Journal of Educational Research and Reviews
Vol. 7(7), pp. 155-168, July 2019
doi: 10.33495/jerr_v7i7.19.128
ISSN: 2384-7301
Research Paper

Predicting developmental degrees of music expression in early childhood by machine learning classifiers with 3D motion captured body movement data

Mina Sano

Osaka-Shoin Women's University, Higashi-Osaka city, Osaka-fu, Japan, 577-8550.

Author E-mail: minasano@yahoo.co.jp

Accepted 26th July, 2019.

Abstract. Interaction between children's developmental degree of music and their musical expression continued to intrigue researchers. Currently, one noteworthy element will be to analyze such interaction from quantitative approach and to detect some predictive methodology to replicate such interaction statistically. In this study, the author extracted developmental characteristics of musical expressions in early childhood from viewpoints of elements of body movement, and applied a classification of machine learning based method on those feature quantities acquired from the participant children. Classification models were applied to the feature quantity for 3-year-old, 4-year-old, and 5-year-old in 2 nursery schools in 2016, 2 kindergartens in 2017 and a certified facility in 2018 utilizing 3D motion capture. In order to highlight developmental degree and to extract feature quantity, a three-way non-repeated ANOVA was applied and a statistically significant difference was observed in the movement data analyzed of the moving average of distance such as pelvis and right hand, the moving average of acceleration such as right hand, and the movement smoothness of right foot. The author classified the developmental degree of children's musical expression by machine learning classifiers using the feature quantities of motion capture data after let classifiers train with categorical variables of developmental degree evaluated by the author with simultaneously recorded video. The author report here that the best classifier is Multilayer Perceptron Neural Network and the second best is Boosted Trees. The sensitivity result showed that the movement of the pelvis was strongly related to the musical development degree.

Keywords: Musical expression in early childhood, 3D motion capture, feature quantity, machine learning, multilayer perceptron neural network.

INTRODUCTION

Young children frequently play songs, pretending while singing in everyday life. Those children voluntarily create musical expressions using the whole body. Musical experience program was devised by the author to encourage musical development by utilizing pretending and dramatization in musical expression. The program was implemented in 3-year-old, 4-year-old, and 5-year-old children in nursery schools and kindergartens and certified facility, and an educational effect was examined through a qualitative and quantitative analysis (Sano 2015; Sano, 2013).

The activity contents of the program were revised in the practice process, and MEB (Musical Expression Bringing-up) program was created with reference to a theory in which dramatization and music were integrated (Rubin and Merrion, 1996). The program was constituted of four phases' activities such as (1) beginning activity, (2) pantomime and improvisation, (3) story creation, and (4) dramatization of the story. MEB program has a main purpose to encourage the recognition of musical elements in early childhood. The program begins to establish the image of a phenomenon of everyday life as the activity, such as sound awareness and advance to the formation of a rhythmic pattern and replay song with body movement such as dramatization. 3-year-old, 4-year-old, and 5-year-old children in nursery schools and kindergartens participated in the musical expression program.

In the first phase, the activity aims to establish an image by reconstructing a sensory impression in everyday life experience. The aims of activity were to provide rhythm and movement experiences in the second phase. The children participated in the music experience including a role play and reply song to encourage the recognition of musical elements in the third phase, and the children tried to integrate dramatization and music experience based on the activities from the first phase to the third phase in MEB program.

The author also devised the music test constituted of six domains including sixty items and the participant children took the music test before and after practice of MEB program to quantitatively analyze the effectiveness of the program. As a result, a change of recognition of musical elements showed a statistically significant difference (Sano, 2013; Sano, 2014). An important relationship between the recognition of musical elements and movement in early childhood has been clarified through experimental studies, although the importance of relationship was presented in previous studies (Jacque-Dalcroze, 1921; Thelen, 1979; Custodero, 1999). The experimental studies showed some results such as infant's recognition of rhythm or melody (Hannon and Johnson, 2005; Phiilips- Silver and Trainor, 2005) and infant's reaction to sounds (Zentner and Eerola, 2010). But, those researches have not intended to capture musical expression in early childhood during the continuous practice of musical expression.

Some of researches using 3D motion capture technology mainly showed a result of movement analysis regarding the specific movement about Japanese traditional dance and sawing in adulthood (Sato *et al.*, 2010). Concerning researches regarding music and movement, reaction to musical sound (Burger, 2013), viewing experiments and video analysis on the relationship between performer's actions and expressions were observed (Dahl and Friberg, 2007; Thompson and Luck, 2012). But the subject of these studies was not a musical expression in early childhood.

The author tried to extract developmental characteristics of musical expression utilizing 3D motion capture technology. Firstly, MTw system, inertial-magnetic motion tracker manufactured by Xsens Technologies B. V. was used to analyze the movement of five children wearing on forehead one by one at the same time in the practical process of musical expressions (Sano, 2016a). As a result, an importance of the moving average acceleration was found out. The author also analyzed young children's movement to calculate the moving trace and moving distance utilizing the MVN system, 3D human kinematic measurement system empowered by wired connection based motion trackers manufactured by Xsens (Sano, 2016b). Based on those inspections, more detailed data regarding the movement changes in musical expression of young children were acquired by the wireless connection type MVN system reduced in size and weight. The MVN system of wireless type was easy to put a motion tracker on young child's body. The wireless MVN system can capture the body movement of 17 places including each measurement part such as head, chest, pelvis, left and right shoulders, left and right upper arms, left and right lower arms, left and right hands, left and right upper limbs, and left and right lower legs. Children in the two nursery schools in 2016, children in the two kindergarten in 2017, and children in the certified facility in 2018 participated in MEB program and the movement analysis of musical expression during the practical process (Sano, 2017; Sano, 2018a). A part of result analysis in 2016 mainly showed some characteristics of elements of body movement in musical expression using a three-way, non-repeated ANOVA by activity phase, age, and facility (Sano, 2017).

In this study, characteristics of musical expressions were extracted from the quantitative analysis results of the elements of body movement in the musical expression of the participant children in U & K nursery schools in 2016, and F & Y kindergartens in 2017, and N certified facility in 2018. Children from 0-year-old to 6-year-old generally attend nursery schools or certified facilities, and children from 3-year-old to 5-year-old attend nursery schools, kindergartens or certified facilities in Japan. The author thought to classify and discriminate the degree of development of musical expression by machine learning based on the feature quantity of body movement in musical expression calculated from the accumulated data. As the method to classify richer musical expression is improved, this study will progress to contribute developmental support for music education in early childhood.

