

センサー・画像・音声のデータ分析による 幼児音楽指導支援システム

清水 宏章[†] 豊田 実香[†] 森本 訓貴[†] 新谷 公朗^{††} 芳賀 博英[†]
金田 重郎[†]

[†] 同志社大学大学院工学研究科

^{††} 常磐会短期大学幼児教育科

あらまし 幼稚園・保育所における幼児教育では、音楽は表現力を養うのに重要な活動である。しかし、経験の浅い保育者にとって、ピアノを弾きながら多数の幼児を観察し、集団の中で幼児1人1人に適した指導をすることは難しい。このような問題を解決するために、本論文ではセンサー、画像、音声のデータ分析による幼児音楽指導支援システムを提案する。具体的には、幼児が振るタクトに着目し、LEDと加速度センサーを付けたタクトをピアノの伴奏に合わせて幼児たちに振ってもらった実験を行う。その様子をビデオカメラで撮影し、データを取得する。実験によりマーカートラッキングからは読み取れない、細かな動きを加速度センサーから抽出できた。

A Support System for Infant Music Education by Analyzing Sensor, Video, and Sound Data

Hiroaki SHIMIZU[†], Mika TOYODA[†], Kunitaka MORIMOTO[†], Kimio SHINTANI^{††},
Hirohide HAGA[†], and Shigeo KANEDA[†]

[†] Graduate School of Engineering, Doshisha University

^{††} Infant Education, Tokiwakai College

Abstract In the education of kindergarten and nursery schools, music is an especially important channel for infant's expression. For inexperienced kindergarten teachers or nurses, however, it is difficult to observe every infant carefully with playing the piano. To resolve the problem, this paper proposes a support system for infant music education by analyzing sensor, video, and sound data. Focusing attention to a motion of the baton shaken by infants, the experiment was conducted by having infants shake the batons, attached LED and the accelerometers, by a piano accompaniment. Consequently, the marker tracking is reliable and the accelerometers detects the detailed motion that could not be observed through the video data.

1. はじめに

子供の音楽活動は、言語の獲得や身体の発達を促す要因として捉えられており、子供の個性を大切に今日の幼児教育では、1人1人の子供をきめ細かく観察し、それに適した指導が望まれ、幼稚園や保育所での音楽指導はクラス単位で行うことが多い。集団の中で指導者が個々の子供の様子を注意深く観察し覚えておくことは難しい。また、実際にはピアノを弾きながら指導する場面も多く、

このような場面において経験の浅い指導者では難しい作業である。

このような状況を回避する手段として、ビデオカメラの利用が考えられる。ビデオカメラは、幼児教育研究の中でも、保育方法指導やその分析などに取り入れられている。ビデオ画像を用いた保育方法の指導・分析は、一定の効果が期待される。しかし、幼稚園・保育所の現場において、大量に蓄積されたビデオテープから、対象となるシーンを抽出し分析することは、熟練を必要とする

だけでなく、労力と時間を要する。

著者らは音楽指導の中で、マーカートラッキングを用いて動きをグラフ化し、経験の浅い保育者・幼児教育科学生を支援するシステムの開発を進めている[1]。予備実験から、視察やビデオ画像からは得られない、子どもの客観的な様子を読み取り得ることを確認できた。しかし、1) 現場の照明条件ではマーカートラッキングが困難、2) 速度が変化する曲への対応ができない、等の問題が残されていた。

そこで、本稿では、これらの問題を解決するため、センサー・画像・音声のデータ分析による音楽指導の支援手法を提案する。具体的には、発光ダイオード(LED)を用いたマーカートラッキングによるタクトの動き、タクトに取り付けた加速度センサーのデータ、伴奏から得られる拍のデータを統合した分析を試みた。その結果、LEDにより確実なトラッキングが実現できた。また加速度センサーからは画像では得られない子どもの細かい運動を抽出できた。

