The Eighteenth International Symposium on Artificial Life and Robotics 2013 (AROB 18th '13),
Daejeon Convention Center, Daejeon, Korea, January 30-February 1, 2013

A study about interactive musical editing system for automatic piano : Inferring performance expression by considering tempos

Tomohiro Inoue[1], Eiji Hayashi[2]

[1]Kyushu Institute of Technology Graduate School of Computer Science and Systems Engineering, Japan
[2]Kyushu Institute of Technology Department of Mechanical Information Science and Technology, Japan
(Tel:0948-29-7775)

[1]inoue@mmcs.mse.kyutech.ac.jp
[2]haya@mse.kyutech.ac.jp

Abstract

We developed a system that allows a piano to perform automatically. To enable an automatic piano to play music with the emotional expression that a skilled human pianist can provide, all of the notes in the score must be arranged. The simulation of the emotional expression in an actual performance by a highly skilled pianist would take a great deal of time by a music editor, and in the present research, we structured an interactive music editing system to edit music more efficiently.

The system is based on a concept, "Similar-patterned phrases within the same composition are performed with similar emotional expression," and is composed of two systems: the "searching system of similar-patterned phrases" and the "inferring system of emotional expression in a performance."

In the searching system, we use dynamic programming (DP) matching as the method for searching for similar phrases. In the inferring system, we use Musical Rules Databases that contain information regarding the musical symbols used by pianists for inferring emotional expression in a performance.

Key words: *automatic piano, knowledge database, computer music, DP matching, Music Interface*

1. Introduction

We developed a system that allows a piano to perform automatically. In this system, 90 actuators are installed in the keys and pedals of a grand piano. These actuators perform the key strokes and pedaling during the piano's performance (Fig. 1).

To enable an automatic piano to play music as performed by a human pianist, we must arrange all of the notes in scores. In the case of piano music, the average score contains 1,000 or more notes. Moreover, a skilled pianist is able to play an unfamiliar piece of music by sight, even if the performance is not completely in accord with a specific musical interpretation. In contrast, current computing systems, such as MIDI, cannot truly perform a new piece of music by sight in that they cannot simulate the emotional expression of a human pianist's performance without direction. Therefore, a simulation of the emotional expression in a performance by a highly skilled pianist would take a great deal of time by a music editor. In the present research, we developed an interactive music editing system to edit music more efficiently and to enable the automatic piano to simulate the emotional expression of a skilled human pianist."[1]

This system is composed of a concept that we arrived at by analyzing pianists' expressive tendencies: "Similar-patterned phrases within the same composition are performed with similar emotional expression." In accord with this concept, the editing system is composed of a "searching system of similar-patterned phrases" and an "inferring system of emotional expression in a performance."

In the searching system, we use dynamic programming (DP) matching. This method maps two character strings dynamically and calculates the degree of disagreement of the strings, which is used as an index to determine resemblance.

In the inferring system, we use Musical Rules Databases that are programmed with the musical knowledge of pianists. Their data express the influence of musical symbols numerically. Our interactive musical editing system refers to these databases and infers similar phrases automatically.

Here we describe and evaluate the results of inferring similar phrases when using the parameter called "Step," which determines the tempo.

Fig.1. The automatic piano.

2. Music Editing Support System

2.1. System architecture

The structure of the music editing support system is shown in Figure 2. The human user edits music via the user's interface on a computer display. The interface edits automatically by referring to the Musical Rules Database which contains information regarding the musical symbols used by pianists. The user's editorial work is greatly reduced and efficient editing becomes possible.

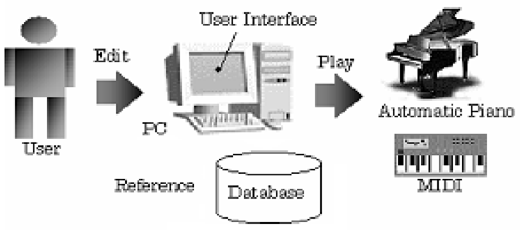


Fig.2. Structure of the editing system.

2.2. Format of performance information

The parameters of the performance information are shown in Table 1.