Purpose of this study

This study aims to explore effective statistical treatments of capturing developmental degree of children as well as to validate a method with higher classification accuracy to classify musical development utilizing machine learning based on the feature quantity of body movement as developmental characteristics of musical expression in early childhood. The key challenges will be to find appropriate feature quantity and flexible classifiers. The MVN system as 3D motion capture was used to acquire the quantitative data of 3-year-old, 4-year-old and 5-year-old in the two nursery schools in 2016, in the two kindergartens in 2017 and the certified facility in 2018.

The author tried to classify the developmental degree of musical expression in early childhood using those acquired data.

METHODOLOGY

Practice of MEB program and extraction of MVN measurement items by activity phase

The primary goal of this study is to observe statistical relationship between scores of music test after MEB program practically applied and kinetic feature quantity processed from motion capture data, and apply such findings to machine learning to forecast developmental degree from body movement data of children.

3-year-old, 4-year-old, and 5-year-old children in K & U nursery schools (n=120) in 2016, in F & Y kindergartens (n=194) in 2017, and in N certified facility (n=90) in 2018 participated in the practice of MEB program. During the activity period, one year is separated by about two months, children participated in the first phase's activity on May and June, the second phase's activity on July and August, the third phase's activity from September to November, and the fourth phase's activity on December with January. The characteristic contents by activity phase were extracted from MEB program as MVN measurement items. MVN measurement data for four times in total, once for each activity phase was recorded. Each participant child performed a music play and self-introduction with the folk song, “What is your name?” in the first phase of MEB program. The children performed song- play with pretend movement of “Shopping at a Bakery” centering on clapping and stepping in the second phase of MEB program. The children created the movement of the lion according to the music “The Grand March of the lion” edited by Tanaka Tsuneo as the excerpt of subject matter of "Introduction and Lion King's March" in "Animal’s Carnival" composed by Saint Saens in the third phase of MEB program. In the fourth phase of MEB program, child also created the movement to play a musical instrument while singing a song including a story and they sang a reply song.

Measurement method by MVN system

17 motion trackers were put on the participant child to monitor full body activity such as head, chest, pelvis, left and right shoulders, left and right upper arms, left and right lower arms, left and right hands, left and right upper limbs, and left and right lower legs for every MVN measurement. After calibrating body lengths such as arm-span, leg length, waist position, one-by-one 3-year-old, 4-year-old and 5-year-old child was inspected in two nursery schools in 2016, two kindergartens in 2017, and a certified facility in 2018. Each child’s movement in musical expression according to the piano accompaniment was captured as collected data at a frame rate of 60 Hz. MVN system is a light weight and compact device which provides less constrained environment for even small sized children. Each child needed 5-10 minutes including the measurement time of 30 seconds.

Only the young child who was authorized by the responsible person of the facility and the parents of the children participated in the MVN measurement. Since measuring a large number of children takes a long time, the measurement date of each activity phase was set for two days each, and the second day was set as a preparation day. Characteristic elements of movement in musical expression in early childhood were extracted to calculate the moving trace, the moving distance, the moving average of velocity, the moving average of acceleration, and the movement smoothness of each body part of measurement from the collected data and were quantitatively analyzed by ages and activity phase of MEB program. The movement smoothness was calculated by the ratio of moving average velocity and moving average acceleration with reference to Burger’s study (2013). The number of measurable persons is 31 in nursery schools (11 in 3 years old, 9 in 4 years old, 11 in 5 years old), 55 in K nursery school (18 in 3-year-old, 17 in 4-year-old, 20 in 5-year-old), 49 in F kindergarten (18 in 3 years old, 14 in 4 years old, 17 in 5 years old), 45 in Y kindergarten (16 children in 3-year-old, 14 in 4-year-old, 15 in 5- year- old), 47 in N certified children's facility (15 children in 3-year-old, 12 children in 4-year-old, 20 children in 5- year- old). K nursery school, Y kindergarten and N certified facility take a childcare form following the Montessori method. U nursery school and F kindergarten take a play centered childcare form. Table 1 shows the dates of MVN measurement in the five facilities conducted by the participant children for each activity phase.

Method of quantitative analysis of acquired MVN data

The author analyzed the measured values of MVN from the viewpoints of the first phase to the fourth phase of MEB program activity phase, the five facilities such as nursery schools, kindergartens, and a certified facility, the age of the participant children (3-year-old child, 4-year-old and 5-year-old child). The analysis method is a three-way ANOVA (non-repeated four standards as MEB phase, non-repeated five standards as facilities and non-repeated three standards as ages) applied to find statistically significant difference between relevant measures.

Classification training and discrimination by machine learning

From the recording data of moving images, the author divided the musical development of each child into three

Table 1. MVN measurement Dates from 2016 to 2018 in the practice of MEB program by activity phase.

	K nursery school	U nursery school	F kindergarten	Y kindergarten	N children's facility
First phase	May 23 in 2016 June 20 in 2016	May 20 in 2016 June 24 in 2016	May 30 in 2017 June 2 in 2017	May 26 in 2017 June 23 in 2017	May 25 in 2018
Second phase	July 11 in 2016 August 15 in 2016	July 15 in 2016 August 19 in 2016	July 4 in 2017 July 11 in 2017	July 14 in 2017 September 8 in 2017	July 27 in 2018
Third phase	September 5 in 2016 October 30 in 2016	September 23 in 2016	October 13 in 2017	October 20 in 2017	October 12 in 2018 October 19 in 2018
Fourth phase	December 26 in 2016 January 25 in 2017	December 16 in 2016	December 5 in 2017 January 16 in 2018	December 15 in 2017 January 12 in 2018	December 14 in 2018 December 28 in 2018

levels of high, medium and low based on the MVN measurement results. (High: 15, Medium: 27, Low: 34 people.) Classification training was conducted using this feature quantity (factor) and three-stage classification (categorical dependent variable) as training subjects of machine learning. The results of training were applied to the data of participant children in 2017 and 2018 to judge the degree of musical development, and at the same time the coincidence rate with the classification result by recorded data of moving images was calculated.

RESULTS

Firstly, the characteristic data will be described by showing a part of the result analysis that combines acquired data in 2016, 2017 and 2018. Secondly, the author describes the machine learning process and the result discussion on the extracted feature quantity.

Characteristic results regarding a three-way non-repeated analysis

A three-way non–repeated ANOVA was implemented to the MVN measurement data (activity phase factor (4 levels), nursery school, kindergarten and certified facility factor (5 levels), age factor (3 levels)). As a result, characteristic results were observed in the moving distance and the moving average acceleration regarding pelvis and right hand movement. Firstly, an analysis result regarding the moving distance of pelvis is showed.

(1) Change of the moving distance of pelvis

The result of a three-way non-repeated ANOVA is shown by change of activity phase of MEB program with ages of participant children. Table 2 shows a change by activity phase regarding the moving distance of pelvis.

