Transactions on Engineering and Computing Sciences - Vol. 11, No. 3
Publication Date: June 25, 2023
DOI:10.14738/tecs.113.14682.
Sano, M. (2023). An Exploration of Prediction Accuracy of Developmental Degrees of Music Expression in Early Childhood Based on A Simultaneous Analysis of Both Eye Movement and Full-Body Movement. Transactions on Engineering and Computing Sciences, 11(3). 10-25.

SERVICES FOR SCIENCE AND EDUCATION

An Exploration of Prediction Accuracy of Developmental Degrees of Music Expression in Early Childhood Based on A Simultaneous Analysis of Both Eye Movement and Full-Body Movement

Mina Sano
Tokoha University, Shizuoka, Japan

ABSTRACT

This article explores validity to introduce eye movement of children in association with body movement as feature quantities of machine learning, in particular, to assess children's maturity level of musical expression. There is a growing awareness of applying machine learning technique to yield practical benefit on predicting developmental degrees of music education by observing body movements with motion capture technology. From professional teaching perspective, it is believed only experienced teachers can evaluate children's music achievement because coordinated functioning of body movements needs carefully be observed in connection with various music factor such as rhythm, beat strength, tones, etc. If new teachers can take advantage of having objective results of evaluation of children's expression level by machines, it will enhance educational efficiency to certain levels which experts could attain. The author aims to improve classification accuracy of machine learning by focusing on eye movements data simultaneously recorded with body movements. In this study, children (n=43) at two child facilities participated in the data capture of both eye movement and body movement during musical expression. Feature quantities were extracted from the results of a three-way of ANOVA, and applied to machine learning to improve the classification accuracy of developmental degree in musical expression. Specifically, when using several classifiers such as NN (Neural Network model), the classification accuracy of developmental degree of musical development was more precisely in the case of both body movement data and eye movement data included in feature quantities than in the case of only body movement data as feature quantities.

Keywords: feature quantities, body movement in musical expression, eye movement, quantitative analysis, machine learning, classification accuracy.

INTRODUCTION

Body movement and eye movement influence the development of musical expression in early childhood. Children try to express using body movement not only their recognition of musical elements but also their own image when listening to the music. Children simultaneously think to judge looking around when they devise their own musical expressions. Therefore, it is necessary to quantitatively analyze the characteristic changes of both eye movement and body movement in musical expression in early childhood at the same time.

Previous researches have showed a relationship between the recognition of musical elements such as rhythms and body movement [1] but have paid little attention a quantitative analysis

of body movement in musical expression in early childhood using the motion capture technique. In educational field applying 3D motion capture, researches on the results of analysis regarding specific movement of traditional Japanese dance and sawing, mainly for adults, and the development of learning support methods can be seen [2,3]. Related researches, exploring relationships between pianists' body movements [4], adulthood reaction to characteristics of music were quantitatively analyzed using motion capture [5,6]. The author quantitatively analyzed body movement in musical expression in early childhood using MVN system as 3D motion capture and verified a close relationship between the recognition of musical elements and body movements [7]. Characteristic changes such as the increase of the moving average acceleration of hands and pelvis were observed as well as differences by age and facility according to the advance of recognition of musical elements in early childhood. The author captured the developmental characteristics of musical expression in early childhood utilizing the motion capture technique, and devised an evaluation method of developmental process by machine learning [8-10]. Previous researches in music education for children have little showed results of machine learning, but some results showed behavior recognition and individual recognition methods in everyday life [11,12], and learning support for elementary school students such as exercise [13].

In the related fields, researches of eye movement have mainly paid to the line of sight of the participant looking at the display. Those previous studies have showed eye movement regarding recognition of letters, and eye movement how affected the judgement of recognition function [14-21].

Regarding research using eye trackers in the field of music, eye movement during rhythm listening and music reading [22-24], and the gaze behavior of musical instrument players [25], eye-hand synchronization when playing musical instruments [26], differences in reading and listening to melodies between music professionals and non-musicians [27,28] and other research reports such as eye contact on the possibility of enhancing music conducting skills through a virtual reality learning environment are presented [29,30]. The author has quantitatively analyzed eye movements during musical expression when listening to the music in early childhood [31,32]. As a result, changes regarding numbers of saccade and total amplitude of saccade showed affection of recognition in tonality of music.

Regarding simultaneous analysis of movement and eye movement in musical expression, no research report has been observed in the field of music education in early childhood. For adulthood participants, synchronization of body movement with eye movements has been verified by quantitative analysis using motion capture technique such as specific skills, piano performance, body reaction to music, and a research of sign language with synchronized motion capture and eye tracking data [33-37]. Therefore, the author tried to conduct simultaneous analysis of body movement in musical expression with eye movement in early childhood and apply those feature quantities to machine learning to classify the developmental degree of musical expression in early childhood. If it is possible to discriminate the developmental degree of musical expression including the body movement and eye movement of young children by processing the feature quantities calculated from the data of 3D motion capture and eye tracking by machine learning, it will be able to classify and discriminate more precisely. If it is

possible to improve the method of classification accuracy, it will be helpful to less experienced educators to practice activities of musical expression for music education in early childhood.

