1. INTRODUCTION

1.1 Research Description

Music is rapidly evolving in our world today. As such, various technologies are also emerging to help in maintaining and improving people's experience with music. Online music databases and websites, online music retrieval systems, and many others now exist and are constantly improving. One of these technologies are media programs that help people learn and create their own music transcriptions.

Music transcriptions or written music are physical representations of the music, similar to how words are translated into text. They exist in many different forms. Two of these are scores and tablatures. Scores or sheet music are generally considered as the standard way to transcribe or write music. Generally, a musical score can be applied to a wide variety of instruments. A sample of a score is shown on Fig 1. Section 1 denotes the clef. Section 2 denotes the beat and tempo for the piece of music. Lastly, Section 3 shows the notes that need to be played.

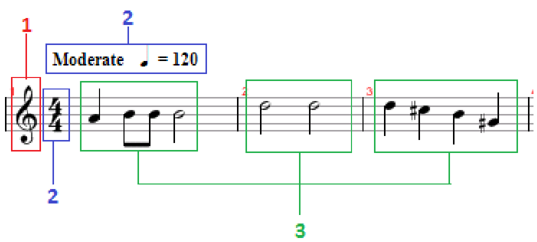

Fig 1. Sample Score.

Tablatures on the other hand are written music specifically for guitar instruments. Each form have their own advantages and disadvantages but regardless, both are created to help musicians learn, play, and share music. A sample tablature is shown in Fig 2. Section 1 denotes the tuning of the guitar. Section 2 denotes the fret numbers to be played at that point in time (in left to right order). Lastly, Section 3 indicates which string to play the frets defined on Section 2.

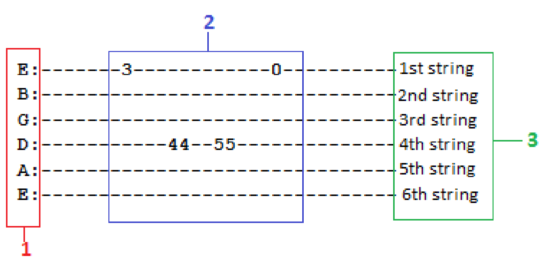

Fig 2. Tablature Sample

There are many useful technologies and systems available that focus on this particular aspect of music, such as research done by Macrae and Dixon, 2010, Kerdvibulvech and Saito, 2007-2008, Tuohy and Potter, 2006, Franklin and Chicharo, 1999, Kashino and Murase, 1998, and Chua et al., 2009.

Creating the music transcription itself can be a tedious task for people. Although there exist systems for creation of transcribed music, almost all generate musical scores, lacking a focus for the guitarist community. As such, the aim of this paper is to develop a system that will focus on the automatic generation of guitar tablature.

1.2 Research Objectives and Scope

The general objective of the research paper was to develop a system that will analyze audio from a guitar and generate the corresponding music transcription for it, specifically guitar tablature and musical score.

The research only focused on the guitar instrument, specifically the classical guitar. The research attempts to continue previous research that has been done by Chua et al., 2009 and focuses on some of the improvements stated on their paper, specifically chord detection, beat and tempo detection, as well as transcription of audio to tablature and musical score format.

The input data was gathered from both input through a microphone and samples of pre-recorded music, which were in .WAV format. These were self-generated, audio recordings from CDs, and data that were available from the internet.

Since guitar tablatures do not necessarily have a standard structure, the generated tablature's structure will be based on observations made on the common characteristics between sample tablatures.

Research into developing new algorithms for the modules were not thoroughly explored, such as for pitch detection. As such, existing algorithms were adopted for the purposes of the research.

Regarding beat and tempo detection, various limitations were taken into consideration, specifically regarding time-signatures and note durations. For any given audio data, it was assumed that tempo will be constant throughout the entire audio. Additionally, due to the almost limitless range of possible tempos, the research limited the tempo range from 40-200, which is the common range for metronomes.

Regarding chord detection, the large amount of possible chord combinations were taken into consideration. As such, the research only covered Major Chords, Minor Chords, and Seventh Chords.

The accuracy of the generated output were evaluated by comparing it with the actual supposed output. To also evaluate the effectiveness of the generated output and determine if it is presented in a familiar and presentable form, a survey was given to actual guitarists to test the system and rate if the results are accurate and can be properly read and played.

2. METHODOLOGY

The framework for the research is shown in Fig 3.

