

# 教育活動映像からの子どもの関心推定システム開発の試み

山田 徹志\*1, 浅利 恭美\*2, 青柳 燎\*3, 宮田 真宏\*2, 大森 隆司\*3

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

## Trial of child's interest estimation system development from educational activity image

Tetuji Yamada\*1, Yukimi Asari\*2, Ryo Aoyagi\*3, Masahiro Miyata\*3, Takashi Omori\*2

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

幼児教育学分野では、子どもの心の状態の記述が「子どもの育ち」を議論する指標の一つとなっている。中でも、心情・意欲・態度の基礎を培う上で「関心」は、子どもの自立能動的な学習の理解に寄与する重要な要因である。一方、心の状態（関心）の記述は現状では人（実践者・研究者）の経験知で解釈する他ない。さらに、その客観性の確保は難しく、その記述作業に膨大な時間と労力を要する。そこで我々は、人工知能を用いて子どもの教育活動映像を解析し、物理的な活動量から半自動的に個々の関心を推定するシステム開発を試みているのでその進捗を報告する。

キーワード: システム開発, 関心推定, 教育支援, 画像認識

### 1. はじめに

近年、乳幼児期からの早期段階での教育環境整備が大きな社会投資効果を有することが明らかとなっている(1)。さらにこれらの投資効果は、「非認知能力」という人生を生きる上で重要な人間性に関わる能力・個性の醸成へ影響を及ぼすことも示唆されている(2)。これを受け、乳幼児の教育の重要性が再認識され、「幼児教育（保育）の質」の向上に関する研究・調査が国内でも本格的に進められている(3)。

一方で、教科教育ではない幼児教育は、特定の教科の学習習熟度の向上から「子どもの育ちの状態」を評価することが難しい。その為、教育現場（幼稚園・認定こども園等）では、子どもの「遊び」・「生活」における姿の変容を日誌等の記録により記述し、教員が評価を行う。また、この手法は幼児教育研究分野においても同様で、多くは質的分析手法により育ちの状態（子ども個々の発達の状態）を評価することで「幼児教育の効果」の向上を計ろうとしている。これらの研究手法は、研究者や教育実践者（幼稚園教諭・保育士）の

経験知からの主観的解釈による子どもの心の状態の記述を基に行われている。この手法は幼児教育において、「遊び」・「生活」から子どもの育ちを解釈する上で有効であるが、その記述・解釈には人手による多大な作業と時間が必要であり、教育現場での事例収集にも制限がある。さらに、分析における結果は定性的なものになりがちで客観性を保つにも限界がある。

この問題に対しいくつかの幼児教育研究・実務支援技術はすでに存在している(4,5)。しかし、これらは、事前に記録されている保育映像や事務記録の保管・管理の機能が主であり、教育実践の振り返りや教育研究において保管された記録を人が見直し解釈するためのツールとして使用されているが、人手による質的分析に替わるものではない。そのため、記録されたデータの解釈はやはり人の経験知に依存し、解釈のための労力・時間は記録の量に比例して増加する傾向にある。これは、現行の幼児教育支援技術の限界である。

そこで我々は、幼児教育研究・実践を支援する新たな分析システムとして、映像情報から子どもの心の状態（関心）を機械的手段で「見える化」する分析技術

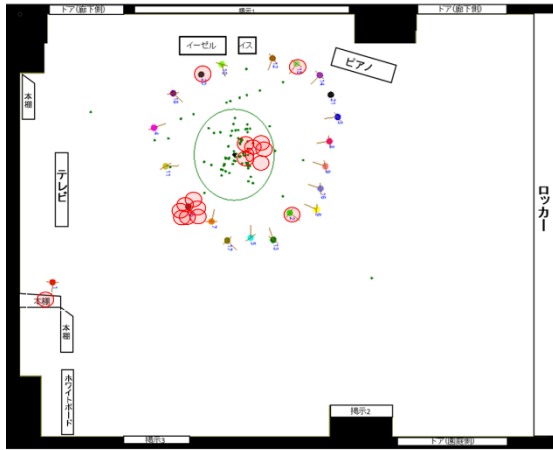


図1 ある瞬間（1秒）の個々の関心対象と  
行動特徴量（位置・向き）

を開発し、ビッグデータを用いた定量的な幼児教育支援システムの構築を目指している(6.7.8).