以下、2章で幼児音楽教育と予備実験について述べ、3章では提案する音楽支援システムの概要を述べる。4章では現場実験による評価について述べ、5章でまとめを述べる。

2. 幼児音楽教育と研究

2.1 幼児の音楽と表現

音楽は、幼稚園・保育所において幼児の日常生活の様々な場面で取り入れられている。音楽を幼児の身体運動に取り入れることで、運動学習効果が上がるという研究事例も報告されている[2]。音楽が扱う分野は体を使ってリズムやテンポを表現するものが多く、発達や身体運動と密接な関係にあるといえる。図1は、幼稚園において実際に幼児がタクトを振っている様子である。



図1 幼児たちがタクトを振る様子

「自分なりに」、「表現する楽しさ」のポイントとして、

- (1) 体を動かす楽しさを感じる。
- (2) リズムやテンポを感じ取れる。

(3) 音楽に親しみ、音楽を楽しむ。

等が指導の重要な要素であるといわれている。また、音や音楽を楽しく表現するためのプロセスとして、一定のリズムやテンポを意識し、表現できることも大切である。ただし、教え込むのではなくリズムが楽しめるように指導することが大事である。

したがって、「幼児が表現したもの」と「幼児にそのような表現をさせたもの」という2つの側面を観察し分析することが、幼児の指導上重要である。また、幼児の場合、大人とは違い、未発達な部分を多く残している。音楽性をもって表現されてものか、本能的に無意識に表現されたものかを量的に判断することは、疑問の生じるところである。つまり、客観化されたデータからいかに量的情報を抽出し、分析するかが課題となる。本稿が提案するシステムはそのような情報抽出の一助となることを目指している。

2.2 音楽指導の現状

実際の音楽指導の現場で、指導者は幼児のどのような部分に注目して指導しているのかを探るため、ベテランの音楽指導者にヒアリングを行った。主要な結果を以下に示す。

(1) ピアノ等を使用した音楽指導では、幼児全体の様子を見ながら指導している場合が多く、それぞれの場面で幼児1人1人を細かく観察することは、ベテラン指導者でも難しい。

(2) 音楽や体育等の身体表現を伴うものは、その場限りであり、指導後に幼児が課題をどれだけ達成できたかは、指導者の観察と記憶に頼るほかない。

(3) 指導者の指導方法が妥当であったか否かについても同様のことがいえる。

以上の結果から、指導者の視点を補う観察方法として、客観的な要素が強く、かつ記録性を有するビデオ記録は有効な方法である。しかし、ビデオ画像を観察記録として用いた場合、視覚データだけに頼ってしまい「できた」「できなかった」というような単純な評価に落ちいつてしまう。ビデオを観察に用いる場合には、画像から読み取ることのできる情報を明確にし、読み取った情報の信頼度を高めるデータ分析、表示方法についても検討が必要である(注1)。このような観点から、以下のような情報の抽出とその活用方法を考える必要がある。また、これらを追跡するには、幼児の動作を何らかの形で定量化・グラフ化する必要がある。

(1) 幼児が音楽により刺激を受けどのような表現行動をしたか

(注1): 実際、後述するように、画像情報からでは見えない差が加速度センサーから得られている。ビデオ画像はかなり解像度も低く、ビデオ画面から得た感触のみで何かを読み取ることは、かなり疑問が残る。

(2) 幼児の表現活動を刺激する音楽、リズムやテンポを、どのように選ばよいか

(3) 幼児の発育段階を踏まえた表現の分析と理解

(4) 音楽によって幼児が、指導者の意図した表現行動を行っているか

(5) 指導者が、指導中に気付かなかった幼児の個々の動きを分析することにより、幼児へのきめ細やかな指導の支援

2.3 予備実験

以上のような観点から、著者らは予備実験として、音楽指導のタクトを振る場面をビデオカメラで撮影するとともに、幼児たちが振るタクトの先端にマーカーをつけて、その動きをトラッキングすることを試みた[1]。タクトの位置データについては、Y軸における極小値から次の極小値までの時間差を求め、曲自体の1拍の平均的長さとの差を計算した。