The data for our automatic piano system are similar to those used by MIDI(Musical Instrument Digital Interface). We defined the performance information by dividing it into two categories: the notes and the pedals. The note information is comprised of the six parameters involved in producing a tone: Key (note), Velo (velocity), Gate, Step, Bar, and Time. "Velo" is the dynamics, given by the value of 1 to 127. "Gate" is the duration of the note in milliseconds (ms). "Step" is the interval of time between notes, and it also depends on the tempo. "Bar" is the vertical line placed on the staff to divide the music into measures.

The pedal information is comprised of four parameters: Key (indicating the type of pedal: Damper or Shifting), Velo (the pedaling quantity), Time (the duration for which the pedal is applied), and Bar.

Table.1. Parameters for automatic piano.

Parameter	Key	Velo	Gate	Step	Time	Bar
Range	21~108	0~127	Depend on Tempo	Depend on Tempo	0~	1~
Unit	–	–	ms	ms	ms	–

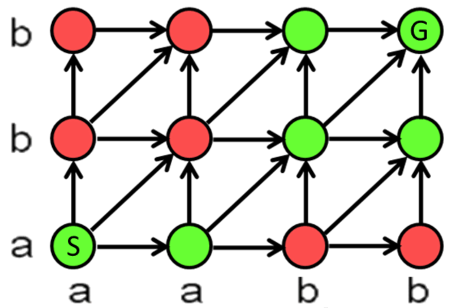
2.3. Searching system of similar phrases

2.3.1. Dynamic programming (DP) matching

We used dynamic programming (DP) matching as the method for searching for similar phrases. This method maps two character strings dynamically and calculates the degree of disagreement of the strings, which is used as an index to determine resemblance.

As an example, we'll explain how the resemblance is determined between two strings, "abb" and "aabb." DP matching is applied through the following steps.

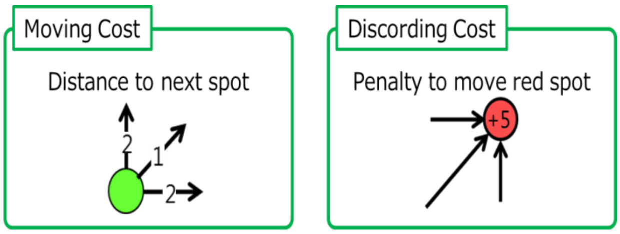
Step 1. Draw a route consisting of the strings.



S: Start, G: Goal

Fig.3. A route consisting of the strings.

Step 2. Set the costs for calculating the minimum value.



Red spot: Spots composed of different characters

Fig.4. The costs for calculating the minimum value.

Step 3. Calculate the minimum value to achieve the goal.

In this example, the minimum value is 4 and the shortest route is shown by the red arrow.

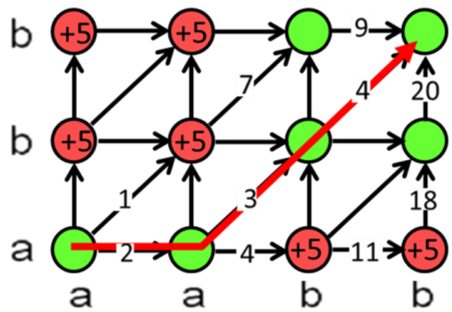


Fig.5. Total costs for reaching the goal.

Step 4. Check the one-to-one correspondence between the strings.

Pursuing the shortest route, we obtain a one-to-one correspondence of the strings. The character string, "abb", is expanded to "aabb" and this string corresponds to "abbb" (Fig.6).

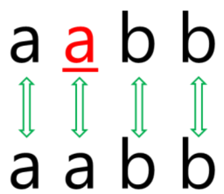
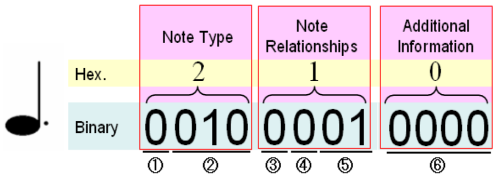


Fig.6. Correspondence between the strings.