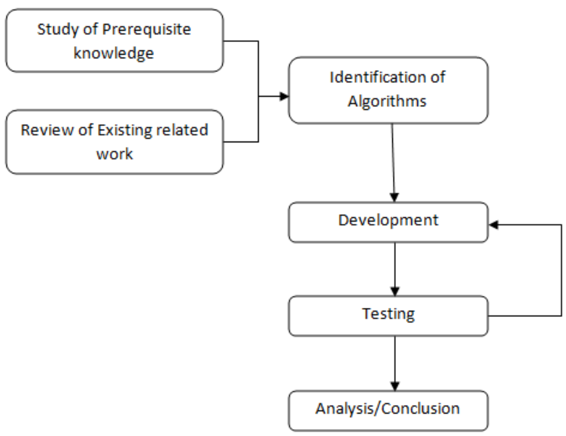

Fig 3. Methodology Diagram

2.1 Study of Prerequisite Knowledge and Review of Existing Related Work

The first step of methodology was to study the prerequisite knowledge required for the research as well as review of the existing related work at the time of the research.

Study of prerequisite knowledge covered the musical concepts that were covered for the research, technical aspects such as digital signal processing, and study of the structures of the music transcriptions that will be generated by the system, particularly for guitar tablatures as there were not yet any common standards for these, hence the need to come up with a particular standard structure that was used for the research.

Notable related work for this research are Paganini (Franklin and Chicharo, 1999) and GMT (Chua et al., 2009), to which this research is an extension to. Additionally, Kerdvibulvech and Saito (2007,2008) attempted to work with visual data. Their work involves working with a guitarist and tracking the guitarist's fingertips, with the aim of helping support beginner guitarists learning the correct finger positioning.

2.2 Identification of Algorithms and Development

The research identified different modules that were needed based on key elements of music and digital signal processing identified during information gathering. These modules are the Pre-processing Module, Pitch Detection Module, Note Segmentation Module, Chord Detection Module, Beat and Tempo Module, and Post-processing Module.

2.2.1 Pre-processing Module

The purpose of the Preprocessing Module is to prepare the .WAV file to be presented as input for digital signal processing. This is to remove irrelevant data from the source.

2.2.2 Pitch Detection Module

The purpose of the Pitch Detection Module is to extract the pitch information from the music. Pitch is defined as the fundamental frequency of a sound, which determines the 'highness' or 'lowness' of a sound. It identifies the tone of the note that was played.

2.2.3 Note Segmentation Module

The purpose of the Note Segmentation Module is to segment each individual note from each other. This is done by performing analysis on note onsets. Fig. 4 shows the characteristics of note onsets.

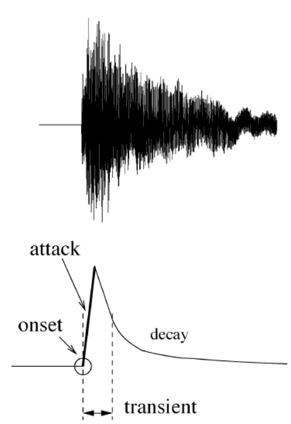

Fig 4. Note Onset (Bello et al., 2005)

2.2.4 Chord Detection Module

The Chord Detection module is used for distinguishing and labelling chords in the signal. The Chord Detection module has two functions:

1. Chord Labelling
2. Chord Distinction

Chord labelling is the aspect of chord detection module that determines which chord label or chord structure is assigned to a particular point in the signal.

Chord distinction is the aspect of the chord detection module that distinguishes whether a sound produced is a single note or a chord. Because of the lack of papers regarding this aspect, algorithms were derived based on a theory made through observation of the frequency domain between single-notes and chords.

2.2.5 Beat and Tempo Module

The beat and tempo module deals with detecting beats as well as estimating the tempo (which is measured in beats per-minute). The beat is defined as a point-in-time that occurs at a regular interval in the music. Most of the musical events (such as the beginning of a new note) fall on the beat or subdivisions of the beat (half-beat, quarter-beat, etc.).

2.2.6 Post-processing Module

The goal of the post-processing module is to gather all of the output from the three modules, pitch detection module, note segmentation module, and beat and tempo detection module, and transcribe these into tablature and sheet music format.

2.3 Testing

Testing was done on 22 audio files with varying elements. A brief overview of this include 5 songs for basic testing, 3 songs that only contain chords, 6 songs that only contain single-notes, 4 songs that contain both single-notes and chords, and lastly, 4 songs that are excerpts from CD quality files (MP3 files which were translated to .WAV files). Additionally, 6 extra audio files with varying tempos are used for further testing the beat and tempo module.