図2、図3、図4は、このようにして得たタクトの時間差グラフである。これらの図は、『幼児が振ったタクトが示す拍の長さ』-『ただし1拍の長さ』を示している。図2は、ビデオ映像から著者らは「リズムに合っている」と感じた幼児である。振れはあるものの（幼児のタクトと正しい拍との間に差がないことを意味する）基線を中心とする値である。正しい拍に対して、早めと遅れが交互に出るのは、幼児の場合に筋肉が未発達で、腕を外から中に入れるほうが早くなり、逆方向が遅くなるためだと思われる。言い換えれば、筋肉の発達の様子が、このグラフから読み取れる。

一方、図3は、ビデオ映像から著者らは元気でよいと感じた幼児である。しかし、グラフを見ると、全体にだんだんと右下へ下がっており、振れも小さくなっている。これは、「パットの素振り」のように、ピアノの演奏とは異なった自分独自の感覚のリズムで腕を振っているために、後半は疲れて、遅れてきているものと想定される。つまり、初学者が最初の一見の印象（「元気そうな子だ」と著者らは思った）に左右されて誤解した例である。マーカートラッキングを行い、タクトの動きを時間軸方向に展開することにより初学者の誤解を解くことがある。

また、図4は「曲のリズムではなく、自分自身の感覚で振っている」と著者らは感じた幼児である。しかし、ベテラン指導者はそれなりにリズムに合っているとした。図から見ると、確かに、一部、大幅に周期がずれた場所があるが、それ以外は、リズムに合っている幼児と類似している。そのように意識してビデオを見直すと、この幼児は曲の『さび』の部分ではピアノの演奏と異なるリズムで腕を振っていた。ベテラン指導者は最初からそれを見抜いていたが、著者らはビデオ画像の初見では、自分の感覚で振っていると誤解していた。

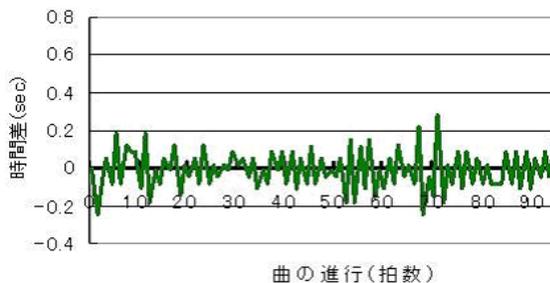


図2 幼児Aが振るタクトの時間差グラフ

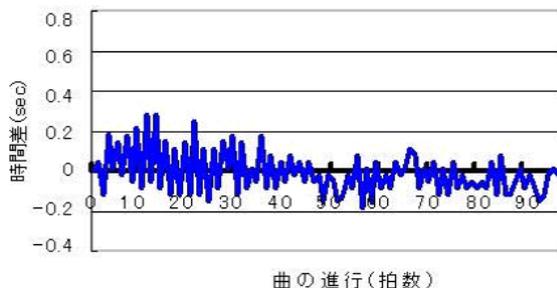


図3 幼児Bが振るタクトの時間差グラフ

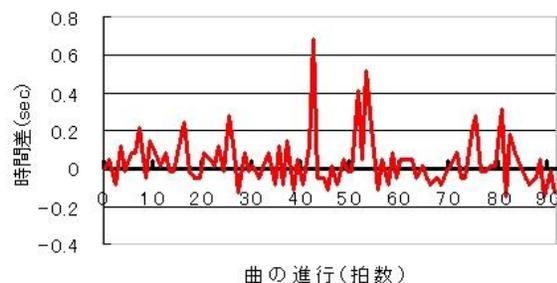


図4 幼児Cが振るタクトの時間差グラフ

2.4 予備実験の結果

著者らのアプローチは、タクトの動きをトラッキングして、時間軸方向に「正解との差」としてグラフ化する点に最大の特徴がある。これにより、初学者が画像から読み取り得ない幼児の状態を、簡単に認識できる可能性がある。言い換えると、ビデオによる視察は、特に初学者は誤解しやすい様であり、このような支援システムの効果は大きい。

