TWO-STEP INPUT METHOD FOR SUPPORTING CONSTRUCTION OF MIDI SEQUENCE DATA

Chika Oshima[†], Yohei Miyagawa[†], Kazushi Nishimoto[†‡], and Takashi Shirosaki[†]

[†] *Japan Advanced Institute of Science and Technology, Hokuriku*
[‡] *PRESTO, JST*
{cooshi, y-miyaga, knishi, t-shiro}@jaist.ac.jp

Abstract   A great number of people desire to actively enjoy music. However, it requires a strong effort for a person to compose or perform a musical piece with conventional musical instruments. In this paper, we propose a "two-step input method" to construct MIDI sequence data more easily than by the conventional input methods. Furthermore, it can give people the pleasure of producing their individual expression. We illustrate the two-step input method and evaluate its effectiveness based on subjective experiments.

Keywords: MIDI sequence data, two-step input method, musical expression, indiscrete musical elements, discrete musical elements.

1.   Introduction

Music is one of the most popular forms of entertainment. Since a downtown area is full of music, we have many opportunities to come in contact with recorded music. However, while most people are content to passively enjoy of music, i.e., just listening to music, many would probably like to actively enjoy music, i.e., composing and performing music, as evidenced by the "karaoke" boom. From our eariest history, it has been natural for people to have the desire to convey their feelings or messages by performing music[3] . However, it requires a great effort for someone to arrange pitches to make or reproduce a melody.

Recently, the computer can perform any musical piece or even an original musical piece based only on the input of MIDI sequence data. Currently, there are two ways of constructing MIDI sequence data, i.e., a real-time input method and a step-by-step input method (e.g., XG-

works[6], SONAR[2]). In the real-time input method, the sequence data are basically constructed by performing on a musical instrument that outputs MIDI data in the usual manner. However, it is impossible for people who are unfamiliar with playing a musical instrument to use this method. In the step-by-step input method, people construct MIDI sequence data by separately inputting the pitch, velocity, onset time, and offset time of each note as numerical values. However, it is difficult to control the balance among the elements, i.e., velocity, onset/offset time, as well as to provide "nuance" to each element. As a solution to the problems of the conventional methods, we propose a "two-step" input method.

2.   SETUP OF COLORING-IN PIANO

In this section, we introduce the Coloring-in Piano (CiP) as preparation to describing the two-step input method. Figure 1 shows the set up of CiP. CiP consists of a MIDI keyboard, a music-database, a function for replacing note numbers, and a tone-generator. Before performing, it is necessary to enter the sequence of MIDI note-numbers (corresponding to pitches) of the piece to be performed into the music-database. While performing, the replacing function replaces the played note-numbers with the note-numbers registered in the music-database,based on the order in which they were entered. Accordingly, the correct note number is always output by touching any key. On the other hand, the expressive elements, i.e., note-on (key down) velocity, note-off (key up) velocity, onset/offset timing, and pedal messages, are output as the performer plays. Consequently, the replaced pitch numbers are input into the sound generator with the expressive elements preserved as they were performed.

3.   TWO-STEP INPUT METHOD

In the two-step input method, the process of constructing MIDI sequence data is divided into two steps. In the first step, only the sequence of pitches is input without giving consideration to the velocity and the onset/offset time data: this can be done in a similar manner to the step-by-step input method. Then, in the second step, the expression of the piece, i.e., the velocity and onset/offset time data, are input with CiP by using the sequence of pitches input in the first step. Then, the performance obtained by integrating the expression input by CiP and the sequence of the pitches is recorded by sequencing software in a similar manner to the real-time input method.
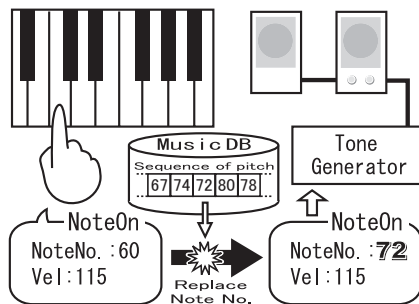
Figure 1. Set up of CiP

4. EXPERIMENTS

We conducted subjective experiments to compare the two-step input method with the two conventional input methods. We employed eighteen subjects who are students at an undergraduate institution as well as our graduate school. The subjects were asked to input "Akatonbo," which is a very famous Japanese children's song, by the three methods. Before the experiments on constructing MIDI data, we gave three basic musical tests, i.e., sight-reading, beating rhythm, and fingering. In the fingering test, the subjects were asked to smoothly play six notes. The six notes were D, G, A, C, F, and G: the interval between the first D-note and the last G-note is the eleventh (one octave and two and a half tones). We denoted a sequence number on the corresponding key for each of the six notes and asked the subjects to play the keys by following the sequence numbers. In order to smoothly play the six notes, it is necessary to use the technique of the thumb passing through the other fingers. We evaluated whether the subjects could use this technique. Eleven subjects who have played keyboard instruments succyeeded in this task.