2.3.2. Converting the notes into character strings

For use with the DP matching method, we built a Score Database which contains information regarding the musical score, with a field called "Note Value." The data in this field indicates a particular type of note, e.g., a quarter note, a triplet, and so on. Our interactive musical editing system converts notes into character strings using the "Note Value data." (Fig. 7).



1. Note or Rest (0: Note, 1: Rest)
2. Note Value (000: A whole note, 001: A half note …etc.)
3. Tie (0: No tie, 1~: The number of ties)
4. Ornament (1: This note has an ornament)
5. The number of dots (except that if it is "11" then the note is a tuplet.)
6. Additional Information

   The number of tuplets (Triplets: 0011)

   The type of ornament (Trill: 0101), etc.

Fig.7. Note values.

2.3.3. Process of searching similar phrases

First, we designate an input phrase for searching similar-patterned phrases and we set a threshold for the total costs.

Then, the searching system searches for phrases that resemble the input, using DP matching. Throughout this process, this system outputs all phrases that have a minimum cost less than the threshold as similar-patterned phrases. In our study, we called these phrases Similar Phrases.

These phrases are not edited, so we infer the emotional expression in a performance by using the inferring system.

2.4. The inferring system of emotional expression in a performance

2.4.1. Parameters used in the inferring system

In this example of the inferring system, we use the parameter called "Step," which is the interval of time between notes and which determines the tempo.

Each note, e.g., a quarter note, has an absolute time that depends on the tempo. We called this time the "Phonetic Value."

However, increases and decreases in this time exist in performance by a human pianist: thus, we introduce the parameter "StepRate"

$$\text{StepRate} = \frac{Step[ms]}{Absolute\ Time[ms]} \quad (1)$$

To compare the tendencies of each note, we store this parameter in the relevant Musical Rules Database.

2.4.2. The Musical Rules Databases

The Musical Rules Databases contain the musical knowledge of pianists. We analyzed the MIDI data of performances of highly skilled pianists in order to observe the pianists' expressive tendencies, and we used these tendencies and musical knowledge to build these databases.

Musical Rules Databases are composed of five tables: Dynamics marks, Articulation marks, Symbols of Changing Dynamics or Changing Tempo (symbols that affect the speed of a note or the increase or decrease of the volume), Time signature, and Tempo marks. Each database has data that expresses the influence of musical symbols numerically.

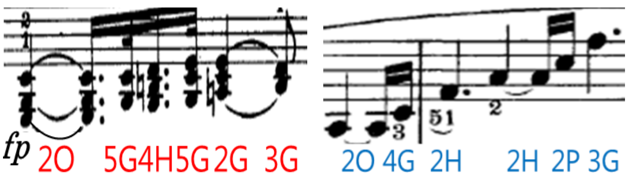
RSoho(60)	番号	記号	VeloRate	GateRate	StepRate	Gate_R_netu	Step_R_netu
1	1	mezzo_staccato	1.00	0.42	0.68	1.00	1.00
2	2	accent	1.00	0.42	0.68	1.00	1.00
3	3	legato	1.00	0.42	0.68	1.00	1.00
4	4	marc.	1.00	0.42	0.68	1.00	1.00
5	11	sf	1.22	0.58	0.70	1.36	1.02
6	5	slur	1.00	0.42	0.68	1.03	1.00
7	12	slur_end	0.95	0.42	0.68	0.95	1.00
8	10	slur_staccato	1.00	0.42	0.68	1.00	1.00
9	6	sostenuto	1.00	0.42	0.68	1.00	1.00
10	7	spiccato	1.00	0.42	0.68	1.00	1.00
11	8	staccato	1.00	0.42	0.68	1.00	1.00

Fig.8. Example of a Musical Rules Database.

2.4.3. The inferring process

Here we explain the process of inferring emotional expressions in a performance, using an example with two phrases: part of the first movement of Beethoven's 'Sonata Pathétique' as an input phrase, and the first movement of his sonata Appassionata as a similar phrase (Fig.9a. and Fig.9b).

Inferring emotional expression in a performance of similar-patterned phrases is accomplished through the following steps.