PURPOSE OF THIS STUDY

This study aims to improve classification accuracy of machine learning by acquired data of body movement with eye movement simultaneously analyzed during musical expression in early childhood.

METHOD OF THIS STUDY

In this study, the author activates the device in order of analysis of movement and eye tracking, connects the MVN system as 3D motion capture to the Tobii Glasses 3 as eye tracking system with a cable by using TTL signals in order to synchronize both those systems.

MVN System as 3D Motion Capture

MVN system is a light weight and compact device which provides less constrained environment for even small sized children. MVN motion capture data include seventeen measurement points of body parts such as the pelvis, head, right shoulder, right hand, and right foot. The children's movement in musical expression were measured one by one during playing the piano accompaniment. In this system, each motion sensor is attached to each of 17 body measurement sites on the participant's child. A computer is activated, wirelessly communicates with the motion sensor, the participant child is calibrated by measurements of height and foot size, and analysis of body movement is carried out in real time in 60Hz by 1/60 second time frame.

Tobii Glass 3 as Eye tracking System

Tobii glass 3 is a glasses-type eye tracker of a wearable device. A participant child put on Tobii glass 3 like as glasses is calibrated while being fixed with a strap, and his eye movement is quantitatively analyzed on PC and audio and video are recorded in 50Hz by 1/50 second time frame, at the same time.

Figure1 shows a picture when a participant child put in motion trackers and eye tracking device for a simultaneous analysis of both eye movement and full-body movement in musical expression.



Figure 1. 4-year-old child put in motion trackers and eye tracking device

Numbers of Participant Children, Date and Time of Measurement

Table 1 shows the participants of simultaneous analysis of eye movement and body movement during musical expression in 2022 year. In Japan, young children's participation in surveys was still constrained by the ensuing COVID-19 in that year.

Table1. Participants in the simultaneous analysis of eye movement and body movement, measurement schedule and songs used for musical expression

	S child facility	T child facility
Measurement day	June, 14th 10:00～12:30 in 2022 year	June, 14th 13:30～15:30 in 2022 year
		June, 28th 13:30～15:30 in 2022 year
Participants	n=17	n=26
Songs	“Yamano onngakuka”	(Lyrics: Shisen Mizuta, German folk song)
	“Darumasan”	(Japanese children's song)

Quantitative Analysis of the Acquired Data by Synchronization of Analysis of Eye Movement and Analysis of Body Movement

The items of calculated data for eye movement are the number of occurrences of saccades, the moving average velocity of saccades, the average size of saccades, the total moving distance of saccades, the time until the first saccade occurs, the moving average velocity and magnitude when the first occurs saccade. Calculated data items for analysis of movement are the moving distance of the pelvis, head, right shoulder, right hand and right leg, the moving average velocity, the moving average acceleration, and smoothness of movement (the moving average of velocity/ the moving average of acceleration).

Concerning the acquired data of eye movement and body movement, a three-way ANOVA (non-repeated two standards as tonality, non-repeated two standards as facilities and non-repeated three standards as ages) was applied in order to find out statistically significant difference between relevant measures.

Applying Affective Feature Quantities to Machine Learning Method to Evaluate the Developmental Degree of Musical Expression

In this study, two educators at S certified facility and an educator at T certified facility were involved in the judgement of children's musical expression with naked eyes at the time of simultaneous analysis of MVN measurement and eye movement. Those educators in two facilities tended to make a similar classification of musical expression of participant children each other. Eye movement data and MVN measurement data were obtained from 3-year-old, 4-year-old and 5-year-old children in S (n=17) & T (n=26) early childhood facilities in June 2022 and in July 2022 (n=43 in total) in conjunction with their body motion video simultaneously recorded. From the data captured by MVN and Tobii glass 3, 14 feature quantities (12 regarding 3 body parts of pelvis, right hand and right foot, for 4 kinetics such as the moving distance, the moving average velocity and the moving average acceleration and smoothness measure) and 2 eye tracking data of total amplitude of saccades and number of saccades 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 by observing video (High: 22 children, Middle: 13children, Low: 8children). Of the 43 cases observed in June and July 2022, the author used 31 cases to train machine learning models in order to discriminate 12 cases into three levels of developmental degrees. The author compared the classifying accuracy of developmental degree with two sets of feature quantities, set A as derived from eye movement data and body movement data against set B as derived from body movement data only. Such comparison were attempted for several machine learning models when the participant children singing "Yamano onngakuka" as a song in major key.

RESULTS

Characteristic Results of a Three-Way Non-Repeated Analysis

In this study, the early childhood children participated in a simultaneous analysis of body movement in musical expression with eye movement. The author extracted feature quantities from acquired data based on results of a three-way non-repeated analysis.

Results of A Three-Way Non-Repeated Analysis Based on Eye Tracking Data During Musical Expression:

Table 2 shows some characteristic data of the average value of measurement items by major key / minor key and child facility. Those data were quantitatively analyzed by a three-way ANOVA (non-repeated factors as 2 levels in major key / minor key, non-repeated factors as 2 levels in S child facility, T child facility, non-repeated factors as 3 levels in 3-year-old, 4-year-old, 5-year-old).