Additionally, supporting data was also gathered in the form of 4 testers of the system. Two of the testers are professionals in the field and two are casual guitar hobbyists. They were given a survey after allowing them to use the system and test its capabilities.

2.4 Analysis and Conclusion

Once the system was evaluated and tested, the results were collated. These are further discussed in Section 3 and Section 4 of this paper.

3. RESULTS AND DISCUSSION

Testing was done regarding the different types of modules namely Pitch Accuracy, Onset Detection Accuracy, Chord Distinction Accuracy, Chord Labelling Accuracy, and Tempo Accuracy.

3.1 Pitch Accuracy

The aim of this test was to determine the accuracy of pitch assigned to the notes in the audio, based on octave errors, pitch errors, and both. The testing was done only on the test data that contained single notes.

Overall, the accuracy results of 69.71% was acceptable. It was observed that the auto-correlation algorithm implemented mainly had problems dealing with the more complex audio files, such as those containing fast played notes, string echo, and those test data from the CD.

3.2 Onset Detection Accuracy

The aim of this test was to determine the accuracy regarding detected note onsets, based on actual missed onsets and ghost note onsets. Missed onsets are onsets or notes that were played in the audio but was not detected by the system. Ghost notes are onsets detected by the system but were note played in the audio. The testing was done on all of the test data.

Overall, the accuracy results of 44.75% combined was not acceptable. Ghost notes were the main contributor due to the nature of the running-mean peak-picking algorithm used. If note onsets appear too slow, there will be a time in which small peaks which are generally just small background noise or static, will pass the threshold. Another analysis is that, in the case of test data containing many missed notes, this was mainly due to the speed at which these notes were played. It is possible to include these missed notes by varying the threshold, but this tends to become a problem because it will introduce the presence of even more ghost notes.

A retest was done by tweaking some of the parameters for the algorithms, which improved the accuracy results to 59.09%.

In summary, although the second test run seems to overall have a better accuracy, the main reason for the low accuracy of the first test run are mainly due to the large amount of ghost notes resulting from two data set, which were mainly composed of chords. Thus, if overall performance are more focused on individual data and detecting actual onsets, the first test run is considered to actually have achieved overall better results.

3.3 Chord Distinction Accuracy

The aim of this test was to determine the accuracy of distinction between single-notes and chords. The testing was done only on test data that contains chords.

Overall, the results are generally good with over 85% accuracy. The errors that occurred were because of inconsistency regarding the theory stated on Section 2.2.4.

3.4 Chord Labeling Accuracy

The aim of this test was to determine the accuracy of the chord labels assigned to chords detected by the system from the data. The testing was done only on the test data that contains chords.

Overall, the results are also good for chord labelling with 85% accuracy correct total. However, it should be noted that, the chords present in these test data consisted majority of Major and Minor chord structures, which are two of the basic chord structures. When tested for Seventh chords, the system was unable to differentiate it from its base chord structure (such as a Major Seventh chord being classified as a Major chord instead, or a Minor Seventh chord being classified as a Minor chord instead).

3.5 Tempo Accuracy

The aim of this test is to determine the accuracy between the estimated tempo of the data compared to its supposed actual tempo. The testing was done on all of the test data.

Overall, the results are acceptable, with 81.14% accuracy in average. It should be noted that, one of the main issues with regards to testing the accuracy of beat and tempo is the subjective nature of beat and tempo itself.

For further testing in this field, additional testing was done using software generated audio files with pre-defined tempos which were then exported into .WAV files and presented as input to the system. The accuracy for these additional tests yielded 81.56%, which is consistent with the previous result and also acceptable.

3.6 Comparison with other software

Comparison was also made regarding some of the modules with some of the existing commercial software available on the internet, Melodyne (Celemony, 2011) for pitch detection, Chordata (CLAM, 2010) for Chord Labeling, and BPM Analyzer (MixMeister, 2010) for Beat and Tempo. A summary of the results are shown in Table 1.

Table 1. Commercial Comparison Results

Category	AGMT	Commercial Software	Software Used
Pitch Detection	69.71%	82.81%	Melodyne
Chord Labeling	85.04%	93.91%	Chordata
Beat and Tempo	81.14%	76.93%	BPM Analyzer

For chord distinction and note onset detection, there were some issues encountered when testing or even finding software that deal with it.

3.7 Survey Results

Based on the results of the survey, all the testers found the system useful, particularly by helping guitarists easily create and thus share their created tablatures. Generally, their evaluation of the system is that there can be further improvement made for the onset detection and pitch detection modules. The testers were happy with the chord detection module as well as the inclusion of a beat and tempo (rhythm) for guitar tablatures.

