

姿勢推定技術を活用した学習者の集中度の計測 Measurement of learner concentration using pose estimation

田上 和輝[†] 野田 昂弥[†] 春日 秀雄[†]
Kazuki Tagami Takaya Noda Hideo Kasuga

1. はじめに

学校の授業において、学習者の行動を把握することは、教育効果の向上において重要である。たとえば、学習態度と成績は密接な関係にあるとされており、服部らの研究^[1]では授業中に居眠りをしていない学習者が居眠りをしている学習者よりも成績が良好であるという結果を得ている。教員が居眠りやよそ見などを行っている学習者をリアルタイムに把握することで、必要に応じて注意を与えるなどして効果的な授業が可能になる。

しかし、学習者の人数が多い場合、教師が学習者全員の行動を把握することは難しくなる。近年では、オンラインでの授業やテストなどの教師が対面していない状況下で活用される教育支援システムがいくつも考案されている。その中には、PC 内蔵カメラを用いて学習者の行動を把握する手法も考案されている^[2]。オンラインの教育支援システムでは章らの研究^[3]のように、学習者の視線方向を推定するものが多い。しかし、対面形式の授業で教室全体の学習者の行動を把握しようとした場合、1 台のカメラで複数人の学習者を撮影する必要があるため、個々の学習者の高品質な目領域の検出は困難である。

そこで本研究では、視線方向ではなく顔の向きから学習者の集中度を計測する。教室の様子をカメラで撮影し、学習者のそれぞれの顔の向きから教室全体の集中度を計測するシステムを開発する。顔の向きの推定には、姿勢推定技術と機械学習を用いる。

2. 学習者の集中度計測システム

教室内のすべての学習者の様子を計測するために、本研究では複数のカメラからの入力をリアルタイムに処理するシステムを構築した。計測した教室内の学習者の集中度はグラフで可視化している。グラフは、教室内の学習者で集中している者の割合の時間変化を表す。また、集中していないと判定された学習者を画像中でマークする機能も実装している。本システムの実行画面を図 1 に示す。

本研究では、学習者の集中度を計測するのに顔の向きを用いる。学習者が正面を向いているときは集中していると考え、正面以外を向いているときは集中していないと考える。顔の向きは、姿勢推定技術で得られた姿勢データからニューラルネットワークを用いて推定する。

2.1 複数の視点からの映像の取得

カメラが 1 台だけの場合、死角に入る学生が多くなる。そこで、本システムでは複数のカメラからの映像を使用して学習者の顔の向きを推定する。ただし、複数のカメラで同じ学習者を処理しないように、映像の使用範囲を制限する機能を実装している。カメラごとに任意の範囲を設定す

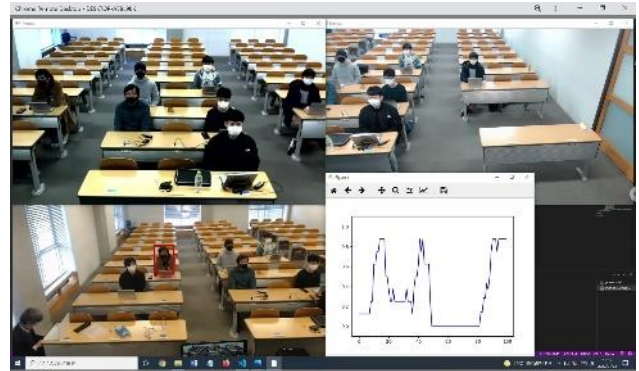


図 1 本システムの実行画面

ることができ、指定された範囲内の人物を対象に顔の向きの推定を行う。

2.2 顔の向きの推定

学習者の姿勢は姿勢推定ライブラリの OpenPose^[4]を用いて取得する。OpenPose は 25 箇所関節等の 2 次元座標をキーポイントとして取得することができるライブラリである。得られたキーポイントを用いて顔の向きを「正面」、「上向き」、「下向き」、「右向き」、「左向き」の 5 方向に識別する。識別にはニューラルネットワークを用いる。

2.2.1 キーポイントの選定

着席している学習者をカメラで撮影した場合、下半身ははっきりと映っていない場合が多い。また、取得できるキーポイントには顔の向きの推定に向いていない部位もある。そこで、どのキーポイントの組み合わせを用いた場合に顔の向きの推定精度が高くなるのかを調査した。比較した組み合わせは以下のようになる。

- 上半身の 12 点のキーポイント
- 上半身から腕を除いた 8 点のキーポイント
- 上半身から腕、肩を除いた 6 点のキーポイント
- 首より上の 5 点のキーポイント

この 4 つの組み合わせを比較した時、B で最も高い結果が得られた。よって、本研究では「鼻」、「首」、「右肩」、「左肩」、「右目」、「左目」、「右耳」、「左耳」の 8 か所のキーポイントを使用して顔の向きを推定する。

2.2.2 ニューラルネットワークによる推定

入力層、中間層、出力層の 3 層の全結合ニューラルネットワークを用いて顔の向きを 5 方向に推定する。入力層は上半身から腕を除いた 8 点のキーポイントの X 座標、Y 座標である。ニューラルネットワークの学習データとして、各方向を向いている人物の映像を撮影し、OpenPose でキーポイントを取得後に、その位置と大きさを正規化したデータを用意した。学習データには約 17,000 個のデータを用いた。