Table 2. The average data of eye tracking of measurement items by major key / minor key and child facility

Facility	Songs	Age	the number of occurrences of saccades	The moving average of velocity of saccade(degree/second)	The average size of saccade (degree)	The total moving distance of saccade (degree)
		3-year-old	27.6	211.934	6.982	191.104
	Yamano	SD	5.31977	21.44636	1.60635	47.68163
	Ongakuka	4-year-old	23.7143	210.6757	6.3843	155.0757
	(n=17)	SD	11.22073	30.08226	1.4734	80.75251
		5-year-old	19.6	226.49	7.358	153.838
S facility		SD	14.25833	25.25016	1.50463	125.7933
		3-year-old	15.8	191.666	6.112	130.912
	Darumasan	SD	18.08867	41.18474	2.76165	170.8425
	(n=17)	4-year-old	14.2857	197.7014	5.2386	85.8729
		SD	14.37259	32.0419	2.19426	87.16182
		5-year-old	16.4	251.634	9.552	120.134
		SD	10.01499	99.43071	6.10761	60.27015
		3-year-old	24.7778	199.8633	6.97	176.1522
	Yamano	SD	10.70955	30.9301	1.43149	86.64843
	Ongakuka	4-year-old	37.8	215.935	7.659	287.904
	(n=26)	SD	14.6424	31.29545	1.44209	114.695
		5-year-old	41.4286	229.3471	8.1343	344.0186
T facility		SD	12.3269	28.04542	1.49652	142.7799
		3-year-old	16.3333	204.8411	6.3367	109.9667
	Darumasan	SD	5.14782	44.77783	2.26418	68.89343
	(n=26)	4-year-old	21.7	222.083	8.176	185.668
		SD	7.63108	47.10029	2.60146	95.19717
		5-year-old	26.8571	246.3814	9.4543	252.4714
		SD	16.66762	20.82267	0.78283	155.315

In this investigation, a statistically significance was characteristic in the numbers of occurrences of saccade and total moving distance of saccade (total amplitude of saccades) as following description showed.

As a result of three-way ANOVA for these acquired data regarding numbers of occurrences of saccade, a main effect between subjects of the test showed a statistically significant difference (child facility factor ($F(1,74) = 10.137$, $p < .005$), songs factor $F(1,74) = 15.435$, $p < .005$). As a result of multiple comparison test by Bonferroni's method, 4-year-old and 5-year-old children significantly showed in T facility larger than S facility when singing “Yamano onngakuka”. In T facility, 4-year-old and 5-year-old children significantly showed the number of occurrences of saccade more when singing “Yamano ongakuka” than “Darumasan”.

As a result of three-way ANOVA for these acquired data regarding total moving distance of saccade, a main effect between subjects of the test showed a statistically significant difference (child facility factor (F (1,74) =13.034, p< .005), songs factor F (1,74) =8.653, p<.005). As a result of multiple comparison test by Bonferroni's method, 4-year-old and 5-year-old children when singing “Yamano ongakuka” significantly showed larger distance in T facility than S facility. 4-year-old children in T facility significantly showed larger distances when singing “Yamano ongakuka” than “Darumasa”.

Results of Quantitative Analysis Regarding Body Movement in Musical Expression:
A statistically significance was observed in 11 items such as the moving average acceleration and the moving average distance regarding full-body measurements. In this paper, some of results regarding pelvis and right hand were showed as table 3.

The Characteristics of Movement of Pelvis:
As a result of three-way ANOVA for these acquired data regarding the moving distance of pelvis, a main effect between subjects of the test showed a statistically significant difference (songs factor F(1,74) =36.692, p<.005), child facility factor (F(1,74) =17.3, p< .005)). As a result of multiple comparison test by Bonferroni's method, 3-year-old children in S facility significantly showed large distance, 3-year-old, 4-year-old, and 5-year-old children in T facility significantly larger distance in major key than minor key. 3-year-old and 4-year-old children when singing “Darumasan” in minor key significantly showed larger distances in S facility than T facility.

As a result of three-way ANOVA for these acquired data of the moving average acceleration of Pelvis, a main effect between subjects of the test showed a statistically significant difference (child facility factor (F (1, 74)=14.083, p<.005)). As a result of multiple comparison test by Bonferroni's method, 3-year-old and 4-year-old children significantly showed when singing in minor key larger acceleration in S facility than T facility.