Then, after the subjects listened to a performance of "Akatonbo" by a professional vocalist [5], they were asked to construct the MIDI sequence data of "Akatonbo" by the three methods. In addition, they were asked to be as expressive as possible in emulating the professional vocalist. Before the experiments, we explained each way of inputting data and had them practice each method with another small piece. The experiment's time limit was thirty minutes for each method. If a subject was satisfied with her/his data or threw up her/his hands in despair over one of the methods, she/he could stop constructing the MIDI data before the thirty-minute session expired. Each subject completed a subjective evaluation for each method after constructing or attempting to construct the MIDI sequence data by all of the methods. The subjects

evaluated the degree of satisfaction (1: non-satisfaction to 5: very high satisfaction) with each set of the constructed sequence data and the difficulty (1: very easy to 5: very difficult) of inputting four elements, i.e., melody, rhythm, Agogik, and Dynamik, by each method. We explained to the subjects that Agogik means the detailed transition of a tempo in a piece. Additionally, the subjects were asked about their experience in playing musical instruments and making MIDI data.

5.   RESULTS

Tables 1 and 2 show the average evaluation values for the real-time input method, the step-by-step input method, and the two-step input method. The results of a t-test indicate that the two-step input method is easier and imposes a lighter burden than the other two methods (In tables 1-6, three asterisks (***) indicate a significant difference at 1%, two asterisks (**) indicate a significant difference at 5%, and a single asterisk (*) indicates a significant difference at 10%). Moreover, the subjects tended to be slightly more satisfied with the sequence data constructed with the two-step input method than those constructed with the other two input methods for the expressive elements of Agogik and Dynamik.

To investigate the kind of people who prefer the two-step input method, the results of the evaluations were divided into two groups from the results of the fingering-test. Except for one subject, those subjects who performed correct fingering have some experience with a keyboard instrument. Tables 3 and 4 show the average evaluation values for the three methods by the subjects who performed incorrect fingering. The results of a t-test indicate that the two-step input method is easier than the other two methods for the subjects who had incorrect fingering. Additionally, they tended to be slightly more satisfied with the sequence data constructed by the two-step input method than that constructed by the real-time input method for the expressive elements of rhythm and Agogik. On the other hand, the subjects who had correct fingering felt no significant difference between the two-step input method and the other two methods. Subsequently, the results of the subjects' evaluations were divided into two groups based on whether they had ever constructed MIDI sequence data by the step-by-step input method of some sequencing software. Tables 5 and 6 show the average evaluation values for the three methods by the experienced subjects. The results of a t-test indicate that the two-step input method tended to be easier than the other two methods for the experienced subjects. Moreover, they were more satisfied with the two-step input method than with the step-by-step input method for rhythm and Agogik. On the other hand, the

inexperienced subjects felt no significant difference between the two-step input method and the other two methods.

Table 1. Average evaluation values for the two-step input method and the real-time input method .

elements	evaluation	two-step	real-time	t-value
melody	difficulty	1.89	3.44	3.50***
	satisfaction	3.22	2.56	2.13**
rhythm	difficulty	2.22	3.67	3.71***
	satisfaction	3.39	2.83	1.82*
Agogik	difficulty	2.50	3.44	2.88**
	satisfaction	3.17	2.50	2.06*
Dynamik	difficulty	2.44	3.56	2.60**
	satisfaction	3.28	2.56	2.06*
emotional burden		2.28	3.22	2.12**

Table 2. Average evaluation values for the two-step input method and the step-by-step input method.

elements	evaluation	two-step	step-by-step	t-value
melody	difficulty	1.89	2.77	2.29**
	satisfaction	3.22	3.28	0.17
rhythm	difficulty	2.22	2.50	0.68
	satisfaction	3.39	3.00	1.07
Agogik	difficulty	2.50	3.72	3.33**
	satisfaction	3.17	2.33	3.39**
Dynamik	difficulty	2.44	3.28	1.87*
	satisfaction	3.28	2.61	1.94*
emotional burden		2.28	3.22	2.12**

Table 3. Average evaluation values for the two-step input method and the real-time input method by the subjects who had incorrect fingering.

elements	evaluation	two-step	real-time	t-value
melody	difficulty	1.30	4.00	7.36***
	satisfaction	3.30	2.10	2.71**
rhythm	difficulty	1.90	3.70	3.25***
	satisfaction	3.50	2.70	1.50
Agogik	difficulty	2.40	4.00	4.00***
	satisfaction	3.20	2.10	2.01*
Dynamik	difficulty	2.20	3.80	2.67**
	satisfaction	3.40	2.40	1.79
emotional burden		2.20	3.60	2.26**

Table 4. Average evaluation values for the two-step input method and the step-by-step input method by the subjects who had incorrect fingering.

elements	evaluation	two-step	step-by-step	t-value
melody	difficulty	1.30	2.80	3.14***
	satisfaction	3.30	3.70	0.84
rhythm	difficulty	1.90	2.80	1.59
	satisfaction	3.50	3.20	0.76
Agogik	difficulty	2.40	3.60	4.13***
	satisfaction	3.20	2.70	0.96
Dynamik	difficulty	2.20	3.60	3.28***
	satisfaction	3.40	2.30	2.70
emotional burden		2.20	3.30	2.18**

