COMPUTATIONAL MODELS AND MUSIC XML DEFINITION FOR COMMON PIPA TECHNIQUES

Shuqi Dai
Peking University
shuqid.pku@gmail.com

Gus Guangyu Xia
New York University Shanghai
gxia@nyu.edu

ABSTRACT

The digitalization of Chinese traditional instruments is still in its infancy. Most synthesizers, though equipped with rich sample libraries, lack basic control models to intimate common performance techniques. In this paper, taking Chinese traditional instrument pipa as an example, we have proposed a series of computational models for common performance techniques by an "analysis-by-synthesis" method, leading to more realistic synthesized pipa performances. Meanwhile, we defined symbols of those techniques in MusicXML to provide readable digital formats for computer programs. Detailed demo, experiment results, and MusicXML symbol definitions see project web site (http://shuqid.net/projects/pipa).

1. INTRODUCTION

The common techniques of traditional Chinese instruments such as Pipa and Guzheng, had neither readable digital format nor basic control models for computer programs. Taking Pipa as an example, it has over 60 techniques for both left and right hand and all sound different[1]. In order to synthesize more realistic pipa performance automatically in computers, we need to build models for realizing precise and smart control over the techniques. For example, if we want to play a pipa technique lunzhi under settled tempo and beats in a specific music style, we should first teach computer how to play it, then precisely calculate and control its time, speed and velocity according to those settings.

In this paper, we have proposed a series of computational models for common pipa techniques derived from four main basic categories: lunzhi, strumming, bending and vibrato, harmonics. They are able to cover the techniques used in most of the pipa pieces. Meanwhile we defined symbols of them in MusicXML. The program reads scores annotated with techniques in MusicXML and then uses computational technique models to generate corresponding MIDI file. We used an "analysis-by-synthesis" method in experiments to adjusting parameters in the models, and compared our model results using a commercial sample synthesizer with real pipa performance and GarageBand iOS virtual instrument performance.

2. TECHNIQUE MODELS

The common techniques used in most of the pipa pieces can be derived from four main categories: lunqzhi, strumming, bending and vibrato, harmonics.[1] We proposed computational models for them together with their combinations according to technique principles and performance experience. An "analysis-by-synthesis"[2] method was used in experiments to promote the model performance by adjusting parameters. Detailed experiment results for model parameters, synthesized demo and the MusicXML definitions can be found on the project website. [1]

2.1 Lunzhi

Lunzhi is performed as rotating right hand fingers on one or more strings and has many subtypes including four-finger single lunzhi, five-finger single lunzhi and long lunzhi. Current lunzhi models in synthesizers use uniform distribution for time and velocity of fingers, which have a huge difference from real human performance and cannot control the precise number and total time of lunzhi. Here we propose computational models controlling time and velocity for both single and continuous lunzhi.

2.1.1 Single Lunzhi

Use right hand fingers to pluck the string one by one and form continuous four or five sounds (four-finger or five-finger single lunzhi). For time control in single lunzhi, the duration of the last finger sound is determined by start time of the following technique in the piece while the time distribution of other fingers has some particular rules on average. In addition, the duration differences among finger become smaller when play in a relatively slow or fast speed. As shown in Eqn(1), T represents tempo which is counted in Beats Per Measure; D represents the total duration of the single lunzhi and is calculated in seconds; L representing the number of beats for the last finger sound; α_i is a float that controls the relationship of durations for the other finger sounds except for the last one; B is the standard number of beats for each finger sound; m indicates the type, four

[1] http://shuqid.net/projects/pipa

fingers or five fingers. α_i conforms to a quadratic function where the variable is the total duration of single lunzhi except for the last finger sound, as Equn(2).

$$L + \sum_{i=1}^{m} \alpha_i * B = 60TD, \;\; m = 3, 4 \tag{1}$$

$$\begin{aligned} \alpha_i = k_{i1}(D - \frac{60L}{T})^2 + k_{i2}(D - \frac{60L}{T}) + k_{i3}, \\ 0 \le \alpha_i \le 1, \;\; i = 1, 2, \ldots, m \end{aligned} \tag{2}$$

For the velocity control, as shown in Eqn(3), v_i represents the velocity for each finger sound, which is a positive integer from 1 to 127; V is the standard velocity and β_i is a float parameter.