しかし、幾つかがの問題が残っている。ひとつは、マーカートラッキングの精度である。市販民生用ビデオカメラではフレームレートが遅く残像が出るために、トラッキングが難しい。また、現場の厳しい照明条件では、マーカートラッキングの障害となる類似色の物体が多数存在している。

一方において、何を持って拍の長さの正解とするかも問題となった。図2、図3、図4に利用している曲（『手のひらを太陽に』）は一定速度の曲であったが、一般には、速度は曲の途中で変化するだけではなく、テヌートなどが存在する。したがって、指導者の弾くピアノの音

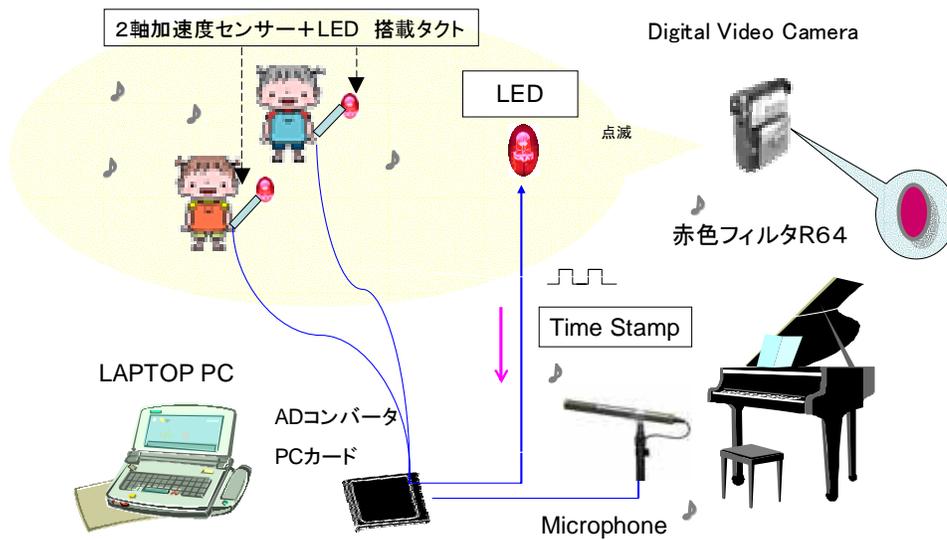


図 5 システムの概要

声からの一拍の長さの取得が望まれる。

3. 音楽指導支援システムの提案

以上述べた予備実験の問題点を踏まえ、現場の照明条件でも確実にマーカートラッキングし、細かな動きまでを抽出することを目的とした、センサー・画像・音声のデータ分析による音楽指導の支援を提案する。図 5 に提案するシステムの概要を示す。本システムを実現するために以下の構成を用いた。

- (1) 発光ダイオード(LED)によるマーカートラッキング。
- (2) 加速度センサーによる詳細なタクト動作の検出。
- (3) ピアノ音から自動的に拍の頭を検出するため、マイクによりピアノを録音^(注2)。
- (4) 現場での運用容易性を考慮して、カメラとしては民生用ハイビジョンカメラの 525p(60fps) を利用。
- (5) マルチセンサでは重要となるタイムスタンプについては、10 秒に 1 回点滅するタイムスタンプマーカー LED を作成してカメラ撮影。

【赤色 LED を用いたマーカートラッキング】

照明条件に左右されない手法としては、マーカーとしては赤外線 LED を用いた手法が一般的である。赤外線フィルターとナイトショットモードを有するビデオカメラの併用が前提となる。しかし、これは「盗撮カメラ」と等価である。保育者が夏場には水着で行動する幼稚園・保育所の労働環境での採用は疑問が残る。そこで、マーカー