The Characteristics of Movement of Right Hand:
As a result of three-way ANOVA for these acquired data regarding the moving distance of right hand, a main effect between subjects of the test showed a statistically significant difference (songs factor (F(1,74) =191.218, p <.005), child facility factor (F(1,74) =71.698, p < .005)). As a result of multiple comparison test by Bonferroni's method, 3-year-old, 4-year-old, and 5-year-old children significantly showed larger distances when singing in major key than minor key. 4-year-old and 5-year-old children when singing in major key significantly showed larger distances in S facility than T facility as well as 3-year-old and 4-year-old children when singing in minor key. 4-year-old and 5-year-old children significantly showed larger distances of right hand when singing in major key than 3-year-old children in S facility. Those participant children more largely moved their right hands when singing “Yamano ongakuka” as a song in major key than “Darumasan” as a song in minor key in both S and T facilities. The children used their right hands to express images related to the content of the lyrics, as if anthropomorphic animals were playing musical instruments.

Table 3 The average value of measurement items as a result of analysis of movement in musical expression

Song tonality	Facility	Age	The moving distance of pelvis	The moving distance of head	The moving distance of right hand	The moving distance of right foot	The moving average acceleration of pelvis	The moving average acceleration of head	The moving average acceleration of right hand	The moving average acceleration of right foot	N
Yamano ongakuka major key	S facility	3-year-old ($\bar{x}$)	0.974	1.353	7.4029	0.3881	0.6503	0.5509	3.9223	0.4552	5
		SD	0.1015	0.27634	2.77601	0.17486	0.18391	0.14724	1.77958	0.46811	
		4-year-old ($\bar{x}$)	0.9177	1.2117	10.2835	0.4373	0.8458	0.6784	4.9877	0.2524	7
		SD	0.53965	0.44181	2.10113	0.42909	0.69716	0.4952	1.13605	0.25014	
		5-year-old ($\bar{x}$)	0.7656	1.24	10.2761	0.122	0.3904	0.422	5.1504	0.0484	5
		SD	0.14269	0.20106	1.89949	0.04068	0.16115	0.05278	1.23255	0.00862	
	T facility	3-year-old ($\bar{x}$)	0.7471	1.1628	6.1839	0.374	0.6344	0.322	3.1869	0.1764	9
		SD	0.32676	0.67008	2.38925	0.43287	0.42842	0.04964	1.26675	0.16238	
		4-year-old ($\bar{x}$)	0.7351	0.989	5.7722	0.1524	0.4692	0.556	3.5568	0.1031	10
		SD	0.23623	0.17953	1.17352	0.06264	0.20035	0.53734	0.90464	0.07672	
		5-year-old ($\bar{x}$)	0.5957	0.9214	4.5899	0.231	0.4587	0.5562	2.4245	0.2029	7
		SD	0.22754	0.40378	2.63442	0.32263	0.14625	0.62916	1.49881	0.32852	
Darumasan minor key	S facility	3-year-old ($\bar{x}$)	0.5615	0.7175	3.4652	0.1376	0.7962	0.4926	3.638	0.1583	5
		SD	0.08023	0.03622	0.95983	0.10644	0.2491	0.17251	0.78855	0.17699	
		4-year-old ($\bar{x}$)	0.6608	0.8041	4.3583	0.4199	1.1757	0.6596	4.6816	0.5535	7
		SD	0.44148	0.41806	1.13466	0.4595	1.00776	0.56095	1.0751	0.87638	
		5-year-old ($\bar{x}$)	0.5135	0.6372	2.7154	0.0934	0.8462	0.2936	2.8501	0.07	5
		SD	0.23624	0.21332	1.74577	0.04059	0.68142	0.10923	2.16243	0.06127	
	T facility	3-year-old ($\bar{x}$)	0.2474	0.2854	0.5889	0.1448	0.2553	0.2024	0.6898	0.2084	9
		SD	0.13953	0.09708	0.38291	0.34858	0.22652	0.10853	0.41109	0.50432	
		4-year-old ($\bar{x}$)	0.3592	0.4445	1.2751	0.0838	0.3382	0.5203	1.3335	0.079	10
		SD	0.14629	0.11766	0.72961	0.0683	0.20141	0.63917	0.7298	0.09334	
		5-year-old ($\bar{x}$)	0.2089	0.3379	1.1144	0.2612	0.3417	0.2552	1.0578	0.7627	7
		SD	0.09896	0.18024	1.01749	0.3882	0.17006	0.07999	0.70504	1.19217	

As a result of three-way ANOVA for these acquired data regarding of the moving average acceleration of right hand, a main effect between subjects of the test showed a statistically significant difference (song factor ($F(1, 74)=34.131$, $p<.005$), child facility factor ($F(1, 74)=71.355$, $p<.005$)). As a result of multiple comparison test by Bonferroni's method, 5-year-old children in S facility, and 3-year-old, 4-year-old, and 5-year-old children in T facility when singing in major key significantly showed larger acceleration than minor key. 4-year-old and 5-year-old children when singing in major key and 3-year-old, 4-year-old and 5-year-old children when singing in minor key significantly showed larger acceleration in S facility than T facility. 4-year-old children when singing in minor key significantly showed larger acceleration than 5-year-old children in S facility.

Characteristics of Movement Both Eye and Full-Body by Simultaneous Analysis in Musical Expression:

Concerning eye movement when children presented musical expression, the numbers of occurrences of saccade and the total moving distance of saccade were characteristic. A result of

three-way ANOVA showed a tendency for the total moving distance of the saccade to be larger for “Yamano ongakuka'” in major key than for “Darumasan” in minor key.

