

Application of Musical Expression Generation System to Learning Support of Musical Representation

Mio Suzuki

Department of Creative Engineering National Institute of Technology, Kushiro College 2-32-1 Otanoshike-Nishi, Kushiro, Hokkaido 084-0916, Japan, mio@kushiro-ct.ac.jp

Abstract

This paper proposes a learning support system of musical representation by piano using teacher's example of musical expression that is generated based on impression expressed by an adjective with our musical expression generation system. The system evaluates learner's performance comparing with teacher's example using a Kansei space and fuzzy rules expressing the relationship between musical expression and impression. The system presents good points of learner's performance and advice by text to a learner for improvement of learner's musical representation. A learner tries to improve his/her own performance based on system's advice, and the system presents other advice again. From the experimental results, it is shown that the proposed system is useful to learn musical representation and an approach of the proposed system is suitable because the affirmative evaluation is obtained from the participants who have taken piano lesson. On the other hand, it is found that to learn musical representation is difficult using the proposed system for learner of low-performance skills.

Keywords: musical expression, impression, fuzzy inference, learning support system

1 Introduction

Musical expression is the deviation of performance from tempo marks and/or dynamic marking in a score, which means a suitable change of tempo or volume as music in a real performance. Musical expression is varied according to performance situations and/or performers. The same musical piece gives listeners different impressions depending on performances or performers. Then, it is said that performing a score correctly does not necessarily present rich musical expression. We propose a musical expression generation system called MUSAI (MUSical expression generation system by Adjective with Interaction) (Suzuki and Onisawa, 2015) that generate a musical expression reflecting any impression expressed by an adjective, where musical expression is performed by a piano. MUSAI has the relationship between musical expression and impression as knowledge, which is expressed using a Kansei space and fuzzy rules, and generates musical expression reflecting impression expressed by an adjective. Furthermore, if a user does not feel that generated musical

expression reflects impression well, MUSAI modifies musical expression based on user's evaluation by the interaction with a user. This paper applies to MUSAI to learning support of musical representation for a piano lesson.

There are some studies on learning support of music as the application of musical performance generation systems (Oshima et al., 2004; Ferrari et al., 2006). These researches make an effect to encourage willingness of a learner to practice because a computer automatically corrects a pitch of musical performance of a learner. If a teacher is not great with pianos, a teacher can show a musical representation without any concerns for a mistake about musical performance. This paper considers the case in which a teacher instructs a learner in musical representation using musical performance and verbal advice, and proposes the learning support system of musical representation with the application of MUSAI.

2 Components and Parameters of Musical Expression

This paper covers the musical expression by a piano. According to (Schmitz, 1977), changes of tempo, volume and length of a note have an influence on the impression of musical performance. Furthermore, a player can change them during the performance of a piece of music. Therefore, in this paper, tempo, volume and length of a note are considered as components of musical expression that are represented by a parameters as shown Table 1.

In this paper, the melody of a piano piece is played with the right hand and the chords are played with left hand. And the musical expression is added to a phrase of piano piece of music.

3 Learning Support System of Musical Representation

The target of the proposed learning support system of musical representation is a piano learner. Figure 1 shows the outline of learning support system of musical representation. The learning support system presents a musical expression as an example to a learner, which is generated by MUSAI. The system also presents impression to a learner using by an adjective. A learner listens to an example of musical expression and practices using a digital piano. A musical expression of a learner is recorded with

Table 1. Parameters of musical expression

Component	Parameter	Value	Meaning
Tempo	<i>TempoBase</i>	[40, 208]	Basic tempo of a performance
	<i>TempoRange</i>	[0.0, 0.6]	Variation range of a tempo
	<i>TempoVar</i>	<i>constant, decrease, increase</i>	Change type of a tempo
Volume	<i>VelocityBase</i>	[16, 127]	Basic volume of a performance
	<i>VelocityRange</i>	[0.0, 0.7]	Variation range of a volume
	<i>VelocityVar</i>	<i>constant, decrease, increase</i>	Change type of a volume
Length of note	<i>LengthSign</i>	-1, 0, 1	Show perform a note for longer or shorter
	<i>LengthBase</i>	[1.5, 10.0]	Show change of note length from in score
	<i>LengthRange</i>	0	Variation range of a length of note
	<i>LengthVar</i>	<i>constant</i>	Change type of a length of a note

