

子どもの動作からの関心推定に向けた センシングシステムの開発

浅利 恭美^{*1}, 山田 徹志^{*2}, 青柳 燎^{*3}, 騰川 裕太^{*3}, 大森 隆司^{*3}

*1 玉川大学大学院工学研究科 *2 玉川大学脳科学研究所 *3 玉川大学工学部

Development of a sensing system for estimating interest from child's behavior

Yukimi Asari^{*1}, Tetuji Yamada^{*2}, Ryo Aoyagi^{*2}, Yuta Togawa^{*2}, Takashi Omori^{*2}

*1 Tamagawa University, Graduate School of Engineering, *2 Tamagawa University, Brain
Science Institute, *3 Tamagawa University, School of Engineering

現在, 幼児教育学領域では質的研究を用いた子どもの成長・発達の評価が行われている。これは, 有識者(教育実践者・研究者)の実践経験に基づいた直感的な解釈よりなされ, 子どもの心の状態を解釈し成長・発達の評価を実施する上で有効な手法である。一方, その作業には膨大な時間と労力が必要であり分析場面にも制約がある。この問題に対し本研究では, 子どもの物理的な動作をセンシングし, 子どもの心の状態(関心)を推定する技術開発を行い, 幼児教育の研究者および実践者の作業支援を目指す。以上より本稿では, 子どもの関心推定に向けた動作のセンシング手法について報告する。

キーワード: アフェクティブコンピューティング, 機械学習

1. はじめに

現在, 幼児教育学では, 子どもの行動を質的分析により解釈し, 育ちの状態(個々の発達の状態)を評価する研究が行われている。この方法は, 研究者や教育実践者(幼稚園教諭・保育士)の経験に基づいた直感的な解釈により子どもの心の状態を記述しており, 教育学研究において重要視される人と事象とのインタラクションの形成過程を紐解く上で有効な手法である。

一方, 課題となるのが人手による膨大な作業と記録事例の制限である。さらに, 分析結果は定性的なものになりがちで客観性の担保にも限界がある。これらの問題に対し本研究では, 幼児教育研究を支援する新たな手法として, 子どもの心の状態(関心)を「見える化」するIT技術を開発し, 定量的かつビックデータを用いた幼児教育研究支援技術の実現を目指す。

幼児教育研究者・実践者は, 子どもの日常的な生活や遊びにおける動作を観察し, その都度, 心の状態(関心)を読み取ることで, 心情・意欲・態度を基礎とし

た子どもの育ちの度合を評価している。これより我々は, 子どもの関心推定へ関与する情報は, 子どもの動作(物理的な行動量)の中にあり, その詳細な観察により関心推定が可能であると考え。そこで, 本研究は工学的な手法により子どもの動作をセンシングすることからの関心推定を目指す(1)。

そのため本研究では, ステレオ画像計測および人工知能による画像処理 OpenPose(2)を用いた3次元骨格検出を行った。その際, OpenPoseによる骨格検出の揺らぎによる骨格の揺らぎを解消するためにテンプレートマッチングによる画像処理を行った。本発表のその結果について報告し, その手法の可能性を議論する。

2. 頭と体の位置・向きによる関心推定

我々は, 子どもの関心を推定するための定量的な指標として, 頭と体の位置・向きに着目した。人は何かに関心を見出している際は, その対象によって行く。また, 関心対象を一定時間注視し続ける。この動作は,



