

## 顔情報を用いた授業活動の自動セグメンテーションの試み

### An Attempt at Automatic Segmentation of Class Activity Using Facial Information

宮田 真宏<sup>\*1</sup>, 山田 徹志<sup>\*2</sup>, 大森 隆司<sup>\*2</sup>  
Masahiro MIYATA<sup>\*1</sup>, Tetsuji TAMADA<sup>\*2</sup>, Takashi Omori<sup>\*2</sup>

<sup>\*1</sup> 武蔵野大学

<sup>\*1</sup> Musashino University

<sup>\*2</sup> 玉川大学

<sup>\*2</sup> Tamagawa University

Email: ma-miyata@musashino-u.ac.jp

**あらまし:**本研究ではこれまで、教室空間にカメラなどのセンシング機器を設置し授業中の様子を計測し、その結果を分析して教員にフィードバックする振り返り支援システムを開発してきた。そこでは講義場面などの特定の場面に対して AI 技術を適用することで子どもたちの授業参加が推定できることを示した。また、授業中に子どもたちが取る行動は授業活動ごとに異なるため、授業参加を示す行動はそれぞれ異なることがわかってきた。これを踏まえて本研究では、個々の活動場面における授業参加度の推定のような行動解釈をする前に、授業活動を自動分類することが必要であると考えた。そこで本発表では、自動分類のための特徴量として顔の検出数に着目し、これを用いた授業内活動の自動分類の可能性を示し、今後の展開について議論する。

**キーワード:**人工知能の教育的応用, 教員支援, 広域計測

#### 1. はじめに

本研究ではこれまで、日々教員が行う授業の振り返りを AI 技術により支援することを目的とした支援システムを開発してきた[1]。そこでは教室の前後に 4K 解像度のステレオカメラを各 1 台ずつ設置し、授業映像を記録した。映像の分析には、顔認識 AI[2]を用いて子ども達の顔の特徴量を抽出し、三角法により個々の子どもの教室内の位置と向きの情報に変換した上で、講義場面における個々の子どもの授業への参加状態を推定してきた[3]。

一方、教員が子どもに対して行う授業は講義だけでなく、アクティブラーニングなどの多様な活動が導入されている。そのため、現在の 1 回の授業には異なる複数の活動が含まれ、それぞれの活動で子ども達が取る行動は異なる。AI などを活用した個別最適化の対応を子どもに取るためには個々の授業活動に対応する分析が必要であるが、その前提となる授業活動の自動分類手法は現時点では未確立である。

そこで本発表では、この授業活動を自動分類するための手法として、顔認識 AI により授業映像より認識された顔の検出数を用いる手法を提案する。そしてその評価として、人手による活動記述と比較することで、提案手法の可能性を示す。

#### 2. 行動計測

行動計測には、LUXONIS 社の OAK-D OpenCV DepthAI カメラ (PoE 版) を使用した。映像は、カラー画像を 4K 解像度 (3840×2160) とし、視差画像を HD 解像度 (1280×720) とした上で 15fps でこれらを同期計測した。本カメラは教室の前方に設置

してある黒板の両端上部から子ども達の方を向けて 2 台と、概ねそれぞれの前方カメラの正面にある教室後方の壁から教室の正面方向に 2 台の合計 4 台を設置した (図 1)。

本稿で分析に使用したデータは、2022 年 10 月 17 日に玉川学園小学部にて行われた小学 2 年生 31 名の算数の授業 (46 分 32 秒) の映像である。

#### 3. 顔の検出数と授業活動との関係

本研究では多くの子どもたちが授業中に行う行動は、教員の指示からその意図を読み取り、それに対応するものであると考えた。そして子ども達は教員が説明している場面であれば教員の方向を向き、ノートを取る場面であれば黒板とノートを一定の時間毎に繰り返し見ると考えられ、グループワークであればグループのメンバーの方を向くと考えられる。とするならば、このような子ども達の取る行動は教