As shown in Table 3, a main effect/ interaction of the test showed a statistically difference (phase: (F(3, 775)=407.216, p<.005), facility: (F(4, 775)=23.314, p<.005), phase * facility: (F(12, 775)=19.923, p<.005), facility*age: (F(8, 775)=5.771, p<.005), phase * facility * age: (F(24, 775)=5.678, p<.005). The test simple main effect with multiple comparisons was carried out by Bonferroni method.

Concerning the phase factor/ phase * facility * age, the simple main effect was statistically significant in K nursery school （3-year-old: (F(3, 775)=85.674, p<.005), 4-year-old: (F(3, 775)=88.664, p<.005),5-year-old: (F(3, 775)=23.699, p<.005)）, U nursery school (5-year-old: (F(3, 775)=5.025, p<.005), F kindergarten（3-year-old: (F(3, 775)=46.684, p<.005), 4-year-old: (F(3, 775)=10.811, p<.005), 5-year-old: (F(3, 775)=46.856, p<.005)), Y kindergarten (3-year-old: (F(3, 775)=8.073, p<.005), 4-year-old: (F(3, 775)=8.495, p<.005), 5-year-old: (F(3, 775)=21.574, p<.005)), N certified facility (3-year-old: (F(3, 775)=32.65, p<.005), 4-year-old: (F(3, 775)=70.468, p<.005), 5-year-old: (F(3, 775)=184.129, p<.005). As a result of multiple comparison, 3-year-old, 4-year-old and 5-year-old during the third phase in K nursery school, F kindergarten and N certified facility showed a significant difference.

Concerning the facility factor/ phase factor * facility factor * age factor, the simple main effect was statistically significant in the third phase (3-year-old: (F(4, 775)=30.04, p<.005, 4-year-old: (F(4, 775)=47.69, p<.005), 5-year-old: (F(4, 775)=52.901, p<.005). As a result of multiple comparison, 3-year-old and 4-year-old in K nursery school and 5-year-old in N certified facility showed a significant difference.

Concerning the age factor/ phase factor * facility factor * age factor, the simple main effect was statistically significant in the third phase (K nursery school: (F(2, 775)=32.451, p<.005), F kindergarten: (F(2, 775)=9.724, p<.005), N certified facility: (F(2, 775)=48.002, p<.005). As a result of multiple comparison, during the third phase, a significant difference was observed in 3-year-old and 4-year-old in K nursery school, 3-year-old and 5-year-old in F kindergarten, and 5-year-old in Y kindergarten with N certified facility. Figure 1 shows change of moving distance regarding the pelvis of 5-year-old by phase of MEB program.

As shown in Figure 1, the moving distance of pelvis significantly increased in the third phase of the five facilities. The purpose of activity in the third phase was mainly to encourage of the recognition of musical elements.

Table 2. Changes in the movement distance of the pelvis of KUFYN facilities by activity phase.

Phase	Facility	Age	Mean	SD	N
		3-year-old	0.81	0.385	17
	K nursery school	4-year-old	0.74	0.39	17
		5-year-old	0.82	0.445	20
		3-year-old	0.65	0.269	11
	U nursery school	4-year-old	0.63	0.136	8
		5-year-old	0.98	1.136	10
		3-year-old	0.28	0.072	18
	F kindergarten	4-year-old	0.37	0.275	14
		5-year-old	0.32	0.15	15
		3-year-old	0.31	0.248	16
	Y kindergarten	4-year-old	0.31	0.171	14
		5-year-old	0.34	0.158	15
		3-year-old	0.28	0.07	13
1st phase	N certified facility	4-year-old	0.27	0.133	12
		5-year-old	0.43	0.302	19
		3-year-old	0.54	0.301	8
	K nursery school	4-year-old	0.58	0.53	14
		5-year-old	0.56	0.202	17
		3-year-old	0.63	0.341	10
	U nursery school	4-year-old	0.59	0.347	9
		5-year-old	0.75	0.295	11
		3-year-old	0.34	0.109	13
	F kindergarten	4-year-old	0.46	0.822	14
		5-year-old	0.3	0.131	16
		3-year-old	0.25	0.096	15
	Y kindergarten	4-year-old	0.4	0.239	14
		5-year-old	0.42	0.312	14
		3-year-old	0.62	0.764	17
2nd phase	N certified facility	4-year-old	0.44	0.302	10
		5-year-old	0.63	0.572	18
		3-year-old	10.56	4.85	18
	K nursery school	4-year-old	10.85	5.239	17
		5-year-old	5.75	3.789	18
		3-year-old	3.11	2.17	11
	U nursery school	4-year-old	2.02	0.792	8
		5-year-old	3.98	2.137	9
3rd phase	F kindergarten	3-year-old	7.6	4.321	17
		4-year-old	4.46	1.573	12

Table 2. Contd

		5-year-old	7.6	4.321	17
		3-year-old	3.69	1.462	13
	Y kindergarten	4-year-old	3.88	2.066	13
		5-year-old	5.78	3.119	14
		3-year-old	6.74	3.454	16
	N certified facility	4-year-old	11.31	5.702	11
		5-year-old	13.68	6.761	20
		3-year-old	1.24	0.597	15
	K nursery school	4-year-old	1.78	0.722	12
		5-year-old	1.49	0.627	19
		3-year-old	1.68	1.004	9
	U nursery school	4-year-old	0.63	0.169	9
		5-year-old	0.89	0.319	10
		3-year-old	0.75	0.372	13
	F kindergarten	4-year-old	0.66	0.54	12
		5-year-old	0.85	0.686	16
		3-year-old	0.75	0.357	13
4th phase	Y kindergarten	4-year-old	0.81	0.465	14
		5-year-old	0.92	0.41	15
		3-year-old	1.01	0.516	15
	N certified facility	4-year-old	0.97	0.335	10
		5-year-old	1.1	0.584	20

Table 3. A main effect/interaction of the test.

Factor	*df*	*F*	Significance probability
phase	3	407.216	$p < .005$
facility	4	23.314	$p < .005$
age	2	1.607	*n.s.*
phase * facility	12	19.923	$p < .005$
phase * age	6	1.061	*n.s.*
facility * age	8	5.771	$p < .005$
phase * facility * age	24	5.678	$p < .005$

(2) Change of the moving average of acceleration of right hand

As a result of a three-way, non-repeated ANOVA regarding the moving average of acceleration of right hand, a statistically significance was observed in changes by age and activity phase. A main effect/ interaction of the test showed a statistically difference (phase: ($F(3, 775)=26.073$, $p<.005$), facility: ($F(4, 775)=60.989$, $p<.005$), age: ($F(2, 775)=11.264$, $p<.005$), facility * age: ($F(8, 775)=5.733$, $p<.005$)). The test regarding simple main effect with multiple comparison was carried out by Bonferroni method.

