ASIA-PACIFIC JOURNAL OF RESEARCH IN EARLY CHILDHOOD EDUCATION
2007, Vol.1, NO.2, pp.113~128

ISSN 1976-1961

# Children's Music Education Support System using Sensor Data, Video Images, and Sound Data

**Kimio Shintani***, **Takehito Yasutaniya,**
Tokiwakai College
**Hirohide Haga and Shigeo Kaneda**
Doshisha University

## Abstract

In kindergartens and nurseries, a musical classroom is an especially important activity for improving children's self-expression abilities. This paper proposes a children's music education support system that uses sensor data, video images, and sound data. Our system can give suggestions to a teacher by observing and analyzing a child's actions. The proposed system focuses on the motion of a baton equipped with an LED and an accelerometer. A child shakes a baton accompanied by a piano. Marker tracking results and accelerometer output are represented graphically: the horizontal and vertical axes are time and observed values, respectively. An experiment at a kindergarten showed that inexperienced teachers recognize more detailed and rich information from data in a graphical format than observation with the naked eye.

***Key words:*** *music education, marker tracking, baton, accelerometer*

## Introduction

### *Research background*

This paper proposes a new system to support early childhood music education using sensor data. Even relatively inexperienced kindergarten or nursery school teachers[1)] can easily observe the music activity and traits of each child. This research

---

* shintani@tokiwakai.ac.jp

1) In Japan, there are two types of childcare persons: kindergarten teachers and nursery school teachers. In this paper, "*teacher*" is employed for both occupations.

contributes to the music education of early childhood education.

In Japan, kindergarten and nursery schools adopt music in many situations. For instance, it is often used as background music (BGM). This paper focuses on music activities or classes in early childhood education. The teacher plans a musical activity in advance and reserves the music classroom. This music class is pre-scheduled, and the children are controlled by the teacher. In the class, children sing or play a simple instrument, such as a castanet, a bell, or drums. Moreover, many kindergartens and nurseries have adopted “Eurhythmics” developed by Dalcroze2) as childcare activities. “Eurhythmics,” which is effective and important for infantile sociality or physical development [1], matches “expression,” one of the five major categories of government guidelines for early childhood education in Japan.

E. Vanderrspar suggested that the “cycle” of lecture, observation, appraisal, and adjustment in music education effectively promotes children’s development [1]. However, since in Japan one classroom often consists of 25 children,3) the teacher needs to grasp the children’s traits by observation for instruction. In Japan, an actual piano player4) is considered effective for making children’s tunes, so teachers are strongly urged to use a piano or electric organ in musical activities. However, observing each child’s behavior and retaining such observations is hard, especially for inexperienced teachers who usually cannot quickly grasp the traits of each child. Under such conditions, the “cycle” has difficulty functioning correctly.

Conventional approaches for such observation of children use a TV camera whose resultant scenes are manually analyzed. However, such a TV camera approach is limited because observations are subjective, not qualitative. Another limitation is the difficulty of extracting a target scene from a large amount of accumulated videotapes. Analyzing the videotape is time-consuming and arduous. A TV camera approach cannot be applied for daily use.

From such a viewpoint, adopting sensor technology for music education is

---

2) A simplified “Dalcroze Method” is often employed for early childhood education.
3) The age is five in this case.
4) Also, an electric organ is often employed. “Electone” is a well-known trademark manufactured by YAMAHA Corp.

effective for observing each child in music class. For instance, a motion sensor system is widely applied to CG generation for movies. Also, sensor technology is employed in a commercial-based game machine manufactured by Nintendo Corp[5]. However, few studies are reported in the early childhood education domain, especially in music education. A study by MIT is one of the few cases. The MIT Media Laboratory is extracting the expressions of the body in musical activities by sensor technology [9]. However, the target of this research is a new instrument, not teacher support. On the other hand, the proposed system can quantitatively analyze infantile sociality or physical development and feedback the results to teachers.

***A new system for observation***

In this paper, the authors propose a new support system for childhood music education. Teachers can easily analyze each child's expression using the sensor data. From a child's motion, the system visualizes rich information, including strength, durability, rhythmic, etc. Such visualization clarifies a child's traits[6] from individual data and in relation to the group in the musical activity. From the visualized results, classroom teachers can adjust their instruction for each child and the group.