Table 5. Average evaluation values for the two-step input method and the real-time input method by the subjects who had previously constructed MIDI sequence data with a step-by-step input method.

elements	evaluation	two-step	real-time	t-value
melody	difficulty	1.60	3.30	3.79***
	satisfaction	3.10	2.60	1.17
rhythm	difficulty	1.90	3.80	4.67***
	satisfaction	3.50	2.70	1.81
Agogik	difficulty	2.20	3.30	4.71***
	satisfaction	2.80	2.70	0.32
Dynamik	difficulty	2.30	3.30	2.02**
	satisfaction	3.01	2.70	1.00
emotional burden		2.80	3.30	0.76

Table 6. Average evaluation values for the two-step input method and the step-by-step input method by the subjects who had previously constructed MIDI sequence data with a step-by-step input method.

elements	evaluation	two-step	step-by-step	t-value
melody	difficulty	1.60	3.10	4.02***
	satisfaction	3.10	3.20	0.22
rhythm	difficulty	1.90	2.70	1.35
	satisfaction	3.50	2.60	1.96**
Agogik	difficulty	2.20	4.00	3.86***
	satisfaction	2.80	1.80	3.35***
Dynamik	difficulty	2.30	3.10	1.21
	satisfaction	3.01	2.70	0.77
emotional burden		2.80	3.70	1.19

6.   DISCUSSIONS

The results, in particular, suggest that the two-step input method is suitable for the subjects inexperienced in playing a keyboard. Then, by
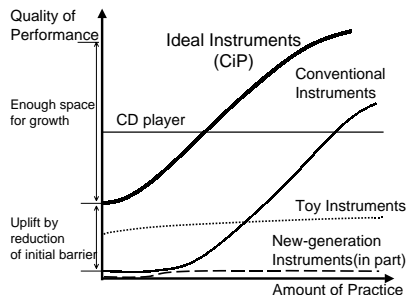
Figure 2. Comparison of musical instruments

alleviating the cognitive load in performance imposed by non-expressive elements, i.e., the sequence of pitches, we can say that it becomes easier for people to externalize more of their own musical expression. However, we would like to emphasize that the two-step input method does not enhance a person's ability in musical expression. Therefore, we think the reason why some subjects were not more satisfied the expression made by the two-step input method than with those made by the other two methods is that the subjects' own musical expressions were originally not so rich. If the subjects had a little time for practicing CiP, more subjects would be satisfied with the MIDI sequence data.

7. IDEAL MUSICAL INSTRUMENT FOR ENTERTAINMENT

Figure 2 shows a comparison of ordinary instruments/appliances with CiP. We think that some requirements must be met for an activity to be considered an ideal entertainment. The first is that the barrier to initial use is low enough for even a novice in the domain to easily participate in the activity. The second is that people have enough room for growth even though the initial barrier is low. Usually, when a person begins to practice a piece that she/he has never performed with a conventional musical instrument, there is the burden of needing to accurately repro- duce the sequence of pitches. Consequently, the first requirement of easy entry is not met although the second requirement of room for growth is met by conventional musical instruments. Some toy instruments, e.g., the Two Finger Piano [4], allow people to perform without the need to reproduce accurate pitches. However, it is impossible for people to con- trol the expressive elements at each note. The CASIO LK-40 Lighted Keyboard [1] is equipped with a "3-Step Teaching System," which is a

self-study system of the keyboard. The first step of this system is quite similar to CiP in outputting only the available pitches by hitting any key, although it always outputs the constant velocity value. Therefore, neither system gives people enough room for growth in their musical expression. Consequently, the characteristics of toy instruments shown in the graph of Figure 2 exhibits a generally flat inclination. Most new-generation instruments do not aim to allow people to reproduce pitches. Moreover, it would be very difficult for people to indicate musical expression even if they practiced quite diligently. On the other hand, CiP meets both of the above requirements. CiP frees people from the difficulty of accurately reproducing melodies. Moreover, people can develop their skills of musical expression with CiP through practice.

8.   CONCLUSIONS

Currently, there are some methods of composing and/or performing music by using a computer. However, they have not yet succeeded in fully entertaining everyone in the musical domain. In this paper, we proposed a two-step input method for constructing MIDI sequence data. We showed experiments to evaluate the effectiveness of the two-step input method in comparison with two conventional methods. The results suggest that the two-step input method is suitable for subjects inexperienced in playing a keyboard. Fundamentally, it is expected to give everyone the pleasure of playing musical instruments.

References

[1] Casio. *LK-40 Lighted Keyboards.* //www.casio.com/musical instruments/section.cfm?section=35

[2] Roland. *SONAR.* //www.roland.co.jp/products/dtm/Sonar.html

[3] A. Storr. *Music and the mind.* Harper Collins Publishers, 1992. (in Japanese)

[4] Y. Takeuchi and H. Katayose. Representation of music expression with Two Finger Piano. *IPSJ SIGNote*, NO.011-006, 1995.

[5] K. Toyoda. *Japanese song without instrumental accompaniment*, compact disc. Victor Entertainment, VICC-169.

[6] Yamaha. *XGworks.*, http://www.yamaha.co.jp/product/syndtm/ p/soft/xgwv4w