Concerning the phase factor/ phase * facility * age, the simple main effect showed a significant difference in K nursery school (3-year-old: ($F(3, 775)=6.537$, $p<.005$), 4-year-old: ($F(3, 775)=8.58$, $p<.005$)), and N certified facility (3-year-old: ($F(3, 775)=16.672$, $p<.005$), 4-year-old: (F(3,

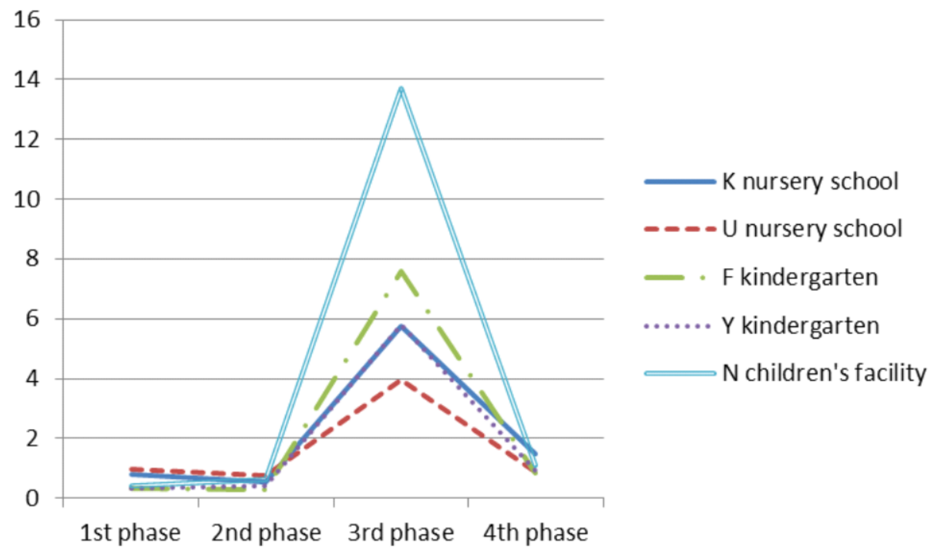

Figure 1. Change of moving distance regarding the pelvis of 5-year-old by phase of MEB program (m).

775)=22.705, *p*<.005), 5-year-old: (*F*(3, 775)=45.913, *p*<.005)). The test regarding simple main effect with multiple comparison was carried out by Bonferroni method. As a result of multiple comparison, 3-year-old, 4-year-old and 5-year-old in K nursery school showed a significant difference in the third phase of MEB program.

Concerning the facility factor/ phase * facility * age, the simple main effect showed a significant difference in the first phase (4-year-old: (*F*(4, 775)=4.672, *p*<.005), 5-year-old: (*F*(4, 775)=21.443, *p*<.005)), the second phase (5-year-old: (*F*(4, 775)=13.218, *p*<.005), and the third phase (3-year-old: (*F*(4, 775)=16.446, *p*<.005), 4-year-old: (*F*(4, 775)=28.353, *p*<.005), 5-year-old: (*F*(4, 775)=40.331, *p*<.005)). As a result of multiple comparison, a significant difference was observed in 4-year-old and 5-year-old in N certified facility during the first phase, 5-year-old in N certified facility during the second phase, and 3-year-old, 4-year-old and 5-year-old in N certified facility during the third phase.

Concerning the age factor/ phase * facility * age, the simple main effect showed a significant difference (the first phase in N certified facility: (*F*(2, 775)=16.131, *p*<.005), the second phase in N certified facility: (*F*(2, 775)=10.176, *p*<.005), the third phase in N certified facility: (*F*(2, 775)=8.05, *p*<.005)). As a result of multiple comparison, a significant difference was observed in 5-year-old in N certified facility from the first phase to the third phase. 3-year-old in U nursery school showed a significant difference in the fourth phase of MEB program.

Figure 2 shows change of moving average of acceleration regarding the right hand of 5-year-old by phase of MEB program. As shown in Figure 2, the moving average of acceleration tended to increase to the third phase, specifically N certified facility showed a large value.

(3) Change of the movement smoothness of right foot

As a result of a three way, non-repeated ANOVA regarding the movement smoothness of right foot, a main effect/ interaction of the test showed a statistically significant difference (phase: (*F*(3, 775)=4.568, *p*<.005), facility: (*F*(4, 775)=81.201, *p*<.005), phase * facility: (*F*(12, 775)=7.034, *p*<.005)). The test regarding simple main effect with multiple comparison was carried out by Bonferroni method.

Concerning the phase factor/ phase * facility * age, the simple main effect showed a significant difference.
In K nursery school (3-year-old: (*F*(3, 775)=4.785, *p*<.005), 4-year-old: (*F*(3, 775)=8.228, *p*<.005), 5-year-old: (*F*(3, 775)=8.174, *p*<.005), U nursery school (5-year-old: (*F*(3, 775)=8.939, *p*<.005)), and N certified facility (5-year-old: (*F*(3, 775)=4.45, *p*<.005). As a result of multiple comparison, 3-year-old and 5-year-old in the fourth phase and 4-year-old in the second with the fourth phase showed a statistically significant difference in K nursery school. 4-year-old and 5-year-old in U nursery school during the second phase, and 5-year-old in N certified facility during the third phase also showed a statistically significant difference.

Concerning the facility factor/ phase * facility * age, the simple main effect showed a significant difference in the first phase (3-year-old: (*F*(4, 775)=7.283, *p*<.005), 4-year-old: (*F*(4, 775)=4.255, *p*<.005)), the second phase (3-year-old: (*F*(4, 775)=9.801, *p*<.005), 4-year-old: (*F*(4, 775)=16.136, *p*<.005), 5-year-old: (*F*(4, 775)=19.466, *p*<.005)), and the fourth phase (3-year-old: (*F*(4, 775)=9.512, *p*<.005), 4-year-old: (*F*(4, 775)=12.327, *p*<.005), 5-year-old: (*F*(4, 775)=20.01, *p*<.005). As a result of multiple comparison, a statistically significant difference was observed in 3-year-old in U nursery school