In the simultaneous analysis of eye movement with full-body movement, the moving distance of right hand and the moving average acceleration as feature quantities of movement were more contributed to development of musical expression than other feature quantities of movement. As a result of quantitative analysis regarding the moving average acceleration of right hand, 5-year-old children in S facility and 3-year-old, 4-year-old, and 5-year-old children in T facility significantly showed larger moving acceleration when singing in major key than minor key.

Machine Learning Method

In this study, the author used several kinds of classifiers for machine learning method such as Support Vector Machine (SVM), k-nearest neighbors (KNN) and neural networks (NN) as Multilayer perceptron (MLP). 31 cases were used as training input of classifiers. 12 cases were used for output to evaluate classification accuracy with regard to respective feature quantities. Developmental degree of musical expression classified into three levels of high, middle and low by way of video observation for 31 cases are (High: 18 (58%), Medium: 9 (29%), Low: 4 (13%)) and for 12 cases (High: 4 (33%), Medium: 4 (33%), Low: 4 (33%)) when participant children singing “Yamano ongakuka” in major key.

Classifiers:

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.

K-nearest neighbors is relatively simple and fundamental method to classify outcome based on distance metrics between learned cases and new query case. Technically it reaches prediction results by comparing Euclidean distance from nearest case to gradually remote specified cases.

Multilayer Perceptron is one of feedforward neural network architecture with unidirectional full connections between successive layers. 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.

Training process of Classifiers:

Training of classification model was performed along with these feature quantities (as factor) as input and three-level evaluation (as categorical dependent variable) as output (machine learning training targets) (n=31). As a result of training process with body movement data (without eye movement data), the classification accuracy based on confusion matrix for the training object was 61% for SVM. KNN showed 50% accuracy. On the other hand, the classification training accuracy of MLP-NN was 74 %. Regarding training process, Neural Networks (MLP) showed best fitting.

In the same manner 31 cases with body movement data and eye movement data were applied as training data for classifiers. SVM showed slight better 64% accuracy, KNN the same 50% and MLP-NN slight better 78%.

Classification Prediction and Validation:

After the training process of classifiers completed, the developmental degree of musical expression of 12 children who were different from the training object (n=31) were forecasted with factor data. Those 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 (High: 4 (33%), Medium: 4 (33%), Low: 4 (33%)). In this paper, the classification accuracy was compared whether the feature quantities included both eye movement data and body movement data or only body movement data.

Results of SVM:

For SVM as classifiers, the result of machine learning using feature quantities of body movement showed 33.3% as classification accuracy shown as table 4-1. The result of machine learning using feature quantities both body movement and eye movement in simultaneous analysis showed 50% as classification accuracy shown in table 4-2 that was more improved than only body movement as feature quantities.

Table 4-1. Body movement analysis

	Predicted		
Actual	High	Medium	Low
High	100%	0%	0%
Medium	100%	0%	0%
Low	75%	25%	0%

Overall Accuracy=33.3%.

Table 4-2. Both body movement and eye movement in simultaneous analysis

	Predicted		
Actual	High	Medium	Low
High	100%	0%	0%
Medium	75%	25%	0%
Low	50%	25%	25%

Overall Accuracy=50%

Results of KNN:

KNN as classifier showed similar tendency as SVM. KNN had better accuracy with eye movement data than that with only body movement data. Accuracy distribution in confusion matrix is slightly better than SVM because KNN predicted “High” level better than SVM.

Table 4-3. Both body movement only analysis

	Predicted		
Actual	High	Medium	Low
High	75%	0%	25%
Medium	0%	50%	50%
Low	50%	0%	50%

Overall Accuracy=58%

Table 4-4. Both body movement and eye movement in simultaneous analysis

	Predicted		
Actual	High	Medium	Low
High	75%	0%	25%
Medium	0%	50%	50%
Low	0%	0%	100%

Overall Accuracy=75%

Results of MLP-NN:

For MLP-NN as classifiers, the result of machine learning using feature quantities of body movement showed 42% as classification accuracy shown in table 4-5. The result of machine learning using feature quantities both body movement and eye movement in simultaneous analysis showed 58 % as classification accuracy shown in table 4-6 that was more improved than only body movement as feature quantities.

Applying both body movement and eye movement data. MLP-NN showed accuracy 58% which is inferior to that of KNN 75%. However, MLP-NN has superior prediction distribution in confusion matrix. MLP-NN never predicted “actual Low” as “High” or “actual High” as “Low”. MLP-NN indicates robust consistency for evaluation spectrum.

Table 4-5. Body movement analysis

	Predicted		
Actual	High	Medium	Low
High	50%	0%	50%
Medium	75%	25%	0%
Low	50%	0%	50%

Overall Accuracy=42%

Table 4-6. Both body movement and eye movement in simultaneous analysis

	Predicted		
Actual	High	Medium	Low
High	50%	50%	0%
Medium	25%	25%	50%
Low	0%	0%	100%

Overall Accuracy=58%

In figure 2 sensitivity of MLP-NN is depicted, which informs relative importance of variables utilized for classifier inputs. Most influential factor was body movement of right foot smoothness (the moving average of velocity/ the moving average of acceleration).