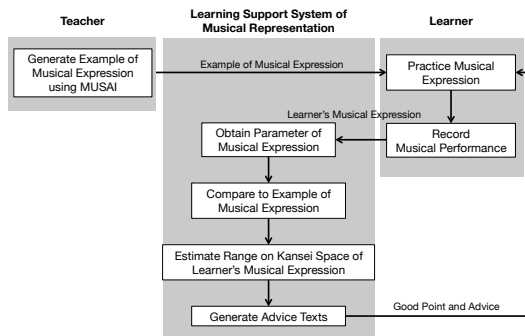


Figure 1. Outline of learning support system of musical representation

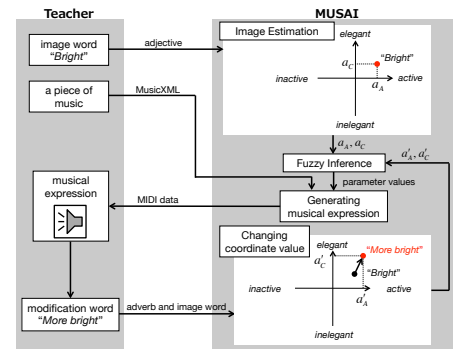


Figure 2. Detail of example of musical expression generation part

MIDI data. The learning support system obtains the parameters values of musical expression from data and compares them with the parameter values of a musical expression example. The system estimates the range on a Kansei space corresponding to learner’s musical expression, where a Kansei space is constructed beforehand. The system presents a good point of learner’s musical expression and an advice for progress in musical representation based on the comparison and the estimation. A learner continues to practice playing the piano by reference to the presented example of musical expression, the good point and an advice given by the system.

3.1 Generation of Example of Musical Expression

Figure 2 shows the outline of generating an example of musical expression using by MUSAI. A teacher inputs a piano piece as an original musical piece and an adjective expressing musical expression impression, which is called an image word in this paper.

The image word is mapped onto a Kansei space in the sense that the coordinates values on this Kansei space are obtained through the image estimation process in Figure 2, using the concept of co-occurrence of adjectives(Shimizu and Hagiwara, 2006). The parameter values of musical expression are obtained by the coordinates values on a Kansei space and fuzzy inference. MIDI data is generated using the obtained parameter values and an inputted piano piece of music, and then, the generated musical expression is presented to a teacher. If a teacher

is not satisfied with the presented musical expression, a teacher inputs a modification word consisting of an adverb and an image word, e.g., *more bright*, which expresses teacher’s evaluation of the presented musical expression. The modification part estimates new coordinates values of musical expression on the Kansei space according to the modification word, and new parameter values of musical expression are obtained by fuzzy inference. Then, the presented musical expression is modified. A teacher evaluates the modified musical expression whether he/she is satisfied with it or not. These procedures are repeated until satisfactory musical expression is obtained.

3.1.1 Construction of Kansei Space

A Kansei space is constructed using data obtained by preliminary experiments. In the preliminary experiments, various musical expressions are generated automatically by setting the parameter values of musical expression at random and are presented to the experiment participants. The participants evaluate their impressions of presented musical expressions using the semantic differential method (Osgood et al., 1957) with a 5-points scale. Then, *an active factor* and *an elegant factor* are extracted by factor analysis of evaluation data. A two-dimensional Kansei space consists of these two factors axes.

3.1.2 Image Estimation

An inputted image word is mapped onto the Kansei space in the sense that the coordinates values on this Kansei space are obtained through the image estimation process.