子ども心の状態の「見える化」に向けて、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て心の状態を読み取ることで、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(9.10.11). その中でも特に注目すべきは、子ども自立能動的な学習の基礎となる事象(物事)への「関心」という心的状態であろう. そして、これまでの研究成果から我々は、子どもの関心の推定(以下、関心推定)には子どもの「位置」・「向き」という物理的な行動特徴量が寄与し、その詳細な観察により関心推定は可能であると考え(6.7.8).

以上より、本報告は工学的な手法により教育活動映像から子どもの動作をセンシングし関心推定を行うシステム構築についての研究の進捗を報告し、本システムの実現の可能性について検討していく.

## 2. 子どもの関心推定システム

### 2.1 子どもの関心と行動特徴量（位置・向き）の関係

本研究では、人の感覚知による子どもの関心状態の評価の結果として、そのための定量的な指標として子どもの頭と体の位置・向きに着目している(参考文献). 我々はこれまで、子どもの関心状態の評価のため複数の保育専門職者によるのべ約 1500 時間のアノテーション(記述)を実施してきた. その結果、子どもが関心に向け対象に主体的に関わろうとするときには、関心対象と子どもの行動特徴量(本研究では位置・向き)

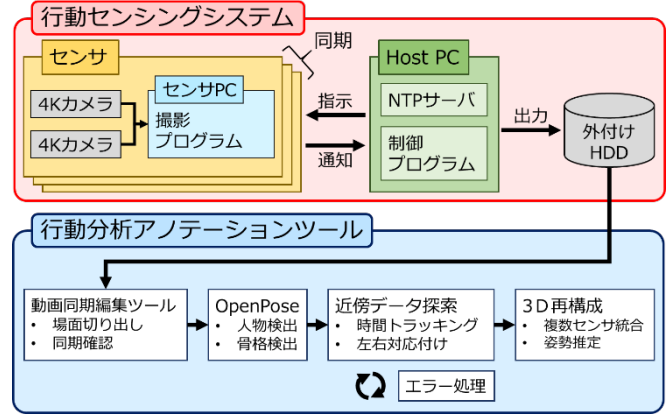


図2 関心状態推定システムの全体像

の間に強い関連があることが示された. 例えば、関心対象の方を向いて近くまで移動する、関心対象をじっと見つめる、という場面がそれにあたる.

図1は、観測・記述した教育活動データ 18分53秒におけるある瞬間の幼児 18名の関心対象と位置・向きの可視化した結果である. 図中で、子どもの関心対象と位置・向きは以下のように示されている.

- ・異なる色で示された点：子どもの位置.
- ・点より延びる線：子どもの頭(前方)と(両肩方向)の向き.
- ・赤色の円：関心対象の位置
- ・緑の点：子どもたちの頭の向きを直線として伸ばした場合の交点.
- ・緑の円：頭の向き方向に伸ばした直線の交点群の標準偏差 $\sigma=1$ の広がり.

このとき、ある一人の幼児(注1、以下A児)を除き、子ども達は円形に立って円の中心方向を向いている. 頭の向き方向に伸ばした直線の交点は円の中心部に集まり、その広がり標準偏差 $\sigma=1$ は小さく、子どもたちの輪の中に納まっている. そして関心対象は、A児を除いては輪の中心・保育者・仲間の子どもの示しており、子ども達が形成する輪の中に収束している. 一方で、この時のA児の関心対象は本棚である.

この場面は「輪になり次の活動を待つ」教育活動場面であり、教員はこの時は子どもたちに輪になるよう指示している. その為、記述された子どもたちの関心は、輪を作る全体の活動と教示を行う教員へ向いていた. これより、「輪になる」と「教員」という関心対象が多数となった一方で、A児は全体の関心傾向とは異なり本棚の絵本へ関心を向けていている.