[†] 神奈川工科大学 Kanagawa Institute of Technology

3. 実験

3.1 顔の向きの判定精度

顔の向きの判定精度を確認するために、広さや机の配置が異なる 3 つの教室で実験を行った。教室の広さ、着席位置の傾向、実験参加者の人数は、表 1 のようになっている。

実験では、教師役の実験監督者の指示で全実験参加者が「正面を向く」、「右を向く」、「左を向く」、「下を向く」、「上を向く」といった動作を数秒間行った。判定精度として、正面を向いている時に「正面」と判定された人数の割合、そして、正面を向いていない時に「正面以外」と判定された人数の割合を求めた。その結果は、表 2 のようになる。

教室 1、教室 2 は人数も少なく、狭い教室または前方に着席していたため、高い精度で正面を向いているかそうでないかを判別できている。しかし、人数が多く広い教室の後方に多くの人が着席した教室 3 では、低い判定精度となった。これは、後ろの席に座っている人物の解像度が低いことが原因のひとつと考えられる。低解像度の画像では姿勢推定の精度が下がり、顔の向きの推定も失敗することが多くなる。また、教室 3 では人数が多く、密集している人物の身体が重なっていることが多かった。身体の一部が見えない場合にも姿勢推定の精度が下がり、顔の向きの推定も失敗したと考えられる。

3.2 処理時間の比較

本システムは、リアルタイムでカメラからの映像を処理するシステムである。そこで、その 1 回の処理に要する時間を計測する実験を行った。本実験で用いたハードウェアの仕様は表 3 のようになっている。OpenPose による姿勢推定、および、ニューラルネットワークによる顔の向きの推定は GPU で計算を行っている。

基準となる条件を人数「1 名」、カメラの解像度「1920×1080」、カメラの台数「1 台」として、人数を増やした場合、カメラの解像度を下げた場合、カメラの台数を増やした場合の処理時間を計測した。計測結果は、表 4 のようになる。

この実験の結果、人数が増えた場合でも処理時間はあまり増えないことが分かった。また、カメラの解像度を下げると処理時間は大きく減少することが分かった。しかし、

表 1 実験に用いた教室

	広さ	着席傾向	人数
教室 1	狭い	万遍なく散らばって着席	9 人
教室 2	広い	前方に多く着席	9 人
教室 3	広い	後方に多く着席	32 人

表 2 顔の向きの判定精度

教室	顔の向き	精度
教室 1	正面	95.9%
教室 1	正面以外	91.7%
教室 2	正面	95.2%
教室 2	正面以外	98.9%
教室 3	正面	63.6%
教室 3	正面以外	90.2%

表 3 実験機器

OS	Windows 10 Education
CPU	Intel® Core™ i5-10400 CPU @ 2.90GHz
GPU	NVIDIA GeForce RTX 2070
メモリ	32.0GB

表 4 処理時間

	条件	処理時間
人数	1 人	0.13s
	15 人	0.14s
カメラの解像度	960×540	0.09s
	1920×1080	0.13s
カメラの台数	1 台	0.13s
	2 台	0.19s
	3 台	0.27s

解像度を下げると姿勢推定の精度が下がり、顔の向きの判定に失敗する可能性が増える。そのため、あまりカメラの解像度を下げたわけにはいかないと考えられる。カメラの台数を増やした場合も処理時間は増加するが、3 台に増やした場合でも単純に 3 倍には増えないという結果となった。カメラ 3 台でも毎秒 3 回以上の処理が可能であり、リアルタイムで学習者の集中度を測定するのに問題ない速度であると考えられる。

4. おわりに

本研究では、対面授業における教室内の学習者の集中度を計測するシステムを作成した。集中度は顔の向きで判定し、顔の向きの推定は十分な解像度の画像なら 90%ほどの精度となることが確認できた。また、本システムはカメラを 3 台接続して使用した場合でも、毎秒 3 回以上の処理が可能であり、リアルタイムでの運用に問題ないことが確認された。

本システムを用いることで教室全体の学習者の集中度がひと目で把握できるようになり、効果的な授業の運営が可能となると考えられる。また、本システムでは教室内の学習者の集中度の時間変化を記録することが可能になっている。この学習者の集中度のログは、授業内容を分析するデータとして利用できる可能性がある。また、授業評価に利用することで、授業の質の改善などにも活用できる可能性があると考えられる。

参考文献

- [1] 服部 伸一, 野々上 敬子, 多田 賢代, “中学生の授業中の居眠りと学業成績, 自覚症状及び生活時間との関連について”, 学校保健研究 54, pp. 305-310, (2010).
- [2] 安彦 智史, 池辺 正典, 丸山 広, 長谷川 大, “PC 内蔵カメラを用いた学習態度把握方式の検討”, 情報教育シンポジウム 2015 論文集, pp. 103-108, (2015).
- [3] 章 毓峰, 田和辻 可昌, 松井 辰則, “視線計測を用いた学習者の集中力の変化の検出と Robot による集中力回復の試み”, 教育システム情報学会 2018 年度学生研究発表会, (2019).
- [4] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, Yaster Sheikh, “Openpose: realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.7291-7299, (2017).