$$v_i = \beta_i * V, \quad 0 \le \beta_i \le 1, \;\; i = 1, 2, \ldots, m \tag{3}$$

2.1.2 Continuous Lunzhi

Linking two or more single five-finger lunzhi together formed the long lunzhi. Even the professional musician could not play two lunzhi with the exactly same speed and volume, we add random noise parameters n_t and n_v to simulate the human bias in time and velocity control. Also, since the number of single lunzhi is various, we cannot control each lunzhi speed according to the total time duration. As a result, we add parameter K to represent number of beats for each single lunzhi and N to control the total number of lunzhi, see Eqn(4)(5)(6).

$$NK = 60TD \tag{4}$$

$$K = \sum_{i=1}^{5} (\alpha_i * B + n_t), \;\; 0 \le \alpha_i \le 1 \tag{5}$$

$$vi = \beta_i * V + n_v, \;\; 0 \le \beta_i \le 1, \;\; i = 1, 2, \ldots, 5 \tag{6}$$

2.2 Strumming

Strumming is preformed as the right hand plucking multiple strings from right to left continuously, including rapid strumming, opposite direction strumming, arpeggio(slow), small strumming(only for three or two strings). According to the experience and experiments, the time interval between each string sound can be treated as uniform distribution. As described in Eqn(7), the time control model is quite similar to Eqn(1) which shares the same definitions for common parameters. For most of strumming techniques(except for arpeggio), the value of L is much larger than B in order to make multiple plucking sound like one.

$$L + 3B = 60TD \tag{7}$$

For the velocity control, the volume of each string in strumming has a certain distribution on average according to different constructions of the strings and human performance habits, whose model equation is the same as Eqn(3).

2.3 Bending and Vibrato

Bending is performed as one left hand finger bending the string in order to increase the pitch to a target one and then return to the original pitch. Current bending models for pipa cannot control the bending speed and target pitch precisely, thus we propose a computational model controlling both two factors based on MIDI message, pitch bend. The vibrato model can be derived from the bending model by controlling the target pitch in a small range and fastening the bending speed. Detailed parameters are described in Eqn(8)(9): definitions for T and D stay the same as models described before; L1 and L2 represent the times that from starting plucking to invoking pitch bending, and from returning to original pitch to the end of sound, both counted in beats; B indicates the beats interval between calling two successive pitch bend messages; mid represents the duration beats from starting plucking to reaching the target pitch; n(t) is the value of the integer parameter in calling pitch bend message, ranging from -63 to 63; p_t and p_o represent the target and original pitch value; k controls the speed of bending and the lager k will have slower bending speed; λ is for generating k. Let $mid = 30TD + \frac{L1-L2}{2}, N = \frac{30TD - \frac{L1-L2}{2}}{B}, k = \frac{\lambda D}{(p_t - p_o)}$,

$$t(i) = L1 + B * i, \quad i = 1, 2, \ldots, N \tag{8}$$

$$n(t) = \begin{cases} \lfloor (p_t - p_o) * log_k(t - L1 + 1) \rfloor, \\ \qquad L1 \le t \le mid \\ \lfloor (p_t - p_o) * log_k(60TD - L2 - t + 1) \rfloor, \\ \qquad mid \le t \le 60TD - L2 \end{cases} \tag{9}$$

2.4 Harmonics

The natural harmonics of pipa discussed here share similar acoustic principles with guitar[3], which is performed as touching(not pressing) the string at specific positions by left hand and plucking by right hand. We tried to calculate the harmonics frequencies and control them in various filters of synthesizers, but failed to simulate a good sound. To maintain the integrity of performance, we truncated the natural harmonics of all possible pitches in real pipa performance samples and linked them to the model control.

2.5 Combination

We also proposed combination models such as lunzhi with vibrato, linking strumming and lunzhi.

3. REFERENCES

[1] China Nationalities Orchestra Society. *Pipa Piece Set*, Beiyue Art Press, 1988.

[2] M. Halle and K. Stevens. "Speech Recognition: A Model and a Program for Research," *IRE Transactions of the PGIT IT-8*, pp. 155–159, 1962.

[3] K. Bradley. *Synthesis of an Acoustic Guitar with a Digital String Model and Linear Prediction*, Ph.D. Thesis, Carnegie Mellon University, 1995.