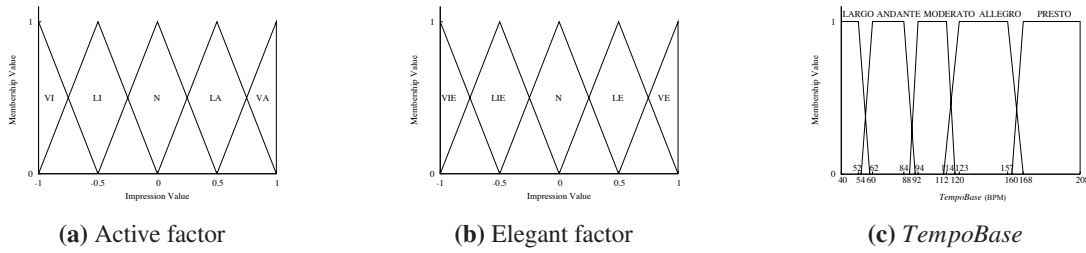


Figure 3. Membership functions correspond to the premise in fuzzy rules

In this process, at first, the co-occurrence phrase of adjectives is prepared using the inputted image word and pairs of adjectives included in the factors composing the Kansei space (Shimizu and Hagiwara, 2006).

The co-occurrence phrase of adjectives is searched using a Web search and the number of Web pages having the co-occurrence phrase is counted. And the similarity degrees of an inputted image word and the adjectives belonging to the factor of the Kansei space are obtained using the number of Web pages having co-occurrence phrase of adjectives. Coordinates values on the Kansei space are obtained using the similarity degree of an image word and adjectives included in the factors composing the Kansei space.

3.1.3 Parameter Values Estimation using Fuzzy Inference

Parameter values of musical expression are obtained using fuzzy inference by Mamdani's min-max-gravity method (Kruse et al., 1994) from the coordinates values obtained by the image estimation process. In this paper, the following fuzzy rule form is used; if the coordinate value of an image word on the *active factor* axis is \tilde{A} and that on the *elegant factor* axis is \tilde{B} , then *TempoBase* is \tilde{C} , where \tilde{A} , \tilde{B} and \tilde{C} are fuzzy sets. Figure 3(a) and 3(b) show membership functions of fuzzy sets of the premise in fuzzy rules. These are related to the coordinates values on Kansei space obtained by the image estimation. Figure 3(c) shows membership functions of *TempoBase* of the consequent in fuzzy rules. The membership functions of *TempoBase* are defined based on tempo mark corresponding to BPM (Beats Per Minute). Table 2 shows constructed fuzzy rules to estimate a value of *TempoBase*. In this paper, twenty five fuzzy rules are constructed per one parameter value. The weights of the fuzzy rules are same for all fuzzy rules. Refer to (Suzuki and Onisawa, 2015) about other fuzzy sets of the consequent in the fuzzy rules and other fuzzy rules.

3.1.4 Interactive Modification of Musical Expression

Generated musical expression is not necessarily satisfied with a teacher. Therefore, the generated musical expression is modified according to teacher's evaluation and a modification word, e.g., *more bright*. The musical expression is modified by changing the coordinates values on the Kansei space according to teacher's evaluation, and

Table 2. Fuzzy rules to estimate value of *TempoBase*

	VIE	LIE	N	LE	VE
VI	Largo	Andante	Andante	Andante	Andante
LI	Andante	Andante	Andante	Andante	Moderato
N	Moderato	Moderato	Moderato	Moderato	Allegro
LA	Allegro	Allegro	Allegro	Allegro	Allegro
VA	Allegro	Allegro	Allegro	Presto	Presto

its algorithm is based on Interactive Particle Swarm Optimization (hereinafter, referred to as IPSO) (Madar et al., 2005).