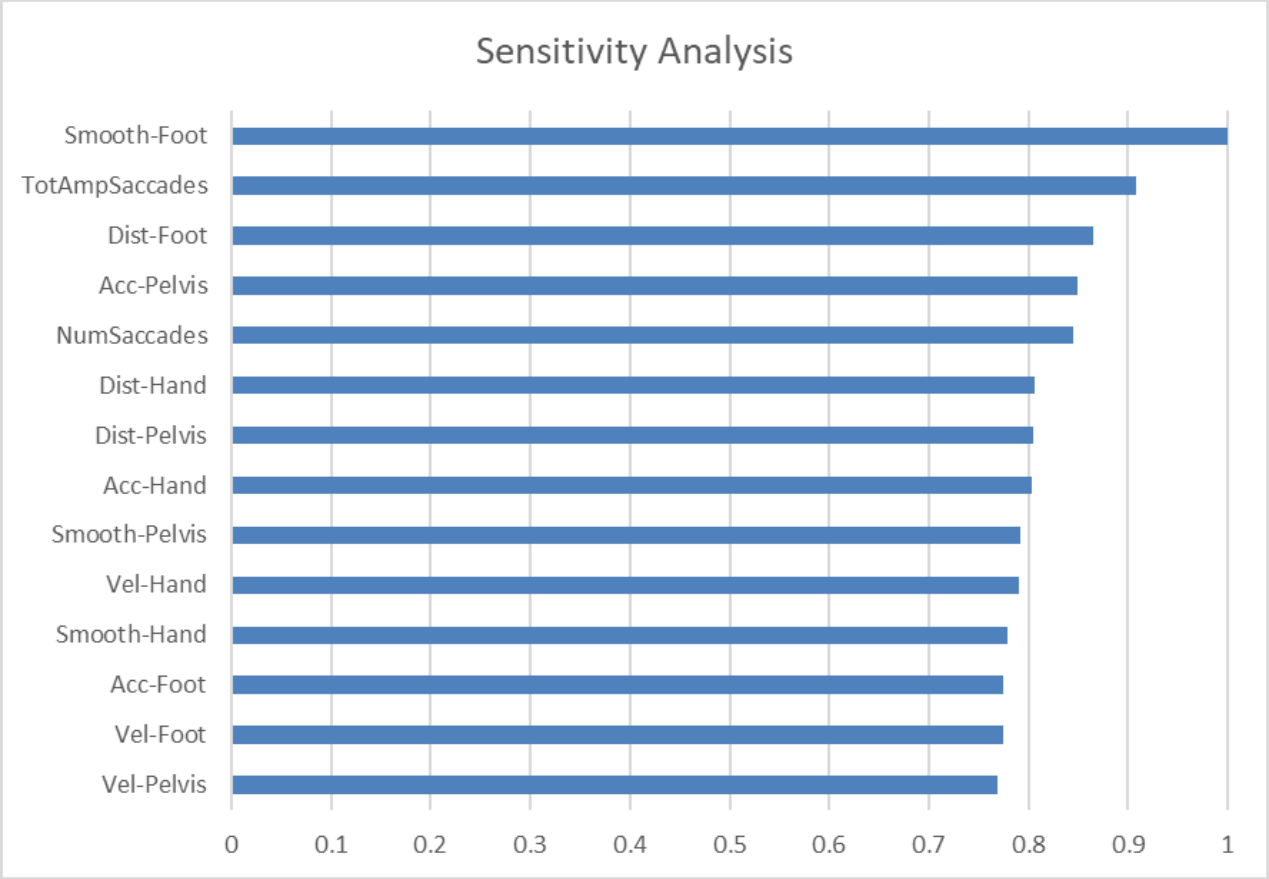


Figure 2. Sensitivity of Analysis

DISCUSSION

In this article, the author tried to improve classification accuracy of machine learning by feature quantities based on acquired data of both body movement and eye movement simultaneously analyzed during musical expression in early childhood.

Firstly, more effective feature quantities were extracted from the result of a three-way of ANOVA based on acquired data in simultaneous analysis of body movement and eye movement. Concerning the results of a three-way non-repeated analysis based on eye tracking data during musical expression, a statistically significance was more remarkably observed in numbers of occurrences of saccade and the total moving distance of saccade when singing a song in major key than minor key. Specifically for the results of a three-way non-repeated analysis based on MVN data in musical expression, a statistically significance was more remarkably observed in the moving distance and the moving average acceleration regarding pelvis and right hand than other body measurement points.