(注2): 音声もタイムスタンプが必要な場合には、音声信号用の AD コンバーターボードにこのタイムスタンプ信号を入力する予定であるが、今回の実験では実施していない。



図 6 赤色 LED/赤色フィルター併用による撮影

として赤色 LED を使い、赤色フィルター (Kenko R64) を装着した民生用 HD カメラでの撮影を試みた。実際にタクトを撮影した様子を図 6 に示す。これにより、輝度のみで 2 値化し白領域の重心を計算することでタクトのロバストな位置情報検出が可能となった。

【加速度センサー】

カメラ画像の周波数特性は高々 15Hz 程度^(注3)であり、実際には画像処理の過程で細かい情報が落ちる。そこで、より精度の高い情報を得ることを目的として、タクトに 2 軸の加速度センサーを埋め込むこととした。方向は、タクトの長手方向に直角方向である。センサーの出力電圧はノイズ対策として平衡伝送に変換してパソコンの AD コンバーターボードへ伝送している。

【音声処理】

(注3): 幼児の様子は一般の民生ビデオカメラで撮影している。30fps である。

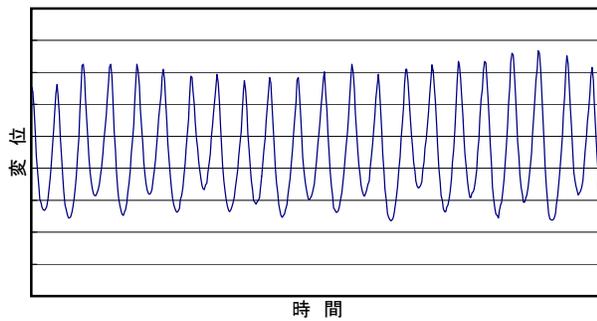


図 7 マーカートラッキング結果 (指導者)

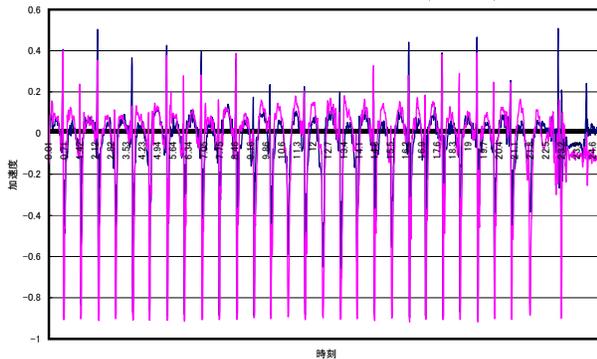


図 9 加速度センサー出力電圧 (指導者)

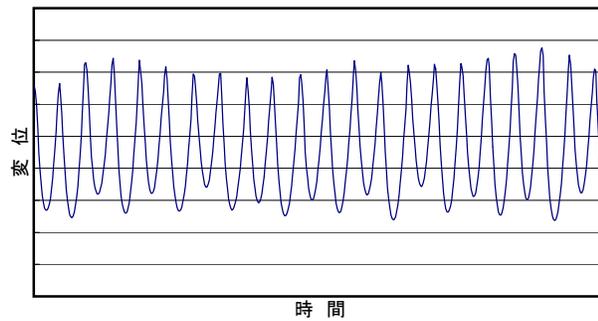


図 8 マーカートラッキング結果 (幼児)

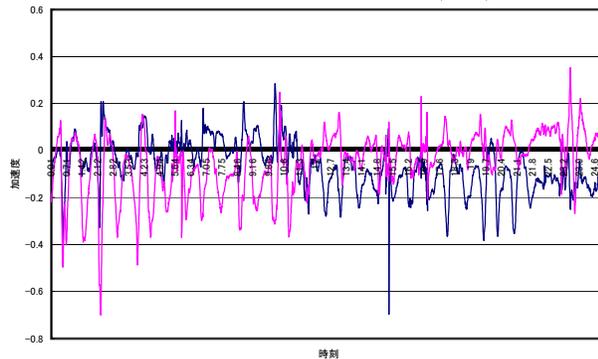


図 10 加速度センサー出力電圧 (幼児)