This paper describes experiments conducted to check the validity of the system. The proposed system focuses on the motion of a baton equipped with a red LED and a 2-axial accelerometer. A child shakes a baton while being accompanied by a piano. Marker tracking results and observed accelerometer voltage are represented graphically: the horizontal axis is the time, and the vertical axis denotes the observed values.

A LED marker results in reliable tracking. The children's detailed motions can also be extracted from the accelerometer data. Thus, the traits of individual children can be easily grasped without special skills. An experiment at a kindergarten showed that an inexperienced teacher could also recognize more detailed and rich information from the data in a graphical format than observation with the naked eye.

---

5) "Wii remo-con" is a remote control system of "NINTENDO DS."

6) This is an item from the Developmental Trait Scale based on Freudian theory. Other traits include: trusting others, manipulating, obeying rules, being punctual, being athletic, ability to concentrate, and ability to communicate ideas.

Section 2 shows the conventional approach and a prototype experiment before the new system proposal. Section 3 proposes our new system, and Section 4 shows experimental results. Finally, Section 5 concludes this paper.

## Conventional Approach and Prototype Experiment

***Musical activity in education of children***

Music is often employed in the daily life of kindergartens and nurseries [6, 7]. Figure 1 shows a musical activity scene at a kindergarten where each child has a baton whose motion is timed to the piano. The music education of children has two major purposes: independence (by oneself) and expressing pleasure. The following points are important.

1. Children feel the pleasure of exercise.
2. They feel the rhythms and tempo.
3. They enjoy music.

For children, it is important to express and recognize rhythm and tempo. Teachers should never inculcate “technique” upon children. They must guide each child to enjoy music. Since children are also often physically undeveloped, teachers have to grasp the physical development level of each child. Subjective observations are often misleading.

Let’s assume that a child cannot control his/her baton. This doesn’t mean that this child’s ability is low. Since children often cannot control objects if their physical coordination is low or at an early stage, the ability to control a baton should be evaluated over a long time span. Objective and quantitative evaluation is required. In other words, extracting quantitative information from the objective analysis of children’s movements becomes the major purpose. A new tool or approach is expected to extract the qualitative information from the activities of children[7].

7) This time we tried to extract the necessary data by comparing graphs generated from short-term observations. But we must observe long time development.

*Figure 1*. Shaking a baton with an LED marker

***Interview with an experienced music teacher***

To design a new tool, the authors interviewed an experienced music teacher to learn what kind of attention he paid to his students. Field research showed the following major results:

1. Even good, experienced teachers have difficulty observing each child in a variety of scenes because a teacher must simultaneously observe all the children.
2. Since a teacher cannot take notes during class, she has to mentally store impressions concerning the achievement levels of each child. Such records are often vague, incomplete, or lost.
3. A teacher has to judge whether his teaching is effective. But that decision is subjective and not qualitative.

Video recording is an effective tool from some points of view because it makes a physical and non-volatile record that assists a teacher's recollection of the actual scenes. However, teachers often depend on their impression of video images, which can be misleading and reflect their individual skill. But the major limitation is that a video image has no qualitative measure.

The above discussion suggests that qualitative and objective measurements should be introduced into the long time observation of the music education of children. If a qualitative measurement is introduced, the following information may be extracted from the data:

1. What kind of expression does each child make while listening to music?
2. What kind of music should be selected to develop self-expression activities?
3. How can a child's expression be analyzed to show personal growth?
4. Does each child achieve the action level that the teacher plans for the music class?
5. What kind of detailed support can be realized by analyzing the child's behavior that cannot be detected with the naked eye?

***Pilot experiment for proposed system***

First, before the proposed new system, we designed a pilot experiment to confirm the kind of information that could be extracted. We focused on baton shaking in a music class. Each child was given a baton at the top of which was a table tennis ball as a marker for tracking. "Baton shakes" were shot with a TV camera, and video images were the source data for automatic marker tracking.

Marker tracking was quite difficult due to the natural lighting conditions of the music room. The daytime sunshine caused brightness changes in the room. The automatic marker tracking often failed, and manual tracking was required.

The height of the baton marker position is shown in Figures 2, 3, and 4. The x-axis is the time progress. The y-axis is the difference between the measure (bar) length of the piano and the child's measure calculated from the marker tracking data. These figures show the differences between the baton's beat and the accurate teachers beat.

Figure 2 is a graph for child A. While watching the video image in the pilot experiment, the authors felt that child A matched the rhythm. The experienced music teacher agreed. The observed difference value almost meets the base line, meaning that child A matched the beat and that the naked eye observation was also correct.