Secondly, the author conducted machine learning using feature quantities extracted from the results of quantitative analysis in which a statistically significance was remarkable in measurement data. The classification accuracy when participant children singing "Yamano ongakuka" in major key was compared between the case of only body movement data as feature quantities and the case of both body movement data and eye movement data. As a result, when using several classifiers such as NN (Neural Network model), the classification accuracy of developmental degree of musical development was more precisely in the case of both body movement data and eye movement data included in feature quantities than in the case of only body movement data as feature quantities. For MLP-NN as classifiers, 42% as the classification accuracy when using only body movement data as feature quantities improved to 58% as the classification accuracy when using both body movement data and eye movement data as feature quantities.

Those results showed a significance of more effective feature quantities extracted from body movement data and eye movement data during musical expression in early childhood as shown figure 2. Moreover, classification and prediction of developmental degree using the machine learning method would be helpful for music education support when each child needs what kind of music experience at individual development.

CONCLUSION

This study verified the increase of more effective feature quantities contributed to improve the classification accuracy in developmental degree of musical expression. In my previous studies regarding the developmental degree of musical expression by machine learning method, feature quantities were constrained to body movement in musical expression (Sano 2020; Sano 2019). A result of quantitative analysis in this study verified an interaction among recognition of musical elements such as rhythm and sound awareness in early childhood child and body movement, and eye movement.

A future issue is to improve the classification accuracy of developmental degree in musical expression by increasing the number of simultaneous analysis data of both body movement and eye movement during musical expression in early childhood. Furthermore, the author thinks it is necessary to acquire more accurate data and explore new feature quantities regarding finger movement for the hand movements in musical expression that were characteristic in this investigation.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 21K02369.

References

[1]. Jaques-Dalcroze, E., *Rhythm, Music and Education*, 1921, G. P. Putnam's Sons.

[2]. Ando, A. and Sumikawa, Y., *Development and functional evaluation of teaching materials for observing sawing motion using motion capture and virtual space*, Journal of the Japan Society for Educational Technology, 2012, 36 (2): p.103-110. https://doi.org/10.15077/jjet.KJ00008274459

[3]. Sato, K., Kaiga, T., and Watabe, S., *Utilization of motion capture to support the mastery of dance,* Journal of the Japan Society for Educational Technology, 2010, 34: p.133-136.

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

[5]. Burger, B., *Move the way you feel: Effects of musical features, perceived emotions, and personality on music-induced movement.* 2013, Department of Music, University of Jyväskylä.

[6]. Burger, B., Thompson, M., Saarikallio, S., Luck, G., & Toiviainen, P., *Influences of rhythm and timbre related musical features on characteristics of music induced movement*, Frontiers in Psychology, 2013, 4 (183): p.1-9.

[7]. Sano, M., *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 Journal of Education, 2018, 8 (3): p.118-130.

[8]. Fernandez-Delgado, M., Cernadas, E., Barro, S., Amorim, D., *Do we need hundreds of classifiers to solve real world classification problems?* Journal of Machine Learning Research, 2014, 15(90): p.3133–3181.

[9]. Sano, M., *Predicting developmental degrees of music expression in early childhood by machine learning classifiers with 3D motion captured body movement data*, Journal of Educational Research and Reviews (International),2019, 7(7): p.155-168.

[10]. Sano, M., *Verification of a classification prediction method for the development of musical expression in early childhood using a machine learning method based on 3D motion capture data*, Advances in Social Sciences Research Journal (International), 2020, 7(9): p.338 -358.

[11]. Kodama, U., Oba, N., and Ishi, N., *Examination of behavior recognition method in general living environment using machine learning*, IEICE technical report, 2015, 114(502): p.73-78.

[12]. Takada, K., Kitasuka, T., and Aritugi, M., *Examination of personal identification method by gait using markerless motion capture device,* Research Report Entertainment Computing（EC）2012: p.1-7.

[13]. Matumoto, A., Mikami, H., Kawamura, H., and Kojima, A., *Examination of form classification method for supporting motor learning of elementary school students who performed motion analysis using image processing and machine learning techniques*, Video Information Media Society Technical Report, 2014, 38, 51(0): p.9-12.

[14]. Higuchi, H., Okumura, Y., and Kobayashi, T., *Development of toddler's hiragana-sound knowledge: An eye-tracking study*, IEICE technical report, 2019, 118（437）: p.83-88.

[15]. Watabe, M., Matumura, A., and Uda, N., *An attempt to extract children's tastes from gaze movements and facial expressions in storytelling of picture books*, Proceedings of the 81st National Convention, 2019, (1): p.563-564.

[16]. Kusunoki, K., Ozawa, W., and Kanemori, Y., *Evaluation of Immigrant Children's Impediments in Reading Japanese*: Comparative Analysis on Reading Horizontal and Vertical Writing Texts, 2017, 6(0): p.7-8.