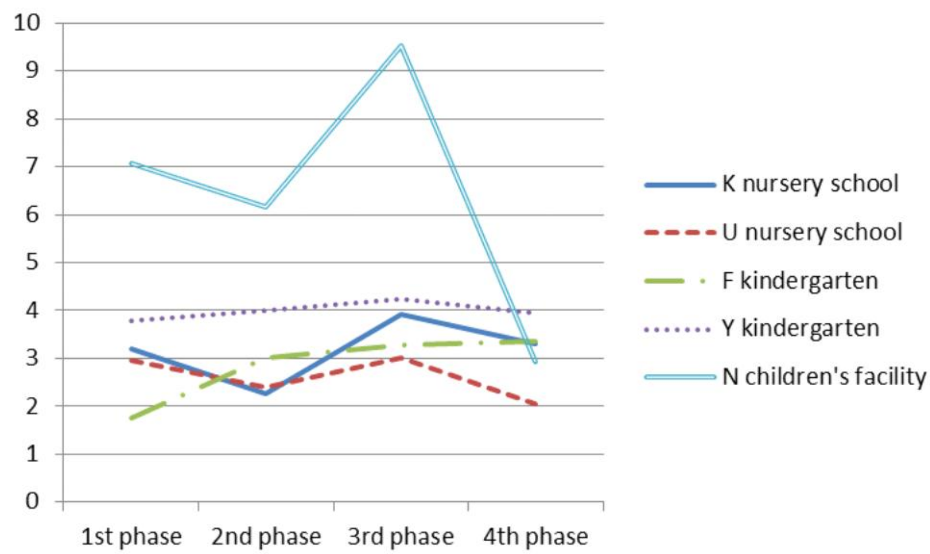

Figure 2. Change of moving average of acceleration regarding the right hand of 5-year-old by phase of MEB program (m/s^2).

and 4-year-old in K nursery school. In the second phase, 3-year-old, 4-year-old and 5-year-old in K & U nursery schools were significantly larger than F & Y kindergartens and N certified facility. The average data in K nursery school was remarkably large in the third and the fourth phase of MEB program.

In this way, based on the results of a three-way, non-repeated ANOVA on all the MVN measurement data, the author tried to classify the feature quantity by machine learning and discriminate the degree of musical development.

Machine learning method for classifying the developmental degree of musical expressions in early childhood

(1) Implementation of machine learning

Machine learning has been implemented in behavior recognition and a method of personal recognition in everyday life as learning support in recent studies (Kodama *et al*., 2015; Takada *et al*., 2012). Machine learning method is also used for motion learning support (Matsumoto *et al*., 2014) such as forward upward circling by elementary school students. Regarding musical expressions, the data classification of instrument performance (Young, 2008) and classification in which decision trees were used in analysis of breathing during music performance by machine learning methods (Igarashi *et al.*, 2001), learning of music performance studies (Widmer, 2001) are observed. Any research report has not been seen about machine learning related to musical expressions in early childhood yet. The author examined the possibility and practicality of the method as follows.

Early childhood educators usually judge musical development depended on body movement in musical expression of young children and harmoniously provide music experiences including musical elements such as strength of sound, duration, pitch and rhythm to young children. Early childhood educators with long educational experiences tend to share similarities to judgments of musical expressions of young children, even in intuitive judgment. Evaluation of musical development degree has no clear scoring rule (ballet posing etc.), but experienced educators generally agree on the evaluation results. In this study, each 3 educators at K & U nursery schools and 1 educators at F & Y kindergartens and 2 educators in N certified facility were involved in the judgement of children’s musical expression with naked eyes at the time of MVN measurement. Those educators in five facilities tended to make a similar classification of musical expression of participant children each other.

If it is possible to discriminate the developmental degree of musical expression including the body movement of young children by processing the feature quantities calculated from the data of 3D motion capture by machine learning, it will be able to classify and discriminate more objectively (clearly indicating sensitivity). If it is possible to improve the method of classification accuracy, it will be helpful to less experienced educators as well as judgement even in the absence of educators, so that they are frequently used to practice activities of musical expression of young children etc. It can be used for music education in early childhood.

Little research in the computational education domain has applied categorical statistics to its experimental outcomes. Therefore, the author referred research in other fields to enhance the classification sophistication. In

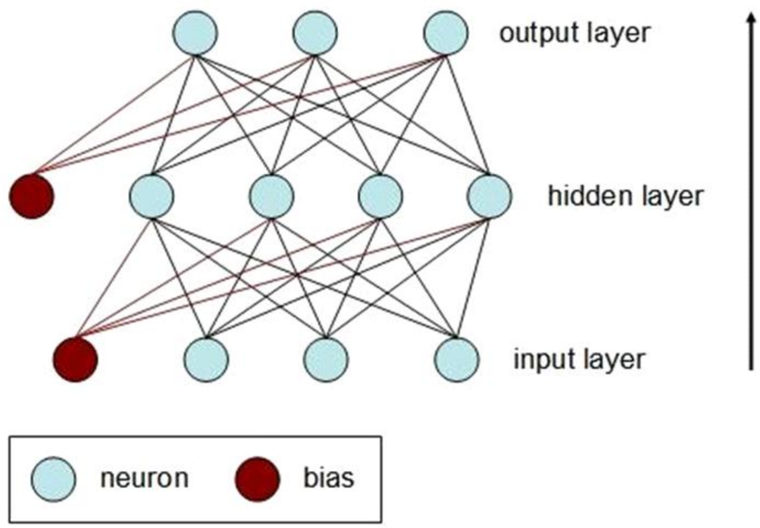

Figure 3. Multilayer Perceptron Neural Networks (Statistica).

particular, the author sets main technical goal is to optimize classification with consistent results. Categorization evaluation is based on the score of correctness of confusion matrix.

In this study, the author used five kinds of classifiers for machine learning method such as decision trees (Boosted Trees and Random Forest), Support Vector Machine (SVM), and neural networks (NN) as Multilayer perceptron (MLP) and Radial Basis Function (RBF). MVN measurement data were obtained from 3-year-old, 4-year-old and 5-year-old children in U (n=28) & K (n=48) nursery schools in December 2016 and in January 2017 (n=76 in total) in conjunction with their body motion video simultaneously recorded. From the MVN data, 13 feature quantities (regarding 3 body parts of pelvis, right hand and right foot, 3 kinetics such as the moving distance, the moving average velocity and the moving average acceleration and 1 smoothness measure, in addition, the moving distance between both hands) were calculated in Excel. At the same time, from simultaneously recorded video of body motion of children, the author also evaluated the developmental degree of musical expression of each child. Developmental degree of musical expression was classified into three levels of high, middle and low as video classification data (High: 15 children, Middle: 27, Low: 34) (Sano, 2018b).

(2)Classifiers

Boosted Trees is one of decision tree classifiers of which fitting methodology is enhanced by additive regression models sequentially fitting parametrized function. Gradient of loss function is generally set as minimization target to improve the process.

Random Forest is one of decision tree classifiers which consist of a collection (ensemble) of simple tree predictors. Such trees are used to vote for the most popular class (classification), or their responses are combined (averaged) to obtain an estimate of the dependent variable. Generally prediction accuracy is improved in comparison to simple trees.

Support Vector Machine is a classifier which performs classification tasks by constructing hyperplanes in a multidimensional space that separates cases of different class labels. Constructing an optimal hyperplane is by minimizing an error function by employing an iterative training algorithm.