ピアノの伴奏からの拍の取り出し方は、FFT等の周波数スペクトル分析は実施せず、波形エンベロープを用いた。打楽器での演奏では明確に現れる波形の立ち上がりを探るという自動分析を行った。この取得は完全にはできず、一部手作業が入っている。これは、後述するように、指導者は単純に楽譜どおり演奏しているのではなく、幼児たちの様子を見ながら、かなりタイミングをずらせていることも起因する。音声処理の改善は今後の課題である。

4. 評価実験

実際に大阪市平野区にある常磐会短期大学・付属幼稚園において5歳児の音楽指導の時間に実験を行った。幼児は指導者のピアノあるいは、歌に合わせて加速度センサー付きのタクトを振り、指揮を行った。その様子をビデオカメラで撮影した。曲は「はたけのポルカ」で2拍子である。ピアノの音は今回はDATに小型マイクを接続しピアノの内部に入れて録音した。図1はその実験風景である。

【マーカートラッキング】

赤色LEDを用いた実験では、タクトの先端のマーカータイムスタンプマーカ（たまたま点灯している）以外は全く見えず、自動的にマーカ位置を画像処理で抽出できた。ただし、輝度のみにより領域抽出をしているため、背景に強い反射や直遮光の差込がある場合には、

トラッキングは難しいと思われる。また、幼児同士の距離を一定距離以上とらないと、タクトの先端が隣の幼児の腕に隠れたり、または隣とタクトが錯綜したりすることがある。このようなケースへの対処は今後の課題である。

【加速度センサー】

加速度センサーは画像よりはるかに細かい情報をもたらす。図7、図8は実際の指導者とある幼児のタクトのマーカートラッキング結果である。この2人の動作に差があるとは思えない。一方、図9、図10は、加速度センサーのデータである。様子は全く異なる。指導者がタクトを振った場合には、センサーの両軸共に波形に「きれいに揃った整ったピーク」が現れ、また値とピークの周期も等しく波形（データ）に現れている。一方、幼児が振ったタクトの方は、2軸のセンサーデータがばらつき、またセンサーのピークの周期、値もばらついている。

この印象をもって画像データを見ると、ほとんどの幼児は指導者のようにタクトに、上手に力を入れて振っていないことに気がされた。タクトの描く軌跡の最下点で、力をいれてリズムを取っているのではなく、タクトが振れる軌跡の鉛直方向の振幅でリズムを取っていた。これは、幼児と指導者とは、タクトを振る時のリズムをとるタイミングが異なることを加速度センサーのグラフから読み取れた例である。この例からも、幼児のタクト