図 1 取得画像



図 2 OpenPose の出力結果

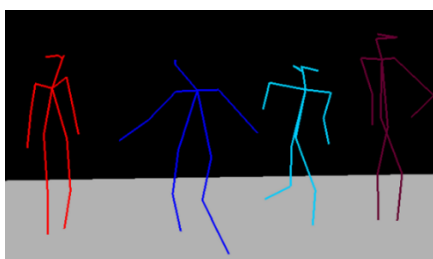


図 3 正面から見た 3 次元再構成結果

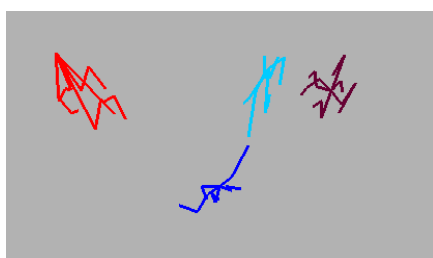


図 4 上部から見た 3 次元再構成結果

頭と体の位置・向きを観察することで検出できる。特に子どもは、正直に関心対象の方を向いて注視する傾向が高いと考えられる。実際、山田らは子どもの関心推定に対する体や頭の位置と向きと有用性を示している(3,4)。これより我々は、物理的な行動の観察を通じた子どもの関心推定の手法を検討した(5)。

3. センシングシステム

3.1 センシングの手法

本手法で人の関心を推定するには、3 次元的な頭と体の位置・向きといった姿勢の検出を連続的に行う必要がある。そこで、複数台の 4K カメラによるステレオ計測を用いて幼稚園での教育活動場面を 10fps で撮影・録画する。4K カメラを用いることで顔などの身体

パーツの検出可能範囲を広げ、ステレオ計測によって立体視を可能とした。得られたステレオ画像に対して、2D 画像から人物の骨格検出が可能な OpenPose を用いて人物の骨格の特徴点を抽出した。この出力から人物の骨格の 3 次元再構成を行い、人物の姿勢推定を可能とした。

3.2 OpenPose による 3 次元再構成

OpenPose による人物骨格の検出結果は、1 つの画像中に複数人いる場合は順不同で出力される。順番が異なると左右画像で対応せず 3 次元再構成を行うことができないため、人物骨格の画像中の位置の変化量により個人の骨格をフレーム順に 2 次元的に追跡し、さらに左右画像での特徴点の対応付けを行った。その後、3 次元再構成を行った。

検証として、成人 4 名に対して 1 台のステレオカメラを用いて計測を行い、6 秒間の場面に対して 3 次元再構成を行った。取得画像 (図 1) より OpenPose を用いて骨格検出 (図 2) を行い、3 次元再構成 (図 3, 図 4) を行った。図 3 より 3 次元的な姿勢推定が可能であることが示された。しかし、手足の位置には揺らぎがあり、さらに揺らぎのために奥行方向に大きなズレが生じていた (図 4)。ズレが生じる原因は、OpenPose による骨格検出の結果の揺らぎであった。OpenPose は左右画像のそれぞれで骨格を検出するため、得られた特徴点がゆらいでもそのことを知ることはできなかった。この問題に対して我々は、骨格検出後の処理を行うことで解決を図った。

3.3 テンプレートマッチングによる揺らぎ補正

OpenPose の揺らぎの問題に対して画像処理の手法のテンプレートマッチングを用いて揺らぎの補正を行った。これは、一方の画像の特徴点の検出位置を基準として、輝度の相関を用いてもう一方の画像から同じ位置を探索する手法である。

この手法を用いて、4 秒間 (40 フレーム) の静止状態の人物に対して 3 次元再構成を行った。また、関心推定には頭と体の向きが重要であるため、肩周辺が重要と考え、図 3 より比較的安定している首、右肩、左肩の連続時間による奥行方向の揺らぎをグラフ化した (図 5, 図 6)。図 6 より、前後フレームと比較して