記述された子どもの関心状態を関心対象に対する位置・向きと比較すると、両者の間には強い関係があることが推測される。これより我々は、全体の関心傾向や個人の関心傾向もまた位置・向きから推定可能と考える。そして、位置と向きという行動特徴量は、現在の工学技術を用いれば行動センシング環境を整備することで抽出可能であると考えられる。

## 2.2 関心推定システムの全体像

関心推定システムの開発は大きく「行動センシング」と「行動分析」の二つの側面で行われている（図 2）。図 2 上部の行動センシングシステムでは、まず教育活動映像を複数台のステレオ 4K カメラ（現在は教室に 3 台、4K カメラ 6 台を使用）を時間同期させて取得する。複数台で異なる方向から観察することで、人物の重なり（オクルージョン）による観察ミスを減少させる。取得された映像はセンシング装置内部のセンサ PC に一時保存され、観察終了後に HostPC を通じて外部のディスクに保存される。

図 2 下部の行動分析システムではまず、取得された教育活動の映像について、全フレームの時間同期の確認と必要な場面を切り出すカット編集を行う。その後、骨格検出ソフト OpenPose (12) により映像内の子どもの骨格を検出する。この段階で、抽出された個々の骨格について左右の映像で対応付けすることで、個々の骨格情報の 3 次元再構成が可能となる。これより、教育活動映像から関心推定に必要な子ども個々の位置・向きの検出が実現され、関心推定が可能となると期待できる。本稿では以降、関心推定システムの試作とその予備的観察による子ども個々の位置・向き情報の 3 次元再構成の結果について報告する。

## 3. 関心推定システム実現に向けた試み

### 3.1 行動センシングシステムの開発

開発を試みたセンシングシステムは、左右 2 台の 4K カメラ (Logitech 社製) を搭載した行動センサ装置で、教育活動の 4K 映像を 10fps で記録できる。制御用 PC には Linux 系 OS を使用し、HostPC を介した NTP (Network Time Protocol) で時間同期したことで、センサ装置の間では 10 ミリ秒以内の精度での同期を実

現した。複数のセンサ装置は有線 LAN を介して Host PC に接続している。

取得される教育活動場面の映像データは、センサ装置 1 機につき 1 時間の計測で約 54GB になる。さらに 1 つの教室にセンサ装置 3 台を用いるため、取得データの総量は 1 時間あたり約 162GB に及ぶ。

このデータは計測後に各センサ装置から FTP (File Transfer Protocol) を用いて HostPC に接続されたディスク装置に自動転送される。そのため、Host PC には FTP サーバを構築した。なお、データ転送に際して各センサ装置ではまず映像データを圧縮してデータ量を削減し、転送後にはセンサ装置内と Host PC の双方で圧縮データのハッシュ値を比較して転送の確実性を確認する。以上より、複数台の 4K カメラから安定的に教育活動映像を記録することが可能となった。

### 3.2 骨格情報からの子どもの位置・向きの検出

取得した映像データから子どもの関心を推定するには子どもの位置・向き情報の抽出が必要である。そこで本研究では、ステレオ画像処理により人物の 3 次元姿勢を連続的に検出する方法を試みる。そのため、2 次元画像から画像中の人物の骨格を検出する人工知能ソフト OpenPose を用いて、身体画像からの身体的再構成を目指した。本稿では、成人 4 人を対象とした 6 秒間のステレオ映像を題材に、人物骨格の 3 次元再構成による姿勢推定の結果を報告する。

OpenPose による画像中の人物骨格の検出は、1 画像中に複数人がいる場合には個々の人物の骨格情報の出力が順不同となり、ステレオの左右画像からの出力はそのままでは対応せず 3 次元再構成は行えない。そのため、各人物の画像中の骨格位置のずれ量より左右画像で対応する骨格を検出して両眼視差を抽出し、骨格の各関節の 3 次元位置を計算して身体姿勢の再構成を行った。

