Guitar recognizer - an artificial intelligence-based teaching software

Mufan Li

Department of computer science and technology, Xidian University, Xi'an, Shaanxi Province, 710126, China

20009201227@stu.xidian.edu.cn

Abstract. This article describes a preliminary idea for guitar teaching software, briefly describing its several components: sound recording, sound recognition, assessment methods, and a feedback system. The voice entry system is as uncomplicated as possible to avoid difficulties in user input. Users can use the most commonly used tool - microphone - to input. Two approaches to voice recognition are presented to suggest possible future implementations. This system is more difficult to implement because of the actual need to consider more external factors, such as noise, volume level, etc. The handling of these issues is the biggest challenge for this system. Three main aspects are considered in the assessment system, and more assessment elements are to be further explored. Some help methods are mentioned in the feedback system and how to test their help efficiency. At the same time, the feedback system turns some difficult techniques that may hinder the learning of beginners into bonus content, so as not to discourage them from learning. The conclusion presents some of the necessary databases for the software and other system issues that still need to be addressed. Some possible methods of implementation are discussed as well as a preliminary plan for future research that could be conducted.

Keywords: software, sound recording, sound recognition, evaluation, feedback.

1. Introduction

At the beginning of learning guitar, people will probably find software to help them hand some basic skills, such as chord, fingering, and basic music theory knowledge. However, some software targeting beginning learners utilize a teaching method, which is easy for a skilled guitarist but hard to understand for the beginning learner. So, this paper come up with appropriate software to help guitar beginning learners.

There are a great number of research on guitar software, such as the novae project [1]. But the people playing the instrument are ignored, which is not a positive trend. Therefore, this paper will consider human factors, like the user's mastery of fingering and the interaction between users and software. Meanwhile, this paper will overview the research status and problems in the same field, analyze and introduce different speech recognition and feedback methods, compare the existing experimental data, and propose feasible solutions. This paper has explored possible solutions for guitar sound recognition and how to give users more effective feedback to help them learn guitar.

This article is about a possible implementation method for guitar teaching software. This software consists of a sound recording system, sound recognition, and a feedback system. The first step is sound recording. This section should not be designed to be overly complex to avoid difficulties for the user at the beginning of use. It intends to utilize users' microphones to get the input, and users also can provide a recording by themselves. The second step-sound recognition and processing are the most challenging. This paper collects several results from other articles and compared their results to find a possible method to achieve it. The third part is the evaluation system, which analyzes the data obtained from the sound recognition to judge the accuracy, timing, and mastery of the user's audio and to provide a basis for the feedback system. The user's mastery of the chord is achieved by interacting with the user, for example by uploading pictures to prove that the user has used the chord correctly. The fourth step is feedback. This is the most important part of the software because it determines the effectiveness of the learner's learning. This paper will propose possible help methods how to help users and use random tests to verify their effectiveness.

In the end, a reward system and a score-making system are proposed to increase the number of scores that can be taught by motivating users, laying the foundation for the software to be widely used.

2. Sound recording

Users have a lot of methods to upload their recording, such as microphone and recording file completed in advance. Users need to be aware of noise reduction, because too much noise can greatly influence the quality of recording identity. Meanwhile, users need to pay attention to the quality of the recording by adjusting the location of the microphone or replacing the microphone with a better quality one. Users with better conditions can go to a more professional venue for recording. For a better recording method, Vincenzo La Spesa has proposed several better ways, such as Mono recording and the Blumlein technique [2].

3. Sound recognition

Matching the right notes correctly is a difficult task because it requires taking into account many distracting factors, such as the height of the notes and the timing of the notes played. The Matching Algorithm (Figure 1) [3] solves the above two problems very well: It sets a waiting time of 50 milliseconds, and if the user plays a note within this period, a checkmark is generated to mark it, and vice versa, a cross is generated to alert the wrong timing. It also approximates the recorded notes to the most similar semitone and then matches them to the correct notes, which ensures to a certain extent that the notes played are correct.