(a) Input phrase (b) Similar phrase

No.	Musical notations: Dynamic	Musical notations: Articulation		No.	Musical notations: Dynamic	Musical notations: Articulation	Position in the bar
1	p	fp	⇔	1	pp	slur	
2	p	–	⇔	2	pp	slur	
3	p	–	⇔	3	pp	slur	
4	p	–	⇔	4	pp	slur	Last
5	p	–	⇔	5	pp	slur	First&Before the rest
6	p	–	⇔	6	pp	slur	

(c) Correspondence between the two phrases

Fig.9. Phrases for inferring system.

Step1. Apply the input phrase data to the similar phrase.

We change the similar phrase's value of "StepRate" into the same value of the input phrase considering the correspondence between the two phrases (Fig. 9c).

Step2. Revise by using "Dynamics marks."

We changed the dynamics marks of the input phrase into the same dynamics marks of the similar phrase.

As illustrated in Figure 9c, we changed the dynamics marks "*p*" into "*pp*."

Step3. Revise by using "Articulation marks."

We changed the articulation marks of the input phrase into the same marks of the similar phrase.

As shown in Figure 9c, we changed the articulation marks of all of the notes to "slur."

Step4. Consider the positions of the notes.

"A note before a rest," "First note of a bar" and "Last note of a bar" are all notes with a particulartendency (Fig. 10).

A note with a tendency has these values if the "StepRate" of the standard note is 1. (Fig. 10)

In this case, note No. 4 in the similar phrase must be changed to the value of the "StepRate" by considering this relation.

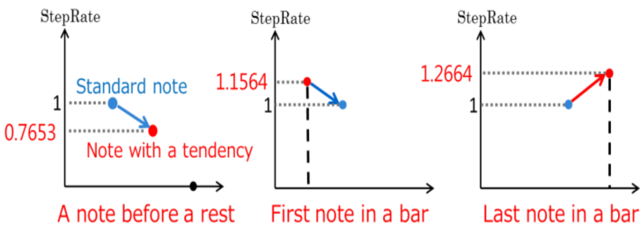


Fig.10. Values of notes with a tendency.

By following the above steps, we obtain results such as those shown in Figure 11. The results provide three values: the "Input phrase," which is non-edited data, the "Result of inferring" and the "Pianist Data," which is the actual emotional expression in a performance by a pianist in the similar phrase.

We have consistently found that the emotional expression represented by the "Result of inferring" resembles the "Pianist Data" more than it resembles the "Input phrase."

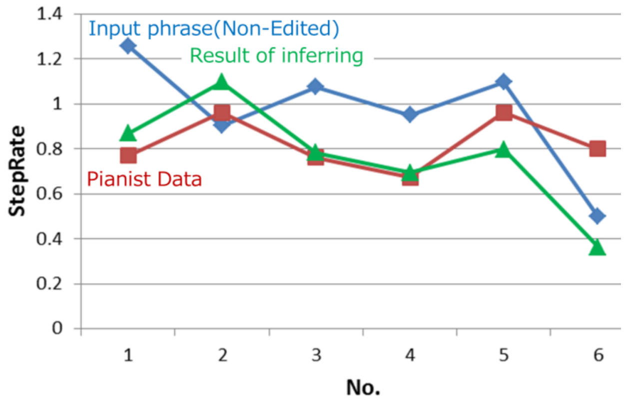


Fig.11. Results of inferring.

3. Conclusions

We developed an interactive musical editing system to edit music more efficiently. This system is composed of a "Searching system of similar-patterned phrases" and an "Inferring system of emotional expression in a performance. The Inferring system infers the emotional expression in a performance of similar phrases by referring to the databases automatically, greatly reducing the time needed for editing. We improved the Inferring system by using StepRate, which is the ratio of the Step to the PhoneticValue and is stored in the Musical Rules Databases.

Moreover, the interactive musical editing system provides edited phrases that resemble the "Pianist Data," which are the actual emotional expressions in a performance by a skilled human pianist.

References

[1] Hayashi, E. et al, "Behavior of piano-action in a grand piano.I", Journal of acoustical Society of America, Vol.105, pp.3534-3544, 1999.