検証として、成人 3 名に対してステレオカメラ 1 台を用いて計測し、OpenPose の検出結果より 3 次元再構成を行った。その結果、ある程度の精度で 3 次元再構成が実現できた（図 3）。しかし、手足の位置が細かく揺らぎ、奥行方向への大きなズレが生じた。これは、OpenPose による骨格検出に揺らぎが生じていたことによる。この問題に対し、骨格検出後に画像処理手法

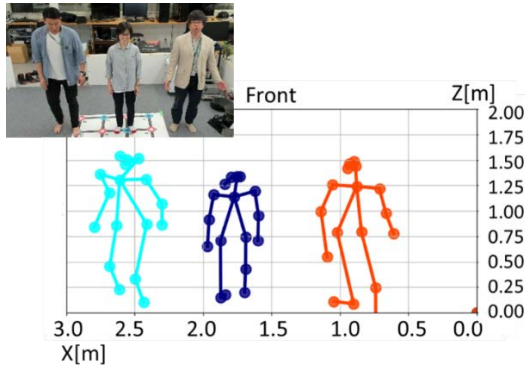


図 3 計測場面および 3 次元再構成結果

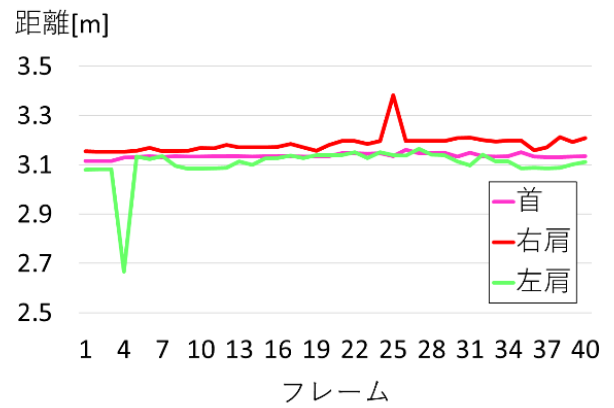


図 5 テンプレートマッチングを用いた際の肩周辺における奥行方向の揺らぎ

表 1 肩周辺の揺らぎの標準偏差 (OpenPose による 3 次元再構成)

特徴点	奥行方向	横方向	高さ方向
首	5cm	0.6cm	1.7cm
右肩	5cm	1.2cm	1.3cm
左肩	7.5cm	0.7cm	2.2cm

表 2 肩周辺の揺らぎの標準偏差 (テンプレートマッチング後の 3 次元再構成)

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

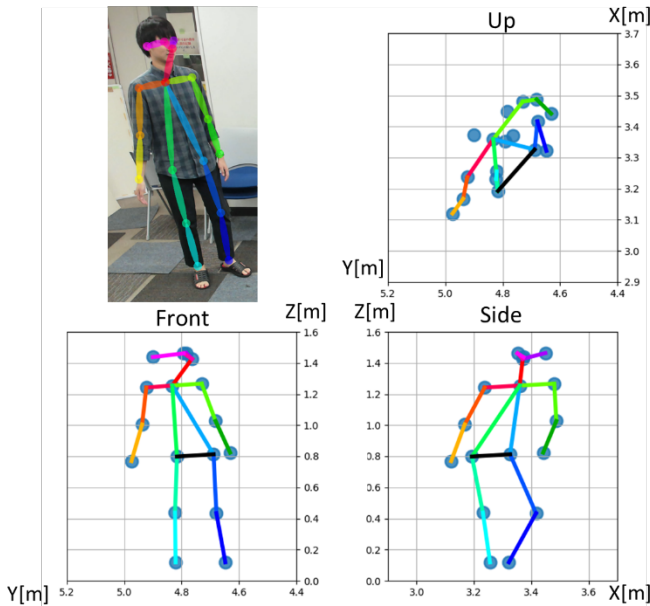


図 4 OpenPose による骨格検出およびテンプレートマッチングによる 3 次元再構成結果

であるテンプレートマッチングを用いた揺らぎの補正を試みた。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の対応する位置を探索する手法である。この手法を用いて改めて関節の対応点の探索を行い、成人 1 名のある瞬間に対する 3 次元再構成を行った結果、姿勢推定の誤差が大きく減少した (図 4)。

