Automatic Generation of Expressive Performance by using Music Structures

Kenzi NOIKE, Nobuo TAKIGUCHI, Takashi NOSE, Yoshiyuki KOTANI, Hirohiko NISIMURA

Department of Computer Science
Tokyo University of Agriculture and Technology

Abstract

Musical performance played by humans have expressions generated by variation of tempo and dynamics of sounds, different from mechanical performances by musical scores. We claim that this phenomenon is caused by the two facts, that a person can recognize whole musical structures, and that he can interpret melody lines of music. Based on this thought, we design an automatic expressive music performing system by using musical structures. We designed an algorithm which automatically analyzes musical structures formalized as binary syntactic trees. Such structures reflect human musical perception. We derive syntactic structures of several music scores using the algorithm, in order to research correspondences between these binary syntactic trees and musical performances played by humans.

I. Introduction

We thought that humans can express performance, which have variation of tempo and dynamics of sounds, since they can recognize whole musical structures, and they can interpret melody lines of music. Based on this thought, we plan to design a system of automatic expressive performing of music in the same manner as human performance, by using musical structures.

There are the researches of [Ohta, 1987] and [Friberg, 1991] and so on, in the research of this field.

The local metronomic speed in Waltz No.10 in b minor by F. Chopin was analyzed in [Ohta, 1987], mean value and standard deviation in every motives and phrases were calculated and it's motives and phrases were interpreted manually. And also, they calculated autocorrelation function, the way of quantitative analyze of performance's individuality was an open problem as a future theme.

He presented performance rules resulting from a project in which music performance had been analyzed by means of an analysis-by-synthesis procedure, in [Friberg, 1991]. It's were the rules to convert the written score to musically-acceptable performance with pitch, duration, chord symbols, phrase markers and key, and so on. Signs for phrase was edited manually.

Our research consists of the four stages as follows:

1. to design the algorithm which automatically analyzes musical structures from musical score information.
2. to derive syntactic structures of music by using the algorithm.
3. to construct performance rules from the correspondences between the syntactic structures of music and musical performances played by humans.
4. to generate artificial expressive performance data by applying these rules.

The analysis algorithm given as the stage 1, can extract musical inner structures, it is simple and mechanical.

We have confirmed in the stage 2 that, generated by the algorithm structures of binary syntactic trees reflect human perception of musical structures, and syntactic structures of music are derived by the it.

The stage 3 is now in progress. We have confirmed that there are the correspondences between syntactic structures of music derived by the algorithm and musical performances played by humans. We are doing to construct the rules of its, now.
The data format for performing system is designed at the stage 4.

2. Analysis of the musical structures

2.1. The Structure Analysis Algorithm

We designed an algorithm for analyzing musical structure. It has the new properties to construct better structures. First “pitch ratio” between adjacent notes is used in it. Second, rests are dealt with in it. The algorithm is shown below:

step 1. To get a symbol string, replaces each musical note by a symbol in the matter that same symbols are used for the note which have the same pitch and duration. Have each rest be dealt as the property of its adjacent note. The note with a rest is replaced by another symbol than any note without a rest.

step 2. To find a pair of symbols, which occurs in the string twice or more times.

step 3. If the pair is found, to replace all its occurrence by a new unique symbol, and to repeat from step 2.

step 4. If there remains no pair in the condition of step 3, then to find a pair of symbols whose occurrences have the same duration pattern and the same pitch ratio. Pitch ratio means the ratio between the pitch of the last note in the first sequence and that of the first in the second that is, difference of semitone new counts.

step 5. If the pair is found, to replace all its occurrences by a new symbol. If there is a pair which has the same condition and has already been assigned a symbol such as “A” by the step 3, then the new symbol is the old one with a prime symbol such as “A’”.

step 6. If there remains no pair in the condition of step 5, then to restart the process from the step 3, making pairs of the newly obtained symbols.

step 7. If no new symbols are generated, then to stop.