In IPSO, an individual, i.e., musical expression in this paper, has the position and the velocity, which are calculated by the following equations:

$$\mathbf{x}_j(t+1) = \mathbf{x}_j(t) + \mathbf{v}_j(t), \quad (1)$$

and

$$\begin{aligned} \mathbf{v}_j(t+1) = & w(t)\mathbf{v}_j(t) + c_1r_1(\mathbf{p}_j(t) - \mathbf{x}_j(t)) \\ & + c_2r_2(\mathbf{g}(t) - \mathbf{x}_j(t)) \end{aligned} \quad (2)$$

where $\mathbf{x}_j(t)$ is the position of j -th individual at t -th iteration, $\mathbf{v}_j(t)$ is the velocity of j -th individual at t -th iteration, $r_1, r_2 \in [0, 1]$ are uniformly distributed random numbers, $\mathbf{p}_j(t)$ and $\mathbf{g}(t)$ are the best position of j -th individual and the best position of population at t -th iteration, respectively, w is defined by (3), and c'_1 and c'_2 defined by (4) and (5) are transformed to c_1 and c_2 .

$$w(t) = (\text{RangeHigh}_j(t) - \text{RangeLow}_j(t)) \times adv, \quad (3)$$

$$c'_2 = \frac{1}{1 + e^{-(\text{Eval}P_j - 5)}}, \quad (4)$$

and

$$c'_1 = 1.0 - c'_2. \quad (5)$$

In this paper, $\text{RangeHigh}_j(t)$ and $\text{RangeLow}_j(t)$, search ranges are defined as $+1.0$ and -1.0 at the first modification step, and the search range becomes narrow at every

iteration, and *adv* is defined according to a modification word as follows: 0.25 when *a little more*, 0.5 when *more* and 0.75 when *very*. As for c_1 and c_2 , c'_1 and c'_2 obtained by (4) and (5) are transformed by adding 0.5 to the values in [0.5, 1.5]. $EvalP_j$ is the number of musical expressions that a teacher's evaluation is affirmative for j -th individual, i.e., j -th musical expression. According to $EvalP_j$, c_1 and c_2 are changed based on (4) and (5) so that musical expression at $(t + 1)$ -th iteration moves to the best position of j -th musical expression or the best position of population at t -th iteration.

Musical expressions are generated at the first modification step as follows. The Kansei space is divided into 25 ranges according to the fuzzy division as shown in Figure 3(a) and 3(b). The position, i.e., the coordinates values of modified musical expression are moved to some range according to the adverb of the modification word as follows: the next range when *a little more*, the next range but one when *more* and the next range but two when *very*. At this time, the coordinates values are determined at random in the range, and the velocity of modified musical expression is defined as zero.

Modified musical expressions are presented to a teacher again and a teacher evaluates whether the modified musical expression reflects impression expressed by an image word or not as follows:

- 1: not reflecting,
- 0: neutral,
- +1: rather reflecting,
- +2: reflecting.

And a teacher chooses one musical expression reflecting an image word best out of presented musical expressions as the best one. When a teacher wants to modify musical expression again, a teacher inputs a modification word. The velocity and the position of each musical expression are updated by (1) and (2) according to the evaluation value for musical expression at the previous step and the inputted modification word. After the second modification step, the number of modified musical expression is set as 10. If the number of musical expressions is smaller than 10, musical expressions are generated at random so that the number is 10. If there are ten or more musical expressions at and after the second modification step, top ten musical expressions having high evaluation survive at the next modification step. Above procedures are repeated until a teacher is satisfied with generated musical expressions. In this way, the MUSAI obtains an example of musical expression of a musical work for practice by a teacher.

3.2 Comparison of Parameter of Musical Expression

The presented learning support system obtains the parameter values of learner's musical expression, compares its parameter values with those of an example of musical

expression, and evaluates whether a learner plays a piano according to an example of musical expression presented by the system. The eight parameter values chosen from ones shown in Table 1, excluding *LengthBase* and *LengthRange*, are considered for the comparison. Obtained parameter values of learner's musical expression are evaluated from the following point of views: whether or not fuzzy sets which the parameter values of learner's musical expression belong to are the ones which the same parameter values of an example of teacher's musical expression belong to, where the fuzzy sets are the ones in the consequent part of fuzzy rules mentioned in Section 3.1. Based on the comparison results the learning support system presents some good points of learner's musical expression and some advices for improvement.