テンプレートマッチングの効果の検証のため、4 秒間 (40 フレーム) の静止状態の人物の 3 次元再構成を行った。評価には、比較的安定して検出できる首、右肩、左肩の特徴点の奥行方向の揺らぎの傾向を見た (図 5)。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 の骨格検出が、対象フレームと前後フレーム間で横方向に約 10 ピクセル以上ゆらいだためと考

えられる。また、25 フレーム目では、人物の服の柄 (輝度の差分) の均一性のためにテンプレートマッチングで誤検出が生じていた。これらの外れ値を検出・除去するため、奥行きデータの標準偏差を算出した (表 1)。テンプレートマッチング後の結果については、 $\pm 2\sigma$  を超えたデータを削除した標準偏差を算出した (表 2)。その結果、外れ値を取り除くことで標準偏差 2.4cm 以下となり、これは人物の姿勢を推定し、位置・向き情報を抽出するのに十分な範囲の誤差だと考える。

以上より、記録した教育活動映像データからの関心推定分析の要となる子どもの位置・向き情報の自動抽出は、骨格検出後に 3 次元再構成の揺らぎ補正することで可能であると考えられる。

## 4. センサ装置開発と分析手法の確立

### 4.1 行動センサ装置の開発と定点観測地の確保

行動センサ装置を用いて教育現場での定点観測を行うため、行動センシングシステムの各種機器を格納した行動センサ装置を企業と共同開発した(写真1)。本行動センサ装置を複数台、幼稚園・保育園の壁に設置して4K解像度のステレオ映像で日々の教育活動を計測・記録する体制を整えた。さらに現在、定点観測地点を拡大し、小学校を含めた複数の教育施設の教室及び保育室への設置を進めている。対象の幼稚園から小学校には多くの子どもが進学しており、縦断の関係にある多様な教育活動のデータ収集が期待できる。

### 4.2 位置・向き情報の評価とオクルージョン問題

抽出した骨格左右の肩の位置が先述のテンプレートマッチングで正しく推定できた場合には、その人物の位置と体の向きが少ない誤差で推定できる。これより、条件が良いときには関心推定に必要な位置・向きの情報は機械的な手段で獲得できると期待できる。

しかし、ステレオ画像からの3次元再構成には避けられない問題がある。一つはオクルージョンによる左右の対応部位の遮蔽、もう一つは均質な画像テクスチャによる画像の対応検出の誤りである。OpenPoseはたとえオクルージョンで人の身体が見えなくとも身体骨格を推定する強力な道具であるが、その推定精度は決して高くない。そのため、左右の対応付け時にオクルージョンや均質画像の判定を行ない、骨格の位置情報の信頼性を評価して問題のある推定値を排除する手法の開発が必要である。

この問題を軽減するため、本研究では複数の行動センサ装置からの3次元再構成情報を統合して使用することを考える。幼稚園などでの教育活動では多くの子どもが集まるためオクルージョンが多発し、完全に解消はできない可能性は高いが、その程度を軽くすることを目指す。どの程度の推定精度があれば関心推定などの利用に資するかは、今後の評価が必要である。

欠損データの後処理として、欠損箇所を人が修正し補完できるアノテーションソフトウェアの開発もまた課題である。そもそも、最新の人工知能やICT技術でも100%の精度の姿勢推定は実現できず、人手による



写真1 開発した行動センサ装置試作機。

カメラマウント部分は上下左右に画角を調整できる。

補完が欠かせない。現在、画像から人の行動をアノテーションするツールは複数あるが、ステレオ画像処理による3次元再構成のような機能はいまだ無い。また、そのような機能を付加できるソフトウェアIFが公開されたアノテーションツールも見当たらない。アノテーションツールのシステム開発の費用・労力を考慮すれば、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(13)。