図5 対象人物の取得画像（左）と OpenPose による骨格検出結果（右）

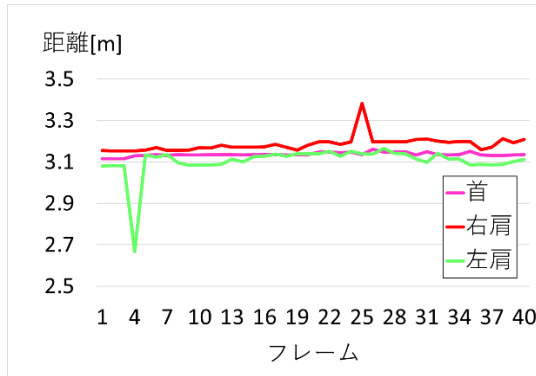


図6 肩周辺における奥行方向の揺らぎ

表1 肩周辺の揺らぎ
(全データの標準偏差)

特徴点	奥行方向	横方向	高さ方向
首	0.9cm	0.5cm	0.4cm
右肩	3.7cm	0.8cm	1.3cm
左肩	7.4cm	0.6cm	1.9cm

表2 肩周辺の揺らぎ
($\pm 2\sigma$ を超えたデータを削除した後の標準偏差)

特徴点	奥行方向	横方向	高さ方向
首	0.6cm	0.5cm	0.3cm
右肩	1.9cm	0.6cm	0.6cm
左肩	2.4cm	0.5cm	0.7cm

4フレーム目では左肩に約40cm以上、25フレーム目では右肩に約18cm以上のズレが生じていることが分かる。これらの原因は、4フレーム目では基準とした画像におけるOpenPoseの検出結果が、対象フレームと前後フレーム間で横方向に約10ピクセル以上ズレたことによると考えられる。また、25フレーム目ではテンプレートマッチングの際に服の柄によって誤認識が生じていた。このような外れ値を除いた全データに対する標準偏差(表1)と $\pm 2\sigma$ の区間のデータを削除した標準偏差(表2)を算出した。これより、外れ値を

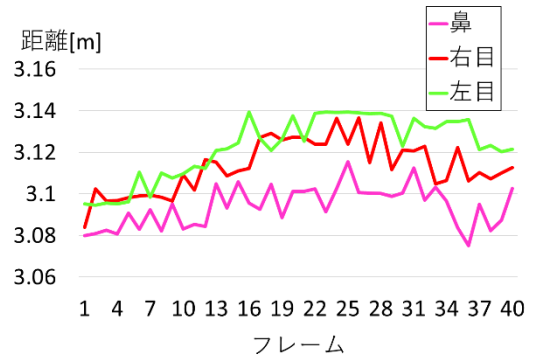


図7 頭部周辺における奥行方向の揺らぎ

表3 頭部周辺の揺らぎ
(全データの標準偏差)

特徴点	奥行方向	横方向	高さ方向
鼻	1.0cm	1.6cm	0.9cm
右目	1.2cm	0.8cm	0.4cm
左目	1.5cm	0.6cm	0.4cm

表4 頭部周辺の揺らぎ
($\pm 2\sigma$ を超えたデータを削除した後の標準偏差)

特徴点	奥行方向	横方向	高さ方向
鼻	0.9cm	1.6cm	0.9cm
右目	1.2cm	0.8cm	0.4cm
左目	1.5cm	0.5cm	0.4cm

取り除くことで標準偏差2.4cm以下となり、これは人の姿勢を推定する上で許容範囲と考える。頭部周辺に関しても同様に評価、検証を行い図7、表3、表4にまとめた。頭部周辺に関しては、 $\pm 2\sigma$ の区間において約1.5cm(センサからの距離にして0.5%)以下という高精度の結果が得られた。

4. 関心推定の半自動化に向けて

ここまで、ステレオ画像とOpenPoseによる姿勢推定について述べた。本研究の目標は子どもの関心の推定であり、特定の対象を注視するという行動が継続していることの検出が必要である。

人が見ている方向を推定するには、視線計測、頭の向き、体の向きなどの計測が必要である。このうち視線の向きは遠隔からの計測は現状では困難である。それに代わっての頭の向きは、多くの表情認識ソフトでは計測されていない。一方で、表4に示された目の位