Multilayer Perceptron is one of feedforward neural network architecture with unidirectional full connections between successive layers. Figure 3 shows Multilayer Perceptron Neural Networks. Neurons and biases are set in layers for feedforward topology where arbitrary complexity with the number of layers and containing units in each layer determining the function complexity.

Radial Basis Function is one of feedforward neural network architecture with unidirectional full connections between successive layers. Figure 4 shows Radial Basis Function Neural Networks. RBF generally employs two distinct stages with hidden layer neurons known as radial basis functions. RBF models the probability distribution of input data and how to relate an input data to classification target to minimize Euclidian distance calculated by radial basis units in hidden layer. RBF can be generally trained faster than MLP (Multilayer Perceptron) for large data sets and linear output activation functions.

(3) Training process of classifiers

Training of classification model was performed along with these 13 feature quantities (as factor) as input and three-

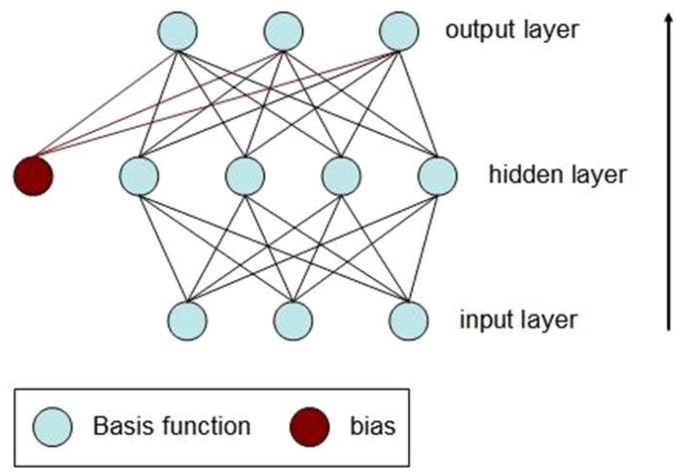

Figure 4. Radial Basis Function Neural Networks (Statistica).

Table 4. Confusion matrix of Boosted Trees.

		Prediction		
		High	**Medium**	**Low**
Actual	High	21	10	21
	Medium	14	8	20
	Low	18	6	10

level evaluation (as categorical dependent variable) as output (machine learning training targets) (n=76). As a result of training process, the classification accuracy based on confusion matrix for the training object was 66.7% for Boosted Trees, 54.0% for Random Forest and 51.3% for SVM. On the other hand, the classification training accuracy of MLP-NN was 64.8% and that of RBF-NN was 64.5%. Regarding training process, Boosted Trees and Neural Networks (MLP and RBF) showed fair fitting. Boosted Trees, MLP-NN and RBF-NN accuracy exceeded 60%.

(4) Classification prediction and validation

After the training process of classifiers completed, the developmental degree of musical expression of 128 children (F: n=41, Y: n=42, N: n=45) who were different from the training object (n=76) were forecasted with factor data and such forecasted results of categorical dependent (classification predictions) were compared with the developmental degree of musical expression of each child determined from the simultaneously recorded videos by the author (as video classification data, high: 52 children, middle: 42, low: 34).

(4)-1. Results of decision trees and SVM: Classification accuracy from confusion matrix of Boosted Trees as the decision tree was 30.5% (high: 21 children, middle: 8: low 10) as shown in table 4. From sensitivity analysis, top 3 contributing factors were the moving average acceleration of pelvis, the moving average velocity of pelvis and the moving average velocity of right foot. Classification accuracy of Random Forest was 25.0% (high: 0 children, middle: 0, low 32). Contributing factors were the moving average velocity of pelvis, the moving average acceleration of right hand, the movement smoothness of pelvis. For SVM, the classification accuracy was 26.6% (high: 0 children, middle: 0, low: 34).

Among these three classifiers, Boosted Trees showed superior categorization because Random Forest and SVM (Osuna *et al*., 1997) could categorize only one level, though, Boosted Trees and Random Forest showed consistent sensitivity focus on pelvis as shown in Figure 5. SVM is usually considered to work well when small numbers of data used. Those data sets might be too large for SVM. As in Table 4, confusion matrix of Boosted Trees showed balanced results which correctly categorized inputs into three levels.

(4)-2. Results of neural networks: Secondly, the author discusses neural networks which had comparatively better fitting in training process than classification trees or SVM. Regarding multilayer perceptron of neural network, classification accuracy was 33.6% (high: 2 children, middle: 26, low: 15) as shown in Table 5. Confusion matrix is shown below. High sensitivity contribution factors were the moving average acceleration of pelvis,

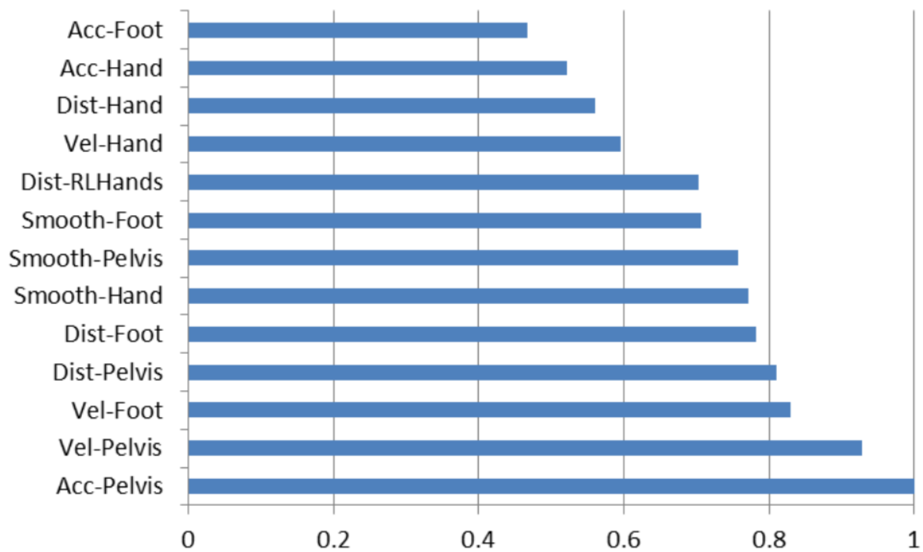

Figure 5. Factor sensitivity of Boosted Tree.

Table 5. Confusion matrix of MLP-NN.

		Prediction		
		High	**Medium**	**Low**
Actual	High	2	38	12
	Medium	1	26	15
	Low	1	18	15

the moving smoothness of right foot, the moving average acceleration of right foot as shown in figure 6. MLP sensitivity is high for pelvis movement which is consistent with classification trees such as Boosted Trees and Random Forest. MLP predicted in balanced manner, most for medium level and less for high or low level.