By this algorithm, binary syntactic trees having syntactic structures of music are generated.

2.2. Evaluation of The Structure Analysis Algorithm

We analyzed syntactic structure of music by applying the structure analysis algorithm to several musical scores.

The result of it was very similar with human perception of musical structures except negligible difference. It can construct better binary syntactic trees from the score of non-stressed start, because the minimum analysis unit of this algorithm is a single note. "Plausing of melody contains the rests works well, since this algorithm deals with rests. It can not construct better trees very well, if it is applied to similar melody of pitch transition pattern played by changing rhythm pattern, because its free degree of rhythm analysis low.

3. Analysis of correspondence between structures and performances of music

3.1. Elements analyzed from performances

We analyze piano performances played by humans to construct rules that describe the relationship between the binary syntactic trees and the human performances. We propose a performance model, based on the three following presumptions:

P1. We discriminate right and left hands.
P2. Each notes occupies the interval between key-on time and its next note’s, [Oba, 1987]
P3. All the keys of notes in a chord are pressed at the same time.

The following parameters define our performance model in these presumptions.

1. MI(i) is the local metric speed value of i-th note in a musical score.  
2. A(i) is the ratio parameter between a performance time of key-on to key-off of i-th note in a musical score and its time is MI(i).  
3. V(i) is the dynamics parameter of i-th note. (equal to the velocity value of MIDI)

One can see tempo change in performances from M(i) and articulations from A(i). The A(i) relates to the musical terms below. We consider it plays an important role to derive the fluent melody lines and beautiful sounds in the musical performance.

A(i) = 1 means legato.
A(0) < 1 means staccato.
A(0) > 1 means legato/slow.

The three parameters were obtained from data disks sold public, called "MUSIC DISK for YAMAHA PIANO PLAYER". It may not modify information of playing time for the reason of difficulties; though we guess the maker may modify several mistaken touches. The highest note among the notes that constitute a chord, determines M(0), A(0) and V(0) in a chord because of the two following reasons.

(1) A human player does not press keys of chord notes at the same time in performances usually. Several inspections indicate he presses all the chord notes when he presses the highest note.
(2) The highest note is highlighted among chord notes.

We regard a rest as a part of its antecedent in investigating M(0) and A(0).

3.2. The relationship between syntactic structure and performance
We research the relationship between the syntactic tree sequence generated by the structure analysis algorithm and the results of analyzing performance by the performance model in section 3.1.

We found the two facts as follows:

(1) The same top trees occurred in the tree sequence are performed in very similar manner as each other.
(2) The similarity of the performance of the same subtrees depends upon the equivalence of the paths between the subtrees and the top nodes.

We conclude that we must consider the hierarchy and branching of binary syntactic trees when we construct the rules from relation of between musical structures and performance.

In this field, [Friberg, 1991] rules for musical performance. The rules is not played expressive performance from syntactic structures of music.

4. The Performance System
We have made a performance system which plays a MIDI piano, in accordance with performance data generated by applying the rules to syntactic structures of music.

We will evaluate the rules which are constructed according to the results in section 3, by listening to the MIDI piano performance this system plays.

5. Conclusion
We designed an automatic expressive music performing system by using syntactic structures of music.

As the first step, we designed an algorithm which automatically analyzes musical structures from musical score information, formalized syntactic structures of music as a sequence of binary syntactic trees. This algorithm derived syntactic structures of music which resembled human perception of musical structures by applying several musical score.

Next, we defined a performance model which has three parameters, and analyzed piano performance played by humans.

We also researched relationship between musical structures and musical performance played by humans from the result data of musical structures and human performance.

We also found out that the difference of performance on the same tree depended upon the location of the hierarchy of the sequence of binary syntactic trees. We are constructing the rules of relation between information getting from binary syntactic trees and expressive performance.

Quantitative description and generalization of the rules for expressively performing from information of hierarchy and branching of binary syntactic trees will be the next themes.

References