3.2.1 Advice for Parameter of Musical Expression

If some parameter value of learner's musical expression belongs to the fuzzy set which the same parameter value of teacher's example of musical expression belongs to, the following advice is presented to a learner: "You play the performance well for the parameter."

If all parameter values of learner's musical expression belong to the same fuzzy sets which the parameter values of teacher's example of musical expression belong to, the text is generated as good point the following and presented to a learner: "Your musical expression reflects impression expressed by an image word."

On the other hand, if the parameter values of learner's musical expression does not belong to the fuzzy set which the same parameter value of a teacher's example of musical expression belong to, advice text is presented to a learner showing how to play a performance.

3.3 Range of Learner's Musical Expression on Kansei Space

Learner's performance of musical expression is mapped to a range on the Kansei space using its parameter values. Text sentences are generated using the coordinate values of teacher's example of musical expression, those of learner's one and information on the practice effect. Text sentences present that what extent learner's musical representation reflects target impression and that what extent impression of learner's musical expression is changed by playing performance of musical expression according to the presented advice.

3.3.1 Estimation of Range of Musical Expression

The mapped range of learner's musical expression on the Kansei space is estimated using the parameter values of learner's musical expression and the fuzzy rules in MUSAI in the following way. At first, as for the parameter value of learner's musical expression that does not belong to the same fuzzy sets which the same parameter values of teacher's musical expression belongs to, the ranges of the active factor and the elegant factor axes composing the Kansei space are estimated from the fuzzy rules having

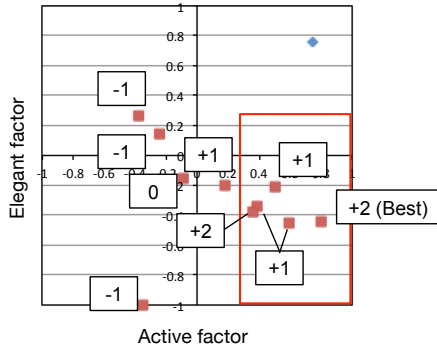


Figure 4. Example of evaluation values in estimated range

the fuzzy sets in the consequent part which the parameter value of learner’s musical expression belongs to. Next, the overlapped range of the obtained range on the active factor axis and that on the elegant factor axis is obtained on the Kansei space.

3.3.2 Advice about Impression

The system verifies whether or not the coordinate values of teacher’s example of musical expression are included in the overlapped range on the Kansei space. If the coordinate values are included in the range, advice about impression of learner’s musical expression is presented to a learner using the evaluation values of teacher’s example of musical expressions included in the range, the number of musical expression example with the same evaluation, and inputted modification words.

The system estimates whether target impression is reflected by learner’s musical expression or not using evaluation values of playing performances of learner’s musical expression. Let teacher’s example of musical expression get evaluation values during generation of teacher’s example as shown in Figure 4. Impression of learner’s musical expression is estimated as *rather* as shown below because the number of evaluation value +1 is the largest.

Your musical performance has **rather bright** impression.

If any teacher’s example of musical expression are not included in the overlapped range, the system does not give any advice.

The system gives a learner another advice using the center coordinate values of the obtained range and the coordinate values of a teacher’s example of musical expression on the Kansei space. According to the distance between the center coordinate values of the range and the coordinate values of a teacher’s example of musical expression, i.e., *dist*, the degree adverb in advice is selected as follows: *a little more* when $0 \leq dist \leq 0.25$, *more* when $0.25 < dist \leq 0.5$, and *very* when $0.5 < dist \leq 0.75$. As for the image word in advice, the system chooses the image word according to the number of affirmative expressions

in the modification words and the number of negative ones in the modification words during practice. For instance, let us assume that the system gives a learner affirmative advice *more bright*, and *a little more bright*, and negative advice, *not bright* during the practice. In this instance, a learner is given affirmative advice twice and negative advice once. Then, the image word *bright* is chosen as advice. If the musical expression is not included in the estimated range or the number of affirmation advices is equal to or less than the number of negative ones, the following text is presented to a learner.