The curve of Figure 2 shows that short and long periods appear alternatively. The short period means that the arm muscles are strong, and the longer one denotes weakness. Generally speaking, movement from outer to inner is high because the downward muscle is strong and develops quicker. On the other hand, movement from inner to outer is slow because the lifting muscles are weak and develop more slowly. In other words, this graph shows the development stage of the child's muscles.

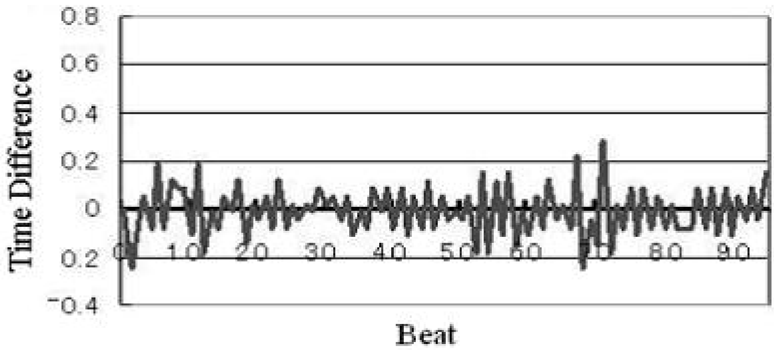


*Figure 2*. Time differences between baton and music (child A)

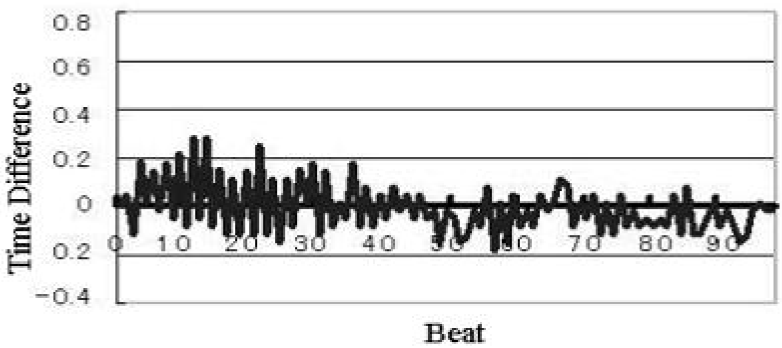


*Figure 3*. Time differences between baton and music (child B)

Figure 3 is the graph for child B. In this case, the authors felt that the child was energetic and accurate. But the expert music teacher disagreed. The curve of Figure 3 gradually falls to the lower right, and the shake amplitude becomes smaller. The boy apparently shook his arm as if swinging a baseball bat. This is an example of how an inexperienced person misunderstands. By graphing the marker tracking results, such misunderstandings caused by inexperience can be overcome.

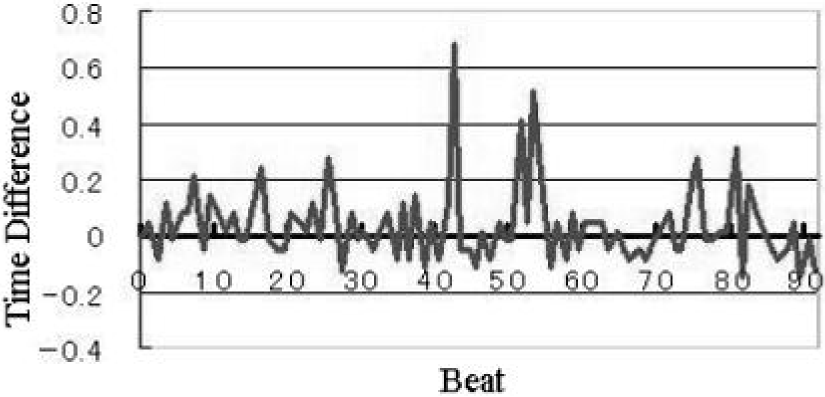


*Figure 4*. Time differences between baton and music (child C)

Figure 4 shows the graph for child C. The authors felt that she didn't keep the rhythm, but followed her own sense of time instead. But the music teacher disagreed, arguing that she nearly matched the rhythm. Figure 4 shows the roughly shifted movement. Except for the illegible parts, the graph is nearly identical to child A (Figure 2), a skilled student. This means that child C almost matched the rhythm. The expert reached that conclusion from the beginning, but the authors (inexperienced persons) misunderstood.