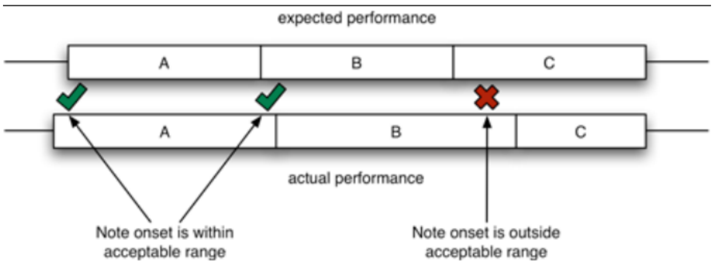
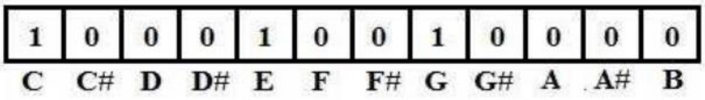

expected performance
A
B
C
A
B
C
actual performance
Note onset is within acceptable range
Note onset is outside acceptable range

Figure 1. Performance comparison using the matching algorithm [3].

In addition, the sound of the guitar is a composite of multiple notes [4], which makes the pitch-matching process difficult. Lance Alcabasa and Nelson Marcos introduced the concept of Chroma Vectors (Figure 2) [5]: it contains 12 arrays, each representing a scale. The sound made by the chords in the audio is decomposed by the system analysis, and then the decomposed information is assigned to the arrays so that the audio can be analyzed more precisely. But this can make the effects of noise very noticeable.



1	0	0	0	1	0	0	1	0	0	0	0
C	C#	D	D#	E	F	F#	G	G#	A	A#	B

Figure 2. Chroma Bin of C Major Chord [5].

The above two ways show how to deal with timing and scale matching respectively, which provides a possible solution to the processing of guitar audio: The software first matches the time of the recorded audio and then matches the pitch afterward (Figure 3).

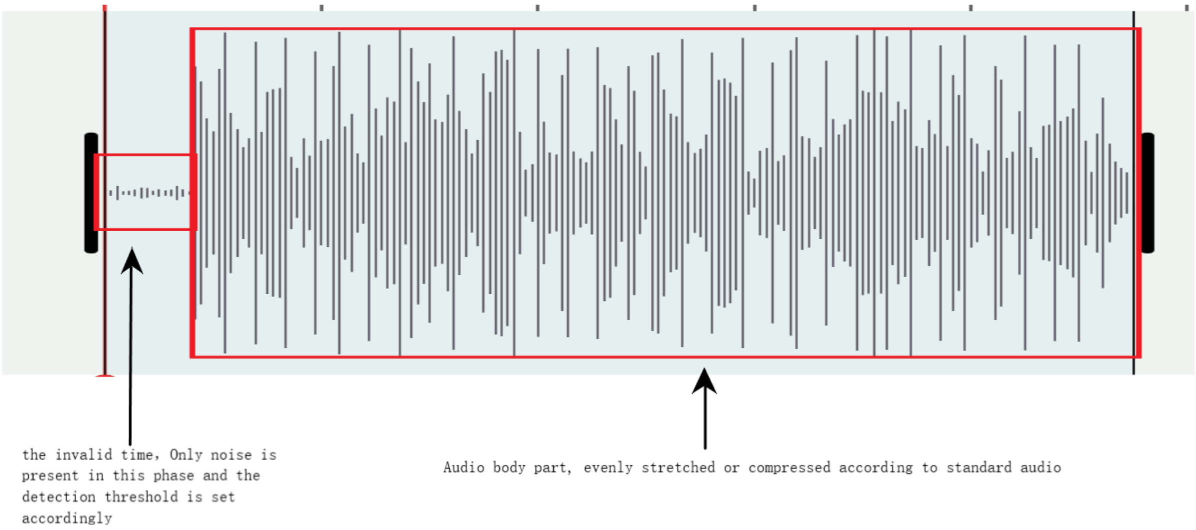

the invalid time, Only noise is present in this phase and the detection threshold is set accordingly
Audio body part, evenly stretched or compressed according to standard audio

Figure 3. Invalid time processing of recorded audio and division of subject audio.