図 1 教室内の 4 方向からの授業計測、および顔認識 AI の検出例

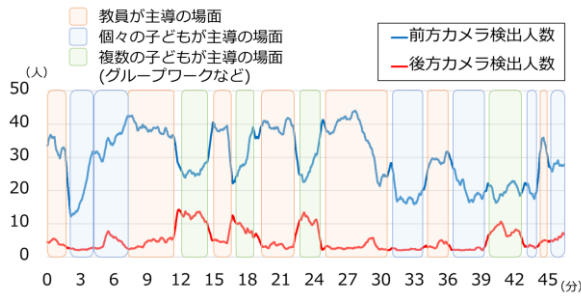


図2 教室の前方／後方カメラ(各2台)に  
おける顔の合計検出数と授業活動

室中に設置した4台の各カメラに映る顔の数として現れると考えた。

そこで分析には、それぞれのカメラ映像に顔認識AIを適用して得られたフレーム毎の顔の検出数を特徴量として採用した。そしてカメラを教室の前方グループと後方グループの2つに分けて、それぞれの合計検出人数を算出した。その後、前後1.5秒分のデータに対して中央値フィルターを取り、1秒当たり1つのデータに間引いた後に、30秒幅の移動平均処理をした(図2)。図2は横軸に時刻を縦軸に顔の検出数を取ったものである。図中の青線は前方カメラの検出数を、赤線は後方カメラの検出数をそれぞれ示している。実際に授業中で起こっていた場面を教員が主導して説明や問いかけをしている場面、個々の子どもが板書の筆記や問題を解くなどの作業をしている場面、子どもが複数人でグループワークをしている場面をそれぞれ図中の橙色、水色、黄緑色の枠で示した。

図2の橙色の枠内では前方カメラの検出数が多く、後方カメラの検出数が少ない傾向があった。水色の枠内では前方カメラの検出数は橙色の枠内より少ないが後方カメラの検出数は変化しない傾向があった。また黄緑色の枠内では、前方カメラの検出数は中程度であるが、後方カメラの検出数は他の枠内よりも多かった。これより、授業中の子ども達の顔の向き、ひいては前方カメラと後方カメラの顔の検出数はその瞬間の授業活動を反映すると考えられる。

これを踏まえて図3にこのデータに対して機械学習手法のクラスタリング手法の一つであるK-means法を適用して3つのクラスに分類した結果を示す。図3の横軸には前方カメラの検出数を、縦軸に後方カメラの検出数を取り各クラスの分類結果を散布図で示した。図3の各点は授業内の各時刻に対応し、その軌道はその時系列を示している。図3の結果は、概ね人手により解釈された場面の分類と一致していたことからK-means法を用いることで場面の自動的な分類の可能性を示す結果であると言える。一方で、本稿で示した結果は他のクラスや科目においても同様に分類できるかについては疑問が残る。この点については今後、計測のサンプルを増やしての検証が必要である。

なお、本稿で示した結果は教室中央に座っている

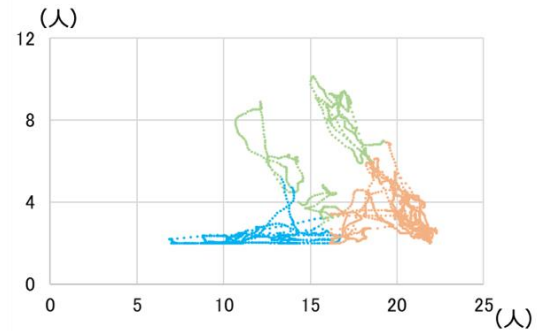


図3 K-means法による  
授業活動の自動分類の例

子どもを中心に、複数のカメラで重複して検出・カウントしたデータによる分析である。このような検出法が結果に及ぼす影響の評価は今後の課題である。

#### 4. まとめ

本稿では、教室の前方と後方にカメラを各2台設置して教室全体の授業映像を記録し、顔認識AIを用いて映像から認識される顔の検出数を数え、そこから授業活動の大きなパターンを分類できる可能性が示された。その検出数を用いて機械学習手法の一つであるK-means法により場面分類を試みた所、人手による場面分類と類似する結果を得た。一方で、人手による解釈において違和感の残る分類がされている箇所も残っている。その点については確率分布に従った分類手法などによる評価も必要と考えるが、今後の課題とする。

本手法で同様に計測した中学2年生の授業においても同様に分類できることを示唆する結果が得られた。これより、分析手法としてもある程度の妥当性があると考えられる。今後はより多くの授業のデータを収集し、さらなる妥当性の検証をする。

本研究は、科研費 若手研究 22K13762, 基盤研究(B)19H01718, NEC バイオメトリクス研究所との共同研究, 及びキャノン財団からの研究支援により実施された。支援に感謝する。また本研究は、玉川大学研究倫理委員会及び保護者の承認のもと実施された。

#### 参考文献

- (1) 宮田真宏, 山田徹志, 大森隆司: “顔情報を用いた授業中の子どもの特性リフレクションシステムの開発”, 教育システム情報学会2021年度第2回研究会, Vol.32, No.2, pp.6-13 (2021)
- (2) Takamoto et al. “An Efficient Method of Training Small Models for Regression Problems with Knowledge Distillation”, 2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), pp.67-72 (2020)
- (3) Mayuko Maruyama, Masahiro Miyata, Tetsuji Yamada, Takeshi Aihara, Takashi Omori, “Estimating children's classroom participation using machine learning methods”, NCSP2023, (2023)