置の揺らぎは両目の間隔 6~7cm と比較してかなり大きく、そのまま両目の位置から頭の方を推定することは難しいと考えられる。結果として、体の向きを表 2 にある左右の肩の位置から推定するというのが最も可能性が高く、その精度評価は今後の課題である。

いずれの方法を取るにせよ、本センシングシステムでの計測はオクルージョンが発生すると機能しないため、オクルージョンの発生時の計測データの評価が必要であり、実用的には全自動での推定は困難と考えられる。そのため、計測の事後に測定結果を評価してオクルージョン発生時のデータは除外あるいは補間する処理を適用する、人手による作業が必要になると考えられる。そのためのツール開発もまた課題である。現在、画像から人の行動をアノテーション（記述）するツールはいくつか市販されているが、このような機能はまだ含まれておらず、またそれを付加するソフトウェア IF が公開されているものも見当たらない。既存のアノテーションツールとの統合ができることが望ましい(6)。

5. まとめと今後の展開

本稿では、関心は動作に現れると考え、骨格検出による 3 次元再構成を行い、連続的な姿勢推定を検討した。OpenPose による骨格検出には揺らぎが生じていたが、テンプレートマッチングによる揺らぎの補正によって高精度での姿勢推定の可能性が示唆された。しかし、テンプレートマッチング後の 3 次元再構成においても大きな外れ値は生じた。そのため検出された骨格の特徴点の位置に対してフィルタ処理を行い、正しい位置を検出する必要がある。また、本稿では 1 台のステレオカメラによる検証であったが、それではオクルージョンの発生は避けられない。それに対してはカメラ複数台を利用した計測とデータの統合が必要であろう。現在、ご協力いただいている幼稚園で複数台での計測を開始した。今後は、画像処理による姿勢推定の精度向上、および複数台のカメラの統合、そのためのツール開発を行い、少ない労力・高い精度での関心推定を実現したい。

画像からの姿勢推定・関心推定がある程度の精度で実現できたなら、その利用範囲は大きい。本研究は保

育・乳幼児教育の実践の場での利用を意図しているが、高齢者ケアもまた有力な応用フィールドである。幅広くはマーケティングなどへの応用も考えられるが、関心という人の心の状態を推定する技術は個人情報の観点からは利用に注意が必要である。この点を踏まえて、慎重に応用領域を検討していきたい。

謝辞

本研究は、産業技術総合研究所人工知能研究センターからの委託研究として実施された。支援に感謝する。また、研究のサポートして下さった宮田真宏氏に感謝する。

参考文献

- (1) 山田徹志,肥田竜馬,宮田真宏,大森隆司,中村友昭,長井隆行,岡 夏樹:“子どもの関心の推定を通じた保育の質の客観化の試み”, 日本教育工学会第 33 回全国大会, 講演論文集, pp.775-776, (2017)
- (2) Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh: “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, cao2017realtime, (2017)
- (3) 山田徹志, 肥田竜馬, 宮田真宏, 大森隆司, “AI による保育研究支援システム開発に向けた予備的調査 - 子どもの関心推定を目指して-”, 人工知能学会全国大会予稿集, (2018)
- (4) 山田徹志, 浅利恭美, 宮田真宏, 中村友昭, 長井隆行, 岡夏樹, 大森隆司: “AI により子どもの発達・教育研究を支援する分析手法の検討- 子どもの位置・向き情報による関心の推定 -”, 日本教育工学会 第 32 回全国大会, pp.51-52, (2018)
- (5) 浅利恭美, 山田徹志, 宮田真宏, 大森隆司: “子どもの関心推定のための行動センシングシステムの開発”, 日本教育工学会第 34 回全国大会講演論文集, pp.679-680, (2018)
- (6) 大村廉: “ATLAYA:アノテーションと行動分析ツールの統合による行動ラベル取得労力の低減と柔軟な分析環境”, 知能と情報 Vol.28, No.6, pp.899-910, (2016)