***Results of pilot experiment***

The pilot experiment showed that baton marker tracking is effective. An inexperienced teacher or students can understand the children at an expert's skill level using the tool. But the proposed approach has some problems. First, achieving marker tracking accuracy was problematic because a store-bought video camera has difficulty tracking the baton motion. This reflects the slowness of the TV camera frame rate. Next, there was a lot of obstruction with similar colors in marker tracking.

Moreover, a problem also exists when selecting the standard measure (bar) length. Fortunately, the song in the pilot experiment had a constant speed. But, generally speaking, the measure (bar) length of songs is not constant. The length of one beat must be obtained from the piano sounds.

# A New Children's Music Education Support System

Based on the pilot experimental results, this paper proposes a music education support system that uses sensor data, video images, and sound data. The proposed system can track the marker position in a wide variety of lighting conditions. Figure 5 shows an overview of the system, which used the following compositions to realize its functions:

1. Marker tracking with a red LED maker
2. Detection of detailed baton motion with a 2-axis accelerometer
3. Recording of piano sounds by microphone and extraction of measures (bars) to automatically detect the head of the measure
4. Commercial Hi-Vision (HDTV) camera for ease of operation in the classroom
5. LED time stamp marker that blinks every 10 seconds to create a time stamp important in multisensor environments

To overcome lighting condition limitations, an infrared LED is often used as the marker. In this case, an infrared filter for TV cameras and a nightshot mode are required. But such "surreptitious" camera shots are not appropriate in kindergartens

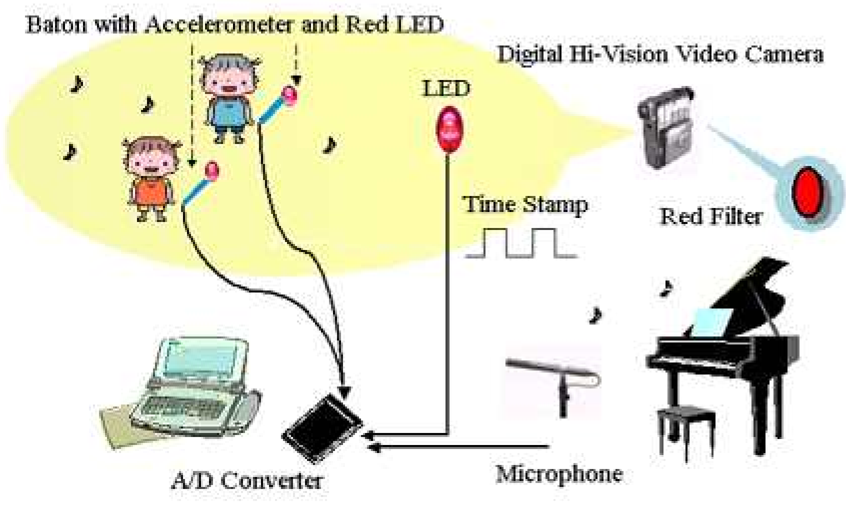


*Figure 5.* Outline of proposed system

and nurseries due to privacy concerns.

Therefore, the authors employed a red LED as the marker and a Hi-Vision camera with a red filter. Figure 6 shows a video image sample. With the filter, only the marker of the baton and timestamp stand out, and the background is dark. Therefore, robust detection of the baton position becomes possible by making binary images in the brightness and calculating the center of the gravity in the white area.

To extract the measure (bar) from the piano sound, the proposed system employs wave envelope analysis, not frequency spectra analysis such as FFT. It detects percussion peaks in the waveform. However, since the algorithm does not detect all the peaks, manual analysis is sometimes required. As mentioned below, an experienced music teacher adjusts the timing while watching the child's response. Since the measure (bar) of the piano sound is not ideal, it doesn't depend upon the original music score.