(Photo Credit: Original)

Before receiving the recording, the user needs to select the range of the score to be played in advance to avoid large errors in sound detection. After receiving the audio, the software will first judge and separate the invalid time at the beginning and end of the audio (no audio recording) and will also judge the noise level as a way to adjust the sound analysis threshold to reduce the impact of noise. After separating the real valid audio segments, the system will analyze the duration of the audio and evenly stretch/compress the audio to the standard time, while recording the original duration information for the evaluation system to judge. The timing of the individual notes is then matched, while the notes in the original audio are matched one by one with the notes in the standard recording for scale analysis. For the scale analysis, the system can use Chroma Vectors [5] to decompose the sound and then judge whether the performance is correct or not.

4. Evaluation

After performing voice recognition, the system evaluates the data already obtained and uses it as a basis for feedback. Here are some aspects to consider when evaluating

When analyzing sheet music data, the most direct demonstration of accuracy is the similarity between user-recorded audio and annotated audio. Accuracy is determined by a combination of playing timing and scale correctness. It is a basic requirement, which directly determines the acceptability of the player to the music score learning. So, accuracy will play a dominant role in this part.

Performance skills show the basic skills of users, such as the familiarity with some chords and fingering, which will affect the speed of users when accepting new fingering and the acceptance of new music scores. For example, this is a picture of fingering, which is the form displayed on the music score. If the user is familiar with this fingering, the efficiency of learning new music scores will be significantly improved. Otherwise, it needs to be considered how to quickly help the user learn fingering.

For example, the system will ask the user if he/she is using the chord, and the user can choose to skip the chord detection by answering yes or uploading a picture of himself/herself using the chord so that the system can evaluate the chord by the picture.

First is the total time. To some extent, the performance time reflects the familiarity of the player with a piece of music. However, it should be noted that not all occasions are faster the better. For example, in the warm-up stage, the playing scale is relatively simple, so it is necessary to increase the speed to quickly enter the performance state; when playing music, try to play according to the speed specified in the music score as much as possible, too fast or too slow will affect the quality of the music.

Next is the timing of playing the scale. Although 50 ms is already a suitable waiting time, for beginners this time can be extended to 100 ms. If the user can play the correct audio within 100 milliseconds, the evaluation system will not consider it an error.

5. Feedback

This is teaching software, so interaction with people is necessary. The software will provide several help methods. Here are some initial ideas on how to help: (1) Show chords: Showing the shape of a chord on a score. (2) Show pictures: Show pictures of chord presses. (3) Play a video: Play a video including the target chord. (4) Play a recording: Play an audio clip of the target chord being played

With these help methods, their efficiency should be measured so that they can be improved better in the future. Reading Tutor [6] has provided a good method-random test: 20 groups of users are randomly offered different methods of help, and the most popular method and the least effective method are summarized at the end of the experiment. In the future, the most popular method will be utilized widely, and the least effective method will be improved.

At the same time, appropriate encouragement is necessary, but it should not be tedious and ineffective. After the user has finished playing correctly, the system generates some bonus content: advanced playing techniques. Most beginners decide to learn guitar after watching a guitar performance, but what interests them are the advanced playing methods used by guitarists, such as sweeping, sliding, and strumming. When users learn these techniques, they will feel like they are one step closer to becoming a guitar veteran. These techniques can motivate people to continue learning guitar, but it is important to note that these techniques are very difficult to learn, some beginners do not master them well, and in the worst-case scenario, this can directly discourage people from learning guitar. There are two reasons for including advanced playing techniques as bonus content: (1) it is not easy to master, and only after the user has correctly completed some simple scores will they be able to learn these techniques (2) the bonus content will specify that the technique is commonly used in guitar playing, which is a good incentive for the user to learn and reduces the likelihood of it becoming a stumbling block.