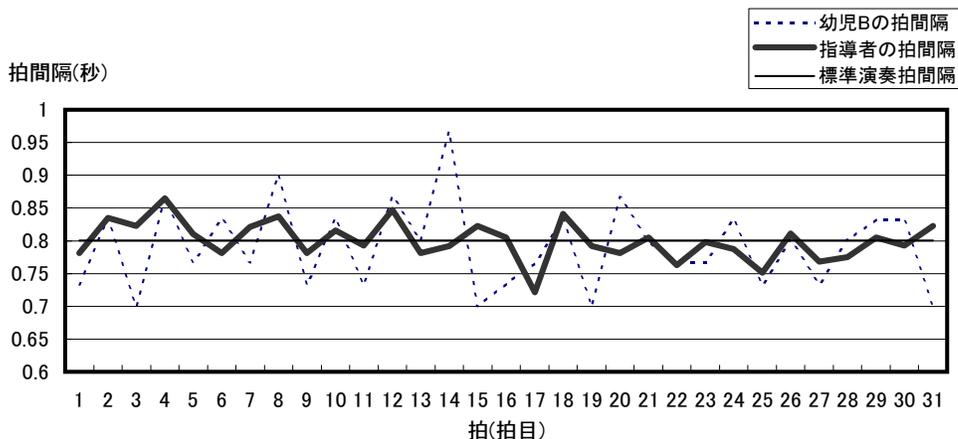


図 11 幼児のタクトとピアノ演奏の拍の比較

の振り方(手首と腕の運動)の客観的な評価尺度として、有効であると考えられる。

尚、加速度センサーの軸の方向はタクトに直行しており、タクトは振り回されるので、カメラ画像との対応は単純でない。少なくとも、タクトに埋め込まれたセンサーの軸方向と親指位置との関係は、握り構造を工夫するなどの対策により、一定に保つ必要が感じられた。

【ピアノ伴奏からの拍抽出】

ピアノ伴奏から小節の頭を取り出し、幼児の一拍のお手本とすることがこの機能のひとつの目的である。しかし、拍抽出結果を見ると、ベテラン音楽指導者は、必ずしも譜面どおりに一定の拍で演奏してはいない。幼児たちの歌の様子を伺いながら、その時々に合わせてテンポを調整している。

図 11 は、平均の 1 拍の長さ (0.8sec) に対する、ある幼児のタクトの拍とピアノの拍の長さを示す。この演奏では 20 拍目あたりからテンポアップしている。幼児は揺らぎながらもテンポアップを感じとり後半は速く振っている。また、センサーデータから幼児は、腕の水平方向の振幅でリズムを取りタクトを振っていると述べたが、図 11 の拍の間隔からは、遅い、速いを、ほぼ交互に繰り返しているのが判る。以上から、指導者のピアノ音声から得た拍の長さデータは、必ずしも幼児が振るべき一拍の長さともいえない。幼児の拍のお手本には、指導者の一拍の長さをマクロにみて平均化したものが必要である。

5. ま と め

本論文では、幼稚園・保育所を対象に、経験の浅い保育士が同時に複数の子供を観察し、指導を行う際の音楽支援システムを提案した。提案手法により、画像・センサーのデータを利用することで、幼児たちの特徴や行動を細かく読み取れることが分かった。ピアノ伴奏との拍

抽出においては、今回は評価基準として演奏者の拍を利用したが、演奏者は必ずしも一定の拍で演奏しているのではなく、幼児たちの様子を見ながら、テンポを調整しているため、実際の現場において最適ではなく、評価基準を見直す必要がある。また、マーカートラッキングにおいては、隣の子供のタクトと交差したり、タクトが隠れたりして、マーカーが見えなくなった場合には実行できないので、その場合にはセンサーデータで補間ができるように今後行っていきたい。

本研究の狙いは、幼児教育科の学生が幼児たち 1 人 1 人の成長、キャラクターをどのようにして把握するのか、あるいは、長期に観測したときの成長をどのように見出すのかを支援することにある。今後は、幼児教育科学生に実際に利用してもらいながら、同一の幼児を長期に観測して波形の変化をみてゆく必要がある。ただし、ケーブルが繋がったタクトでは、幼児たちの自由な活動を制約している。今後は、無線で加速度センサーの情報を伝達できるより実践的なタクトを開発して行きたい。なお、本研究の一部は学術フロンティア『知能情報処理とその応用』(主査:同志社大学工学部・三木光範)による。

本研究を進めるあたり協力をいただいた、常磐会短期大学・安谷屋教授、ならびに、常磐会短期大学附属常磐会幼稚園・植田園長に深謝いたします。

文 献

- [1] 渋谷真人, 新谷公朗, 坂東敏博, 金田重郎, 柳田 益造 “幼児を対象としたマーカー追跡による音楽指導支援”, FIT2003, 講演番号 LN-002(CD-ROM 予稿集) 2003 年 9 月
- [2] 三宅一郎 “BGM が運動技能獲得に及ぼす影響”, 日本保育学会第 55 回大会発表論文集, pp758-759, 日本保育学会, 2002 年 5 月