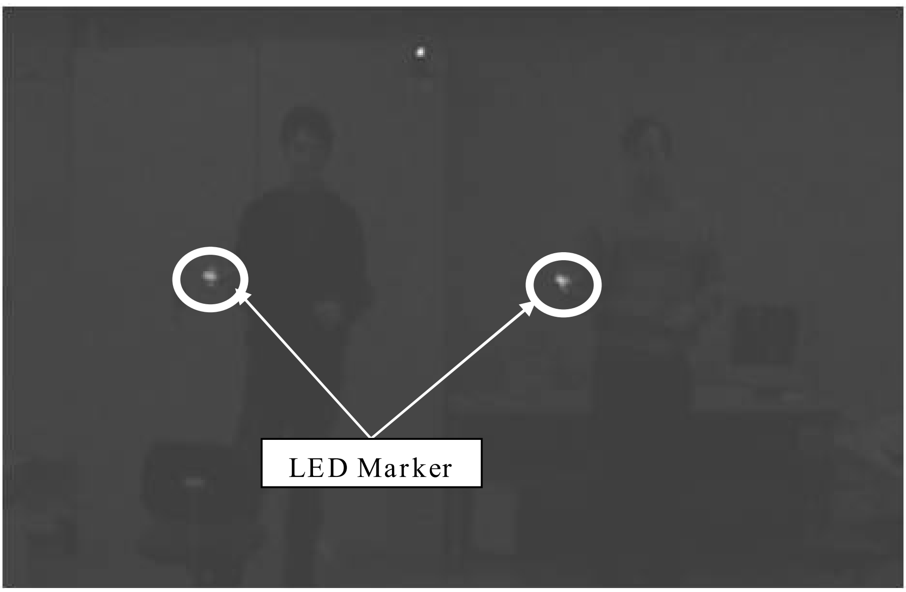


*Figure 6*. Video image observed through red filter

***Experimental Evaluation***

The authors implemented a prototype system and introduced it to a music class of five-year olds at the Tokiwakai Junior College Kindergarten in Osaka. The children sang songs of two beats and shook batons to piano accompaniment. The scene was shot by a High-Vision (HDTV) video camera. A DAT recorder recorded the piano with a small microphone inside the piano. Figure 1 shows the experiment scene.

1. Circuit design of baton

The system's main device is a baton equipped with a 2-axis accelerometer and a red LED. When shaken intensely, the direction in the baton's space is not constant. On the other hand, the emitting angle of the LED is generally narrow. Therefore the authors selected a full color LED manufactured by Nichia Corp., NSTM515AS, with a wide emitting angle. The red LED was activated in this system. Five LEDs are attached to the baton's top. The accelerometer is "ADXL202E" of the Analog Devices Corp., and the bandwidth is up to 6 kHz from DC. The bandwidth is 10 to 100 times wider than a TV camera.

These accelerometer outputs are amplified by two OP-amps: one has unit gain, and the other is an inverter. Since the baton is several meters from the computer, balanced transmission is employed to get lower signal impedance. Electric power for the LEDs and the sensor is supplied by a computer that employs a five leads shielded cable for transmission to prevent noise. These five LEDs, a sensor, and operational amplifiers are composed on one small universal circuit board.

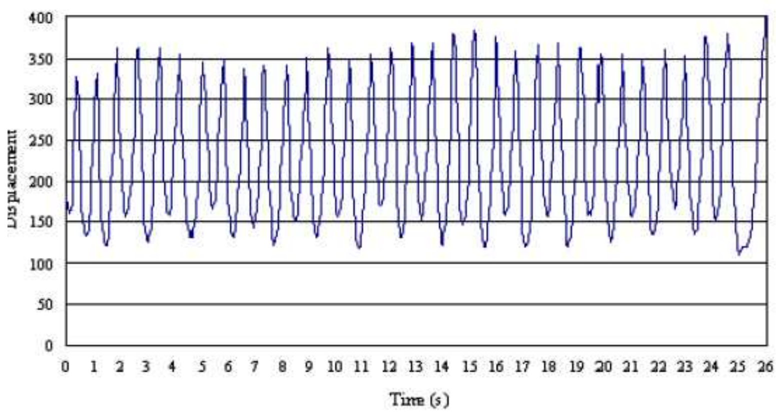


*Figure 7.* Marker height of teacher's baton

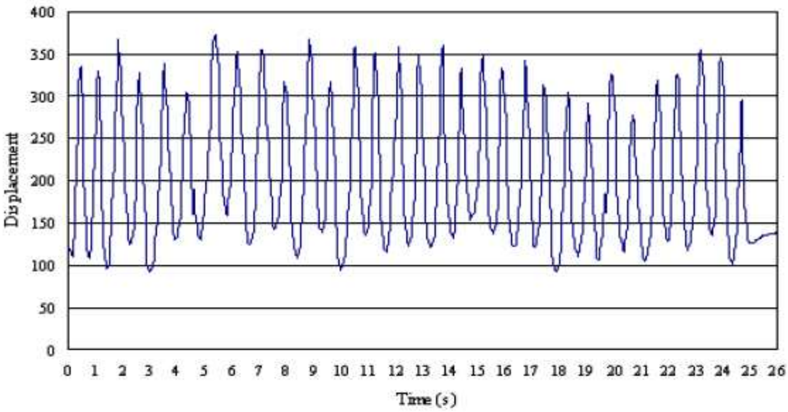


*Figure 8*. Marker height of child's baton

Marker tracking is reliable. By using a red LED and a red filter for the TV camera, only the markers and the time stamp appeared, so the positions are automatically detected. But if the TV image gets direct sunlight, tracking the marker becomes difficult.