6. Project optimization

For a better user experience, it needs an easy-to-understand user interface for communication like Virtual Guitar Teacher [7]. In addition, it also needs to consider compatibility on different devices (e.g., Mac, PC) [8] and a reliable database to store user information, sheet music, technical videos, audio, etc.

If possible, it is also a good option to include a virtual guitar [9] in the software to assist the user in learning. The virtual guitar should have all the features that a real guitar has, but given the limitations of operating on the software, the left-hand chord section will be listed by the software and the user will select and have the software press out the chords instead, while the playing of the six strings can be controlled directly by the user. The sound of the guitar panel will be different when struck at different positions, so they will hit different positions of the guitar panel to simulate drumming during playing, and the software will record the sound of the drums at these different positions for the user to learn better.

After this software is actually implemented and optimized, the next goal is to consider doing what Interactive Teaching Guitar [10] did: combining software with a real guitar.

7. Conclusion

The guitar is one of the most popular musical instruments in the world, and more and more people are going to get in touch with and learn guitar. But learning the guitar is not an easy process, especially when it comes to learning some of the more difficult chords. This software can lower the threshold for users to learn guitar and motivate them while allowing them to master more content.

In the future, the software will have a search function that allows users to search for the score freely. The basis for implementing this feature is that we have a large enough number of scores. This brings us to a pressing problem: how to efficiently expand the number of scores included in the software. To solve this problem, the software will introduce a score-making system, which will provide similar functions to the software, so that the score made by the user will also have recording, recognition, and feedback functions. To motivate the users to make scores, the software will additionally design a reward system to provide gold coins to exchange for some rewards. The content of the score-making and reward system will be further developed in the future.

It needs to be focused on experimenting and improving the sound recognition part because the sound played by the guitar is very complex, one or a few mathematical models may not be able to analyze all the data, and more experiments are necessary. When looking for testers, both professionals and non-professionals need to be considered. The former can provide some professional technical guidance on the software, such as how to present the bonus content more understandably. The latter can provide some first-time experience and offer some subjective opinions for system optimization.

References

[1] Burns, A. M., Bel, S., and Traube, C. (2017). Learning to play the guitar at the age of interactive and collaborative Web technologies. In Proceedings from Sound and Music Computing Conference (pp. 77-84).

[2] La Spesa, V. (2017). The classical guitar recording.

[3] Smith, G., and Johnston, A. J. (2008). Interactive software for guitar learning. In Australasian Computer Music Conference. Australasian Computer Music Association.

[4] Héroux, I., Giraldo, S., Ramírez, R., Dubé, F., Creech, A., and Thouin-Poppe, L. É. (2020). Measuring the Impacts of Extra-Musical Elements in Guitar Music Playing: A Pilot Study. Frontiers in Psychology, 11, 1964.

[5] Alcabasa, L., and Marcos, N. (2016, March). Simple audio processing approaches for guitar chord distinction. In Proc. the DLSU Research Congress.

[6] Heiner, C., Beck, J., and Mostow, J. (2004). Improving the help selection policy in a Reading Tutor that listens. In InSTIL/ICALL Symposium 2004.

[7] Popa, E. M., Georgescu, A. V., and Barbat, B. E. (2007). E-learning with protensional agents: playing guitar. In Submitted to 12th WSEAS International Conference on COMPUTERS (ICCOMP'07), Heraklion, Crete.

[8] Kato, M. (2011). Teaching Guitar in an Online Environment.

[9] Moraru, S., Stoica, I., and Popescu, F. F. (2011). Educational software applied in teaching and assessing physics in high schools. Romanian Reports in Physics, 63(2), 577-586.

[10] Young, J. and Ullmann, M. Interactive Teaching Guitar.