(4)-3. Results of Radical Basis Functions of neural network: For Radial Basis Functions of neural network (Memarian and Balasundram, 2012), classification accuracy was 46.1% (high: 42 children, middle: 11, low: 6) as shown in Table 6. Although accuracy value of RBF-NN is better than that of MLP-NN, forecast was heavily biased to "High" degree as shown in Figure 7. The contribution factor of sensitivity was the moving average acceleration of right foot, the moving average velocity of right hand, the moving average velocity of pelvis. Sensitivity analysis of RBF-NN did not show relevant results with any other classifiers.

DISCUSSION

The idea of exploiting statistical methods and devices to support educators of music would be anxiously awaited, but it could not be much examined practically due to its difficulty of extracting characteristics of body movement kinetics and modeling classifiers to fit captured observations. In the previous study, in order to develop methodology to support educators, the author combined statistical analysis of body movement and classroom test results of MEB (musical expression bringing up) program. Statistically significant characteristics of full-body motion backed by test results were adopted as feature quantity of series of classifiers and the author successfully presented potential of AI application to music education.

In this article, the author processed more data with certified facility to refine characteristics of feature quantities and introduced classifiers with enhanced architecture to improve categorization accuracy and consistency by validating from factor sensitivity. The data of 3-year-old, 4-year-old, and 5-year-old children in two nursery schools in 2016, two kindergartens in 2017, and a certified facility in 2018 was classified to discriminate the developmental degree of musical expression.

Firstly, results of a-three-way non-repeated ANOVA were shown. A statistically significant difference was observed in the moving distance of pelvis, the moving average acceleration of right hand and the movement smoothness of right foot. As a result, the moving distance of pelvis significantly increased in the third phase of the five facilities. The purpose of activity in the third phase of

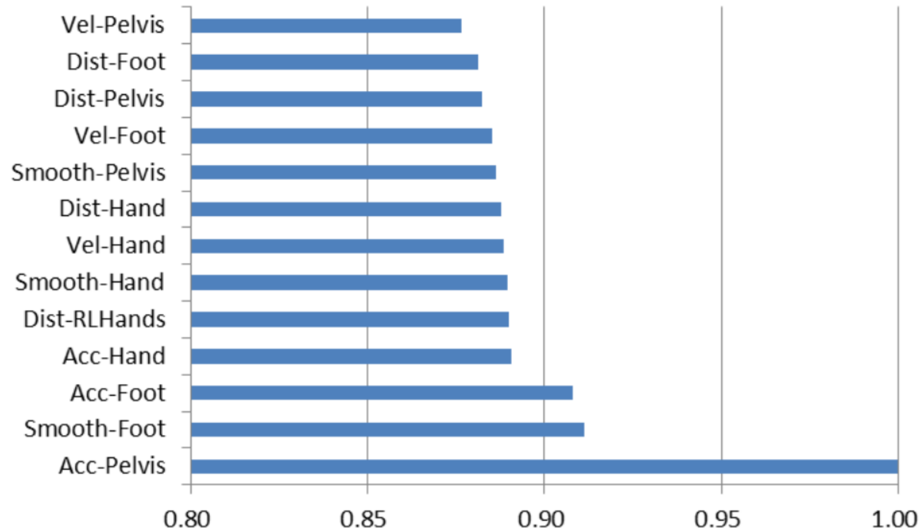

Figure 6. Factor sensitivity of MLP-NN.

Table 6. Confusion matrix of RBF-NN.

		Prediction		
		High	**Medium**	**Low**
	High	42	9	1
Actual	Medium	25	11	6
	Low	21	7	6

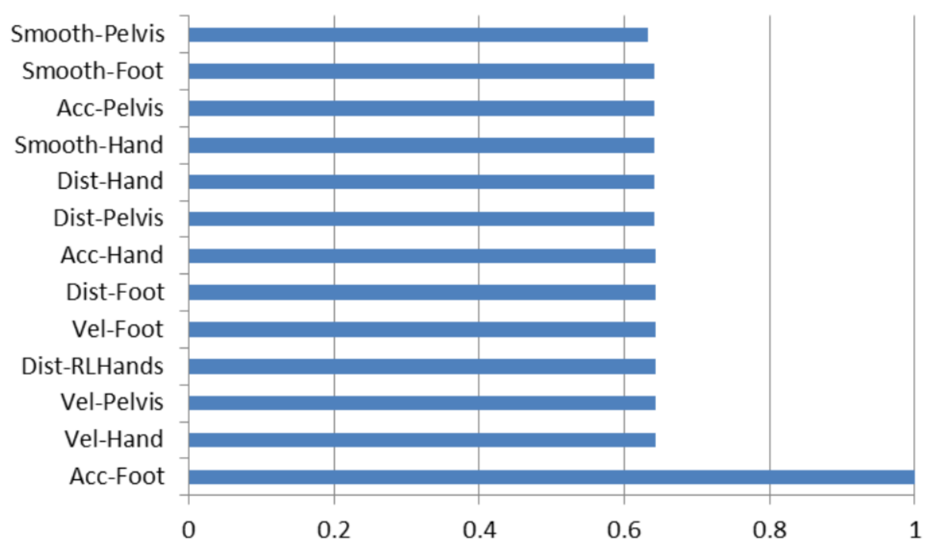

Figure 7. Factor sensitivity of RBF-NN.

MEB program was mainly to encourage of the recognition of musical elements. Increase of pretending movement to music and constant movement to take rhythm of music were observed in remarkable change of the moving distance of pelvis. Concerning change of the moving average of acceleration during the third phase, specifically a large value was observed in N certified facility. In general, the N certified facility, Y kindergarten and K nursery school taking Montessori method showed a large value in the third phase. This result also showed the difference in change due to difference in childcare forms. The children in K nursery school, Y kindergarten, and N certified facility showed more sensitivity by the regularity and contrast of music such as rhythm, strength of sound, duration of sound, and pitch at the activity of third phase. Furthermore, the change of movement smoothness in right foot tended to be divided into K & U nursery schools, and F & Y kindergartens, N certified facility, and K & U nursery schools were large. The data in F, Y and N facilities gradually increased to the third phase although the data in K & U nursery schools were large in the second phase of MEB program. The result showed that children in K & U nursery schools tended to express the recognition of musical elements by making a pretend movement of image of music, rather than taking a beat while regularity moving at a constant velocity at the third phase. The children in K nursery school expressed both pretend movement according to music showed by increase of the moving average of acceleration of right hand and the movement smoothness taking a beat by regular movement of constant velocity.