Play the piano as given advice, then your performance of musical expression has **bright** impression.

4 Experiment

Experiment is performed in order to verify the usefulness of the proposed learning support system of musical representation. Participants are 10 males or females of 16 through 26 years old. 8 out of all participants have taken a piano lesson and other 2 participants have never taken it. In the experiment, two examples of musical expression are generated by an experimenter using MUSAI. That is, an experimenter is a piano teacher. And the participants have practice in musical representation using a digital piano and proposed learning support system until the system evaluates that learner’s musical expression reflects impression, where the upper limit number of practice repetition is fixed at 10.

Table 3 shows questionnaire items in this experiment. The participants answer the questionnaire items from Q1 to Q4 every one practice, and answer Q5 and Q6 after the experiment with the 7-points scale. Points from 1 to 3 mean negative evaluation and points from 5 to 7 mean affirmative evaluation, and point 4 means neutral. Musical pieces used in this experiment are Gnessin No. 1 written by Erik Satie and Ecossaise in G Major, WoO.23 written by Ludwig van Beethoven. The example of musical expression of Gnessin No. 1 reflects hard impression and the example of musical expression of Ecossaise in G Major, WoO.23 reflects light impression.

Table 3. Questionnaire items

Item	
Q1	Is advice helpful?
Q2	Do you think the contents of advice is suitable?
Q3	Do you have a easy to understand advice?
Q4	Do you think you can perform musical expression close to an example by advice text?
Q5	Do you think you will learn musical representation using this system?
Q6	Do you think you want to practice musical representation using this system?

4.1 Results and Discussions

Figure 5 shows the evaluation results. Although all participants musical expressions are not evaluated by the system that they reflect target impression, it is found that from Figure 5 that affirmation ratios in Q1, Q2 and Q4 are 90.0%, 95.0% and 88.2%, respectively. This means that the participants feel that presented advices are useful for the practice of musical representation using the learning support system. It is also found from Figure 5 that the affirmation ratios in Q5 and Q6 are 80.0% and 90.0%, respectively. This means that the participants feel that the learning support system is useful to learn musical representation. The participants have taken a piano lesson who have the following free descriptions about advice presented by the learning support system: “The system is useful because the system evaluates my performance objectively when practicing the piano alone.” and “It was easily practice because the system presents good points and advice for my performance.” Furthermore, the two participants have never taken a piano lesson who have the following free descriptions: “I think good about the system presents quantified my performance.” and “It was easy to understand that attend to my performance because the system presents example of musical expression and advice.”

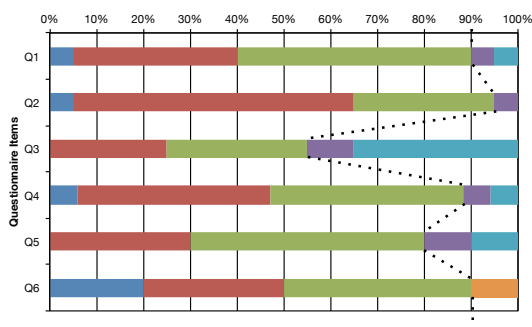


Figure 5. Evaluation results for learning support system

On the other hand, the affirmation ratio in Q3 is 50.0% lower than its in other questionnaire items. It is found the participants feel that it is hard to understand advice presented by the learning support system. Some participants have free descriptions about the easiness of understanding advice: “I understand the presented advice, but I don’t imagine how I should improve my performance.” One of the reasons is that although the system compares learner’s performance with a teacher’s example of musical expression, it does not compare the current learner’s performance with learner’s former ones.

From the above experimental results, it is said that proposed learning support system is useful to learn musical representation. And an approach of learning support of musical representation by the proposed system is suitable because the affirmative evaluation is obtained from the participants who have taken a piano lesson.