[17]. Mpofu,B., *University Students Use of Computers and Mobile Devices for Learning and their Reading Speed on Different Platforms*, Universal Journal of Educational Research, 2016, 4(4): p.926-932. https://doi.org/10.13189/ujer.2016.040430

[18]. Seong-un Kim, Sung-man Lim, Eun-ae Kim, Il-ho Yang, *An Analysis of Eye Movement and Cognitive Load about the Editorial Design in Elementary Science Textbook*, Universal Journal of Educational Research, 2016, 4(3): 570-575, https://doi.org/10.13189/ujer.2016.040314

[19]. Moreno-Estevaa, E., White, S., Woodc,J., and Blackc, A., *Application of mathematical and machine learning techniques to analyze eye tracking data enabling better understanding of children's visual cognitive behaviours*, Frontline Learning Research, 2019, 6(3): Special issue: p.72 – 84. https://doi.org/10.14786/flr.v6i3.365

[20]. Remount,F., and Budke,A., *Strategies for Successful Learning with Geographical Comics: An Eye-Tracking Study with Young Learners,* Education Sciences, 2020, 10: p.293, doi:10.3390/educsci10100293

[21]. Valtakari,N., Hooge1,I., & Viktorsson,C., & Nyström,P., Falck-Ytter,T., and Hessels, R., *Eye tracking in human interaction: Possibilities and limitations,* Behavior Research Methods, 2021, 53: p.1592–1608.https://doi.org/10.3758/s13428-020-01517-x/ Published online: 6 January 2021

[22]. Plöchl, M., Obleser, J., *Do auditory rhythms influence eye movement statistics?* Conference on Music & Eye-Tracking, 2017, Max Planck Institute for Empirical Aesthetics, p.19.

[23]. Lörch, L., Fehringer, B., Münzer, S., *Reading music. How tonality and notation influence music reading experts' eye movements, pupil dilation and performance in a pattern matching task.* Conference on Music & Eye-Tracking, 2017, Max Planck Institute for Empirical Aesthetics, p.17.

[24]. Fusase, E., *In "Childcare Expression Technology" Music I Consideration of specialized skills-line of sight when playing the piano: Through the difference in behavioral proficiency*, Nagoya Yanagishiro Junior College Lab Bulletin of Research, 2017, (39): p.313-331.

[25]. Vandemoortele, S., Feyaerts,K., Reybrouck, M., De Bièvre, G., Brône, G., & De Baets, T., *Gazing at the partner in musical trios:amobileeye-trackingstudy*. Journal of Eye Movement Research, 2018, 11(2):p.6. DOI: 10.16910/jemr.11.2.6

[26]. Marandola, F., *Eye-Hand synchronisation and interpersonal interaction in xylophone performance: A comparison between African and Western percussionists*, Conference on Music & Eye-Tracking, 2017, Max Planck Institute for Empirical Aesthetics, p.18.

[27]. Drai-Zerbib, V. & Baccino, T., *Cross-modal music integration in expert memory: Evidence from eye-movements*. Journal of Eye Movement Research, 2018, 11(2):4, p.1-21. DOI: 10.16910/jemr.11.2.4

[28]. Puurtinen, M., *Learning on the job: Rethinks and realizations about eye tracking in music-reading studies*, Frontline Learning Research, 2018, 6 (3): p.148- 161.

[29]. Orman, K., Price, H., Russell, R., *Feasibility of Using an Augmented Immersive Virtual Reality Learning Environment to Enhance Music Conducting Skills*, Journal of Music Teacher Education, 2017, 27 (1): p.24-35.

[30]. Orman, K., *Effect of Virtual Reality Exposure and Aural Stimuli on Eye Contact, Directional Focus, and Focus of Attention of Novice Wind Band Conductors*, International Journal of Music Education, 2016, 34(3): p.263-270.

[31]. Sano, M., *Quantitative Analysis of Eye Movements during Singing of Early Childhood Children*, Research Journal of Education, 2021, 7(3): p.125-140.

[32]. Sano, M., *Characteristics of Eye Movements during Musical Expression in Early Childhood by Tonality: Through Quantitative Analysis by Eye Tracking*, Frontiers in Education Technology, 2022, 5(2): p.67-93.

[33]. Nishina, M., Kume, M., Yoshida, T., Takai, Y., and Goto, A., *Process Division of Soldering Work by Skilled Workers and Cooperative Relationship between Motions and Eye Movements in Making Metal Fittings of Flags,* J. Science of Labour, 2013, 89(1): p.1-11.

[34]. Bishopa,L., Cancino-Chacón, C., and Goebl, W., *Eye gaze as a means of giving and seeking information during musical interaction*, Consciousness and Cognition, 2019, 68: p.73-96.https://doi.org/10.1016/j.concog.2019.01.002

[35]. Burger, B., Puupponen, A., & Jantunen, T., *Synchronizing eye tracking and optical motion capture: How to bring them together*, Journal of Eye Movement, 2018, 11(2): p.1-16.

[36]. Fink, L., Lange, E., & Groner, R., *The application of eye-tracking in music research*, Journal of Eye Movement Research, 2018, 11(2)1: p.1-4. doi: 10.16910/jemr.11.2.1

[37]. Jantunen, T., Puupponen, A., & Burger, P., *What Comes First: Combining Motion Capture and Eye Tracking Data to Study the Order of Articulators in Constructed Action in Sign Language Narratives*, Proceedings of the 12th Language Resources and Evaluation Conference, 2020, p.6003–6007. https://aclanthology.org/2020.lrec-1.735