## 5. 今後の課題と発展

### 5.1 関心推定システムの開発課題

今後の主要な課題は、複数の定点観測地における行動センサ装置の運用と分析システムによる子どもの行動特徴量(位置・向き)の抽出評価である。

行動センサ装置については、社会実装を見据え適切かつ安定的に多様な教育活動の映像を取得できることを確認していく必要がある。この時、子どもの個人情報が含まれる可能性があるため、暗号化等により安全に管理する手法についても検討を要する。また、装置自体のコストの削減もまた普及のための課題である。なお、多様な教育施設での運用を目指し、遠隔操作による行動センサ装置の稼働実験も継続していきたい。

分析システムについては、子どもの位置・向き情報を高い精度で検出する分析手法の確立が望まれる。そのためには、骨格情報のオクルージョン問題を解決する画像処理アルゴリズムの開発が必要である。今後開発するアノテーションツールで、如何にして人の作業



を軽減し、半自動化処理を実現していくか、本システムの実現に重要な課題であろう。

## 5.2 社会実装に向けた議論

近年の AI 技術の進展は急速であり、人の活動を高精度で認識する技術はいずれ実現されよう。教育活動場面の高度かつ長時間の記録分析もまた可能になると予想される。そうなれば、幼児教育・保育の分野でも新たな科学的発見や保育者の作業軽減、専門性の向上といった効果が期待できる。本研究はその初期の試みと位置付けられよう。

映像からの行動特徴量としての位置・向き の計測を通して関心推定が可能であるなら、教育実践現場において新しい教育支援システムの開発に繋がるだろう。特に、AI, IoT, ビッグデータは急速に我々の生活に浸透し発展し続けている。本研究は、これらの技術発展に伴い、これまで教員や研究者が行ってきた育ちの記述の作業を半自動化できる近未来を目指している。

例えば、特徴的な発達傾向を示す子どもの関心や個性の推定、さらに、教員間での関心の読み取り技能の共有や教育実践現場における子ども育ちの振り返り資料への活用など、子どもとの関わりや教員同士の相互理解を深めながら「教育の質」を議論する定量的資料がより少ない労力で得られる可能性がある。

しかし、本研究が目指す子どもの関心推定技術が実用的な教育支援システムとして社会実装されるためには、教育実践現場での有用性についての議論が求められよう。その為には、教育実践者・研究者と共にこれを如何に活用して教育の質を向上させるか、その方策と効果の議論が必要である。また、このような新しい技術の確立と実装は、社会に対し想定以上の大きな影響力を持つ可能性がある。新技術により得られる結果の解釈を誤れば、子どもの育ちの安易な評価や教員の職務評価も起こりうる。それを防ぐには、システムが現実のものとなった後での議論では手遅れであろう。そのため、技術開発のみではなく実装する教育領域の将来の姿を見据えた開発が必要である。

## 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。

## 参考文献

- (1) OECD Starting Strong 2017 Key OECD Indicators on Early Childhood Education and Care OECD (2017)
- (2) Heckman, James J. and Tim Kautz, "Fostering and Measuring Skills: Interventions that Improve Character and Cognition" NBER Working Paper Series 19656Zhe (2013)
- (3) 秋田喜代美 淀川裕美(訳)『「保育プロセスの質」評価スケール』 明石書店 (2016)
- (4) 刑部育子, 戸田真志, 植村朋弘, 佐伯胖 : 観察中の「瞬間リフレクション」記録・分析のためのツール開発, 日本認知科学会 第 26 回研究大会 研究大会発表論文集 pp3-5 (2009)
- (5) MKI 三井情報株式会社 きっずノート  
<https://www.mki.co.jp/solution/kidsnote.html> (2019.2.4)
- (6) 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 - 日本人工知能学会 第 32 回大会研究大予稿集 103-OS-15b-03 (2018)
- (7) 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)
- (8) 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)
- (9) 文部科学省 『幼稚園教育要領』(2017)
- (10) 厚生労働省 『保育所保育指針』(2017)
- (11) 内閣府 『認定こども園 保育・教育要領』(2017)
- (12) Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh: "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", cao2017realtime, (2017)
- (13) 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)