Figure 6 show the changes of the average number of

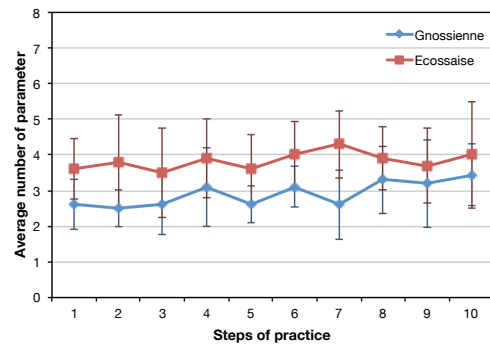


Figure 6. Change of average number of parameters of Gnessienne’s and Ecosystem’s musical expression

parameter values of participant’s musical expression for Gnessienne and Ecosystem matching to the parameter values of teacher’s example in the sense that fuzzy sets in the consequent part of fuzzy rules mentioned in Section 3.1 which the parameter values of participant’s musical expression belong to are the same as the ones which the parameter values of teacher’s musical expression belong to. Although there is no significant difference between the matching degree at the first practice and that at the tenth practice for the both performances, the increasing tendency is observed from these figures. In fact, the correlation coefficient between the number of times of practice and the average value of the matching degree is 0.771 for Gnessienne, and 0.496 for Ecosystem. This means that the parameter values of participant’s musical expression approach to those of teacher’s example of musical expression as the participants repeat practice. On the other hand, in the participant H and I who have never taken a piano lesson, the correlation coefficient between the number of times of practice and the matching degree of participant H for Gnessienne and Ecosystem are -0.572 and -0.078 respectively, and that of participant I for Gnessienne and Ecosystem are 0.232 and -0.087 respectively. One reason for this is the piano performance skill of these 2 participants is not high.

From the above, although to learn the musical representation is difficult for a learner having low performance skills by learning support system, it is said that a learner can learn how to play performance of musical expression by learning support system’s advice.

5 Conclusions

This paper proposes the learning support system of musical representation using the musical expression generation system called MUSAI and verifies the validity of the proposed learning support system by the experiment. The learning support system obtains the parameter values of musical expression from learner’s performance. And the system compares the parameter values of learner’s musical performance with those of an example of musical expression and gives a learner an advice about good points and improvement of learner’s musical expression. From

the experimental results, it is found that proposed learning support system is useful to learn the musical representation. And it is said that the approach of proposed learning support system is suitable because the affirmative evaluation is obtained by the participants who have taken a piano lesson. On the other hand, it is found that to learn the musical representation is difficult for a learner of low performance skills using proposed system. A performance skills of a learner is considered as future works.

References

- L. Ferrari, Anna R. Addressi, and F. Pachet. New technologies for new music education: The continuator in a classroom setting. In *Proceedings of ICMPC 06*, 2006.
- R. Kruse, J. Gebhardt, and F. Klawonn. *Foundations of FUZZY SYSTEMS*. JOHN WILEY & SONS, England, 1994.
- J. Madar, J. Abonyi, and F. Szeifert. Interactive Particle Swarm Optimization. In *Proceedings of 5th International Conference on Intelligent Systems Design and Application*, pages 314–319, 2005.
- Charles E. Osgood, George J. Suci, and Percy H. Tannenbaum. *The measurement of meaning*. University of Illinois Press, 1957.
- C. Oshima, K. Nishimoto, and M. Suzuki. Family Ensemble: A Collaborative Musical Edutainment System for Children and Parents. In *Proceedings of the 12th Annual ACM International Conference on Multimedia*, pages 556–563, 2004.
- Hans-Peter Schmitz. *Singen und Spielen : Versuch einer allgemeinen Musizierkunde*. Symphonia, 1977. S. Imoto and K. Takii, trans.
- K. Shimizu, and M. Hagiwara. Image Estimation of Words Based on Adjective Co-occurrences. *The IEICE transactions on information and systems*, 89(11), pages 2483–2490, 2006.
- M. Suzuki, and T. Onisawa. Musical expression generation reflecting user’s impression by interaction. *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, 27(4), pages 651–668, 2015.