As for the generated results, they were satisfied with the resulting guitar tablature but had mixed opinions regarding the music score. A possible cause of this is that the music score generated by the resulting MusicXML was a bit confusing because of the missed notes and ghost notes, which made the resulting score seem a bit complex.

3.8 Sample Output

Fig 5 and Fig 6 show a sample output (guitar tablature and score) of the system. The sample output is based on the input data "Happy Birthday" song played on the guitar through a microphone.

```
Estimated Tempo: 80BPM

|----------------|----------------|
|----------------|----------------|
|----------------|----------------|
|---0-------3---2|----------0-----|
|3-3----3--------|-------3-3----3-|
|----------------|----------------|
 1&&&2&&&3&&&4&&& 1&&&2&&&3&&&4&&&
```

Fig 5. Sample Tablature Output

Fig 6. Sample Score Output.

4. CONCLUSIONS

Results of the research and system were generally acceptable. The problems were commonly found in the pitch detection and onset detection module, which generally occurred due to complex test data. On the other hand, chord detection and beat and tempo module produced good results.

The resulting tablature and music score generated by the system proves to be consistent with general tablature and music score structures, as noted also by the testers, which is essential to be of contribution to the guitarist community. However, while the tablature is able to be generated without any other software, generating the music score or sheet music (which uses MusicXML) requires additional software.

Based on the results of the survey, all the testers found the system useful, particularly by helping guitarists easily create and thus share their created tablatures. Additionally, although not included in the survey results, some of the testers expressed an opinion of making a similar software for commercial music, which can be a direction for further improvement in the future. Overall, this research was able to improve on the previous research by adding chord detection and beat and tempo module, as well as being able to generate guitar tablatures.

Future works regarding this research may include improvement regarding some of the problems encountered during the research. One important aspect is improving performance with fast played notes, which was also a problem for the previous research. Improvement regarding analysis of similar chord structures is also recommended for future work as well as partial chords (chords that comprise of additional notes or lack thereof). Likewise, improving and working on additional analysis on chord distinction to determine other elements that distinguish chords from single-notes can also be improved upon. Possible improvements for estimating the tempo as well as evaluation of it is also a good point for future research. In general, improvements for working with more complex data is the basis for future work.

5. REFERENCES

Alonso, M., David B., & Richard, G., Tempo and beat estimation of musical signals.

Bello, J., Daudet, L., Abdallah, S., Duxbury, C., Davies, M. & Sandler, "A tutorial on onset detection in music signals", Speech and Audio Processing, IEEE Transactions, 13 (5), Sep 2005, pp. 1035 - 1047.

Burk, P., Polansky, L., Repetto, D., Roberts, M., & Rockmore, D., Music and computers.

Celemony, Melodyne, 2011.

Chua, E., Hernandez, R., Lautchang, W., & Ong, R., Guitar music transcription, 2009.

CLAM, Chordata, 2010.

Davies, M. E. P., & Plumbley, M. D. “Beat tracking with a two state model”, International conference on acoustics, speech, and signal processing, 2005.

Ellis, D. P. W., “Beat tracking by dynamic programming”, J. New Music Research, 2007, pp. 51–60.

Franklin, D., & Chicharo, J., “Paganini-a music analysis and recognition program”, Signal processing and its applications, 1999, ISSPA ’99, Proceedings of the fifth international symposium, Vol. 1, 1999, pp. 107 -110.

Kashino, K., & Murase, H., “Music recognition using note transition context”, Acoustics, Speech and Signal Processing, 1998, Proceedings of the 1998 IEEE International Conference, May 1998.

Kerdvibulvech, C., & Saito, H., “Vision-based detection of guitar players’ fingertips without markers”, Computer Graphics, Imaging and Visualisation 2007, CGIV ’07, Aug. 2007, pp. 419 -428.

Kerdvibulvech, C., & Saito, H. “In Markerless guitarist fingertip detection using a bayesian classifier and a template matching for supporting guitarists”, 2008.

Macrae, R., & Dixon, S., “A guitar tablature score follower”, Multimedia and Expo (ICME), 2010 IEEE International Conference, Jul. 2010, pp. 725 -726.

MixMeister, Bpm analyzer, 2010.

Patin, F., Beat detection algorithms, 2003.

Tuohy, D., & Potter, W., “GA-based music arranging for guitar”, Evolutionary Computation, 2006, CEC 2006, IEEE Congress, 2006, pp. 1065 -1070.

Stark, A. M., & Plumbley, M. D., Real-time chord recognition for live performance.