Secondly, when the degree of musical development obtained from videos was classified by machine learning using feature quantities of motion capture data, the author used five kinds of classifiers for machine learning method such as decision trees (Boosted Trees and Random Forest), Support Vector Machine (SVM), neural networks (NN) as Multilayer perceptron (MLP) and Radial Basis Function (RBF). Regarding training process of classifiers, Boosted Trees, MLP-NN and RBF-NN fitting exceeded 60%. As a result of classification prediction and validation, confusion matrix of Boosted Trees showed balanced results which correctly categorized inputs into three levels. High sensitivity contribution factors were the moving average acceleration of pelvis, the moving smoothness of right foot and the moving average acceleration of right foot. Regarding multilayer perceptron of neural network, classification accuracy was 33.6%. High sensitivity contribution factors were the moving average acceleration of pelvis, the moving smoothness of right foot, the moving average acceleration of right foot. Regarding Radial Basis Functions of neural network, confusion matrix score was actually high 46.1% but the prediction credibility was weakened due to its output imbalance and inconsistent profiles of sensitivity. Consequently, the author confirms Multilayer Perceptron Neural Network is the best classifier and Boosted Trees is the second best classifier.

CONCLUSION

In this study, the author inspected a method to achieve higher classification accuracy to predict musical development utilizing machine learning based on the feature quantity of body movement in musical expression in early childhood.

As a result of technical goal, it was concluded that the best model is Multilayer Perceptron. The second best is Boosted Trees. Boosted Trees showed inferior accuracy to that of MLP-NN but its distribution of number of predictions is similar to that of MLP-NN and sensitivity of those two classifications share highest contribution of acceleration of pelvis. The author demonstrated objective methodology to support educators of music.

In order to improve classification accuracy, classifiers such as deep learning etc. can be considered in addition to the further review of kinetic feature quantities, or increase of training sample data etc. In the future, such AI-based methodology can be an effective tool for educators with little experience to exploit standardized skill for evaluation.

ACKNOWLEDGEMENT

This work was supported by JSPS KAKENHI Grant Number 16K04579.

REFERENCES

Burger B (2013). Move the way you feel: Effects of musical features, perceived emotions, and personality on music-induced movement. Department of Music, University of Jyväskylä. http://urn.fi/URN:ISBN:978-951-39-5466-6.

Custodero L (1999). Construction of musical understandings: The cognitive- flow interface, Cognitive processes of children, England in musical activity Conference, p. 25. https://files.eric.ed.gov/fulltext/ED460061.pdf.

Dahl S, Friberg A (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, 24:433-454. 10.1525/mp.2007.24.5.433.

Hannon E, Johnson S (2005). Infants use meter to categorize rhythms and melodies: Implications for musical structure learning, *Cognit. Psychol.* 50:354-377.

Igarashi S, Ueno K, Furukawa K (2001). Analysis of Respiration during Musical Performances by Machine Learning, *SIG Technical Reports, MUS*, (82(2001-MUS-041)):27-32.

Jacque-Dalcroze E (1921). *Rhythm, Music and Education*, GP Putnam's Sons.

Kodama Y, Oba S, Ishii S (2015). Recognition of daily-life actions by machine learning methods, *IEICE technical report, IBISM*I, 114(502):73-78.

Matsumoto A, Mikami H, Kojima A (2014). Support for motor learning by clustering of sports forms: Examine effective image features as a target children back hip circle, *ITE technical report*, 38, 51(0):9-12. https://doi.org/10.11485/itetr.38.51.0_9.

Memarian H, Balasundram S (2012). Comparison between multi-layer perceptron and radial basis function networks for sediment load estimation in a tropical watershed, J. Water Resour. Protect. 4:870-876, doi:10.4236/jwarp.2012.410102.

Phiilips S, Trainor L (2007). Feeling the beat: Movement influences infant rhythm perception. *Science,* 308:1430. https://doi.org/10.1126/science.1110922.

Rubin J, Merrion (1996). *Drama and music methods*, Portsmouth: Linner Professional Publications.

Osuna E, Freund R, Girosi F (1997). Support Vector Machines: Training and Applications, *A.I. Memo* No. 1602, C.B.C.L Paper No. 144:1-43. https://dspace.mit.edu/handle/1721.1/7290.

Sano M (2013). Quantitative analysis about the educational effect of the music expression program, *The 9th Asia*-Pacific Symposium on Music Educ. Res. 39:1-7.

Sano M (2014). Contents of the music test concerning the recognition of musical elements in early childhood, *Research bulletin of Osaka*

Shoin Women's University, 4:67-74.

Sano M (2015). Characteristics of musical expressions showing the formation process of beat perception in early childhood: Through the practice of Musical Expression Bringing-up program for 5-year-old children in K nursery school, *Japanese J. Music Educ. Pract*. 12(2):120-131.

Sano M (2016a). Quantitative analysis concerning the movement elements of musical expression in early childhood utilizing motion capture. *Research Bulletin of Osaka Shoin Women's University* 6:133-143.

Sano M (2016b). Quantitative analysis of body movement in musical expression among three nursery schools in the different childcare forms utilizing 3D motion capture, *Information and Communication Technologies in the Music Field (ICTMF)*, Media Musica, Central and Eastern European Online Library, 7(2):7-18.

Sano M (2017). Analysis of elements of movement in musical expression in early childhood. *Japan J. Educ. Technol*. 41:5-8. https://doi.org/10.15077/jjet.S41003.

Sano M (2018a). Development of a quantitative methodology to analyze the growth of recognition of musical elements in early childhood from a viewpoint of change of body movement. *Asia-Pac. J. Res. Early Childhood Educ*. *(International)*, 12(1):61-80. http://dx.doi.org/10.17206/apjrece.2018.12.1.61.

Sano M (2018b). Statistical analysis of elements of movement in musical expression in early childhood using 3D motion capture and evaluation of musical development degrees through machine learning. *World J. Educ.* 8(3):118-130. https://doi.org/10.5430/wje.v8n3p118.

Sato K, Kaiga T, Watabe S (2010). The utilization of motion capture to support proficiency of the dancing. *Japan J. Educ. Technol*. 34:133-136. https://doi.org/10.15077/jjet.KJ00007086695.

Takada K, Kitasuka T, Aritsugi M (2012). A consideration of an individual identification method by gait using a marker less motion capture device, *IPSJSIG Technical Report*, 2012-ED-25, 9, 1-7.

Thelen E (1979). Rhythmical stereotypes in normal human infants". *Anim. Behav.* 27:699-715. https://doi.org/10.1016/0003-3472(79)90006-X.

Thompson M, Luck G (2012). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, 16:19-40. https://doi.org/10.1177/1029864911423457.

Widmer G (2001). Using AI and machine learning to study expressive music performance: project survey and first report. AI Communications, 14(3):149-162.

Young D (2008). Classification of Common violin Bowing Techniques Using Gesture Data from a Playable Measurement System. Proceedings of the 2008 conference on new interfaces for musical expression (NIME08), Genova, Italy.

Zentner M, Eerola T (2010). Rhythmic encouragement with music in infancy. PNAS, 107(13): 5768-5773. https://doi.org/10.1073/pnas.1000121107.