2. Experimental Results

Figures 7 and 8 are the marker tracking results for the teacher and the children, respectively. The y-axes show the height of the marker in the video image. There doesn't seem to be major differences in their actions. The authors and an expert music teacher could not find differences in the classroom, either. However, the above subjective impressions were misleading. Figures 9 and 10 show a major difference in accelerometer data between an experienced teacher and a child, respectively. The y-axis is the output voltage of the sensor. The teacher's data (Figure 9) have a cycle and many peaks with constant intervals.

On the other hand, the child's data (Figure 10) have no peaks. The length of one measure (bar) is not constant. The child doesn't take the rhythm to the bottom point of the baton movement because she is following the teacher's movements. She can't successfully keep the rhythm, even though she is following the teacher as well as she can. But she doesn't seem to enjoy the music or the baton shaking. This example clearly illustrates that we can observe more information with the accelerometer than with marker tracking.

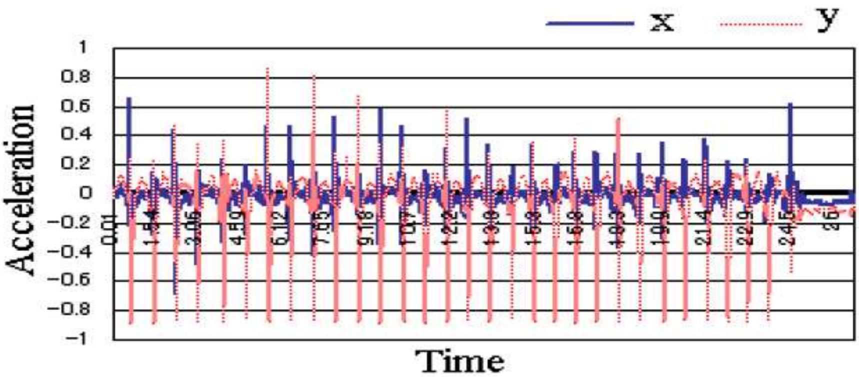


*Figure 9*. Two-axis accelerometer data of teachers baton

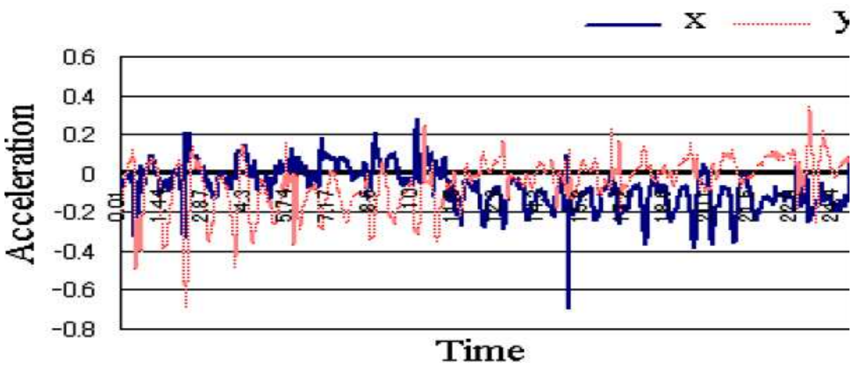


*Figure 10*. Two-axis accelerometer data of childs baton

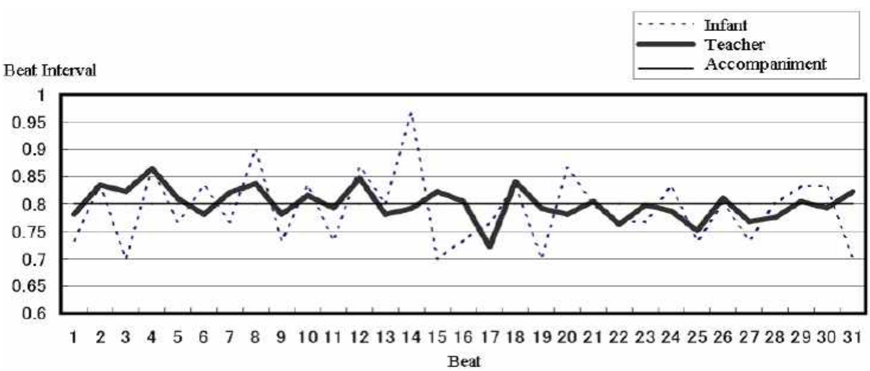


*Figure 11*. Time differences between piano and children

This accelerometer system has one problem. Since the sensor's axis is perpendicular to the baton when vigorously shaken, the sensor output signal value has an indirect relationship to the camera image. This limitation of the baton must be improved. Finally, baton movements were compared with the measure length calculated from the piano sound data, as shown in Figure 11. The Y-axis values are the measure length of the piano musical score, the experienced teacher, and the child. Figure 11 shows a very interesting phenomenon. The teacher's playing speed does not equal the values of the musical score. The interval of the measure often varies. Since the teacher is very skilled, this is not his mistake. He adjusts the tempo while observing the children. If the children cannot follow him, he delays the beat. The children notice the tempo changes and shake the baton faster. In this performance, the baton became uptempo after the first twenty beats.

## Conclusion

This paper proposed a music education support system for children. An inexperienced teacher can observe and understand individual children in detail at kindergartens or nurseries using this proposed system[8]. Children's traits and their activity can be minutely understood by marker tracking and sensor data.

In an experiment at a kindergarten, marker tracking data showed that an energetic child shook his baton with an original and inconsistent tempo. These results agreed with the expert's opinion. However, inexperienced teachers (the authors) regarded this child as skilled in the classroom. Also, the accelerometer detected drastic differences between the baton shaking of a teacher and a child. Observations using marker tracking failed to detect any differences. In some cases, the accelerometer offered richer information than marker tracking.

---

8) It contains the attention points of graphs appreciated beforehand that must be analyzed.

Acknowledgements
The authors would like to thank Prof. Akeru UEDA, Tokiwakai College, for her kind guidance and cooperation during this research. Also, the authors would like to express our deep appreciation to the teachers and the children of the Tokiwakai Kindergarten for their support and cooperation.

The system employed automatic extraction of measures (bars) from the piano sounds. However, not only is extraction technically difficult but also the expert music teacher intentionally varies the tempo by controlling the song speed of the children with tempo deviation. This means that the tempo of the teacher's piano is not a good example for automatic analysis.

## References

Vanderspar, E. & Ishimaru, Y. (2002). A Dalcroze handbook: Principles and guidelines for teaching eurhythmics(In Japanese), Tokyo: Doremi Music Publishing.

Aronoff, F. W. (1988). Reaching the Young Child through Music: Howard Gardner's Theory of Multiple Intelligences. Model, International Journal of Music Education, 12, 18-22.

Massachusetts Institute of Technology Media Laboratory (2002). http://www.media.mit.edu/

Shibuya, M., Shintani, K., Bando, T., Kaneda S., & Yanagida, M. (2003). A Music Education Approach using Marker Tracking for Infants(In Japanese), FIT2003, LN-002.

Mihon, M. & Nishida, S. (1999). Information Media Technology(In Japanese), Tokyo : Ohm Publishing.

Oda, Y. & Kaminaga, M. (1999). A New Explanation of Kindergarten Education Guideline(In Japanese), Tokyo: Daiichihoki Publishing.

Kojima, T. (2002). Planning and Practice of Music Expression(In Japanese), Japan Society of Research on Early Childhood Care and Education, The 55th Domestic Workshop, pp.758-759.

Mistsumori K. (2000). Present Situation of Music Expression(In Japanese), Japan Society of Research on Early Childhood Care and Education, The 53rd Domestic Workshop, pp. 512-513.

Feldmeier, M., Joseph, A., & Paradis, A. (2004). Gieaway Wireless Sensors for Large-Group Interaction, http://www.media.mit.edu/resenv/pubs/papers/2004-05-lb632-Feldmeier.pdf

Feldmeier, M. (2007). Large Group Musical Interaction using Disposable Wireless Motion Sensors, http://www.media.mit.edu/resenv/pubs/theses/Feldmeier-SM.pdf