

令和 5 年 6 月 22 日現在

機関番号：18001

研究種目：基盤研究(C)（一般）

研究期間：2018～2022

課題番号：18K02901

研究課題名（和文）映像解析による人物の集中度推定の向上に関する研究

研究課題名（英文）Improvement of Personal Concentration Estimation by Image Analysis

研究代表者

姜 東植（KANG, Dongshik）

琉球大学・工学部・准教授

研究者番号：00315459

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究では、教師の授業改善を支援するため、授業中に撮影された映像から生徒の姿勢や動きから集中度を推定し、授業について客観的に評価する方法を模索したものである。具体的には、まず深層学習を有するニューラルネットワークにより授業映像から生徒の姿勢や動き（動作）の候補領域を検出し、その候補領域の映像を6つの動作にラベリングを行った。次に、時系列のディスプレイ信号（動作譜）を生成することで、集中度の時間変異を考慮した推定を可能にした。また、識別器により3つの状態（集中、曖昧、退屈）にクラス化を行い、可視化を実現した。さらに、各生徒の動作譜から相関関係を求め、学習集団がもたらす個人への影響を推定した。

研究成果の学術的意義や社会的意義

本研究で提案したシステムは、教室前方に設置した1台のカメラで撮影した映像を入力とし、映像から人物を検出し、検出された人物一人一人に対し学習時の姿勢や動きの度合いを時系列化することで集中度の変化を可視化することが可能である。また、これらをラベリングにより数値化することで、授業時間の経過において、どの時刻からどの程度時間を集中していたかをスコアで表すことが可能になった。授業の時間内に生徒がどのタイミングで、どれ位の集中時間が持続できたかを知ることが大変重要な要素である。本提案システムは教員が自らの授業を振り返るきっかけを提供することができ、授業改善に役立てていただけると考えている。

研究成果の概要（英文）：In this research, in order to support teachers in improving their lessons, we explored a method of objectively evaluating the lesson by estimating the degree of concentration from the posture and movement of the students from the video taken during the lesson. Specifically, we first detected candidate regions for student postures and movements (movements) from class videos using a neural network with deep learning. Next, by generating time-series discrete signals (movements), we enabled estimation considering the temporal variation of the degree of concentration. In addition, classifying into three states (concentration, vagueness, boredom) was performed using the NN, and visualization was realized. Finally, we calculated the correlation from each student's movement score, and estimated the influence of the learning group on the individual.

研究分野：教育工学

キーワード：集中度推定 授業映像 学習姿勢 動作譜 深層学習 候補領域

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

近年、学力向上のための授業改善の取り組みが様々な教育機関で重要視されている。授業には、ある程度、決まった進め方(パターン)はあるが、絶対的な進め方は存在しない。教師の説明や発問、話し合い活動、思考を整理する活動など、その時々授業のなかで、有機的に生かされているかが授業の善し悪しを決める。したがって、授業の時間内に生徒がどのタイミングで、どれ位の集中時間が持続できたかを知ることが大変重要な要素である。

一方、現行の集中度の計測方法としては、授業中に撮影された映像から生徒の姿勢や動きから集中度を推定し、客観的に評価を行う方法が模索されている。しかしながら、これらは特別なセンサや授業時間全体の集中度を推定するものであり、授業中に集中度が、どの程度、どのタイミングで時系列的に変異しているか測定することは困難であるため、実現において妨げになっている。

このことを踏まえ、本研究では情報処理技術を導入し、学習時の姿勢や動きの度合いを時系列に数値化した動作譜を生成し、授業時間の経過に置ける集中度の変異を客観的に評価するための指標作成や測定手法の提案を目指す。

2. 研究の目的

映像解析に基づく人物行動認識技術は、多くの研究機関で長年にわたり研究されている。これまでの発表された研究成果として、学習時における集中度の計測方法や集中度に影響を与える要因について検討された研究が多く存在する。学習者が課題に取り組んでいる様子を映像に記録する方法で集中度の計測をしている研究では、第三者が観察し、主観評価で集中度を評価している。しかし、主観評価による評価方法では、第三者の主観が評価結果に影響してしまう。また、行動観測指標を作成し、観察した学習者の行動から集中状態を示す行動と非集中状態を示す行動の回数を計測、その結果から集中度の推移を求めている研究がある。このような研究では観察者の主観による評価も含まれるため、行動したという基準が観察者によって異なってしまうことや授業全体に対する集中度を評価している。

一般的に集中力の持続時間の平均は50分、限界は90分と言われている。また、集中力には周期があり、最も集中できている状態というのは15分間程度で、波を描くように15分ごとに集中力が高まったり切れそうになったりを繰り返され、15分の波を3回繰り返した45分にプラス5分の50分とされている。しかしながら、実際どのように生徒が授業中に集中力を維持しているか？どのように時間的に変化していくのか？また、生徒が集中している時にどのような姿勢や動きを示すのか明らかにされていないのが現状である。

そこで本研究では、この問題を打開するため、教室前方に設置した1台のカメラから撮影した授業映像から生徒の姿勢や動き(動作)の候補領域(Candidate Location of Audiences, CLAs)を検出し、深層学習(Deep learning)により各CLAsの映像を6つの動作(向き)にラベリングを行い、時系列のディスクリット信号(以下、動作譜)を生成する。これにより、集中度の時間変異を考慮した推定が可能となる。そして、最終的に識別器により(1)集中状態、(2)曖昧状態、(3)退屈状態の3クラスに分類するシステムを構築する。さらに、各生徒一人一人の動作譜から相関関係を分析し、学習集団がもたらす個人への影響を考慮することで、集中度の推定精度の向上を図り、数値表現による可視化を実現する。

本提案システムを用いることで、教師が自らの授業を振り返るきっかけとなり、日々の実践につなぎ、授業づくりに役立てていただけることを期待している。

3. 研究の方法

本研究で提案するシステムは、大きく分けて (1)キーポイント検出部、(2)人物検出部、(3)姿勢推定部、(4)集中度算出部の4つの処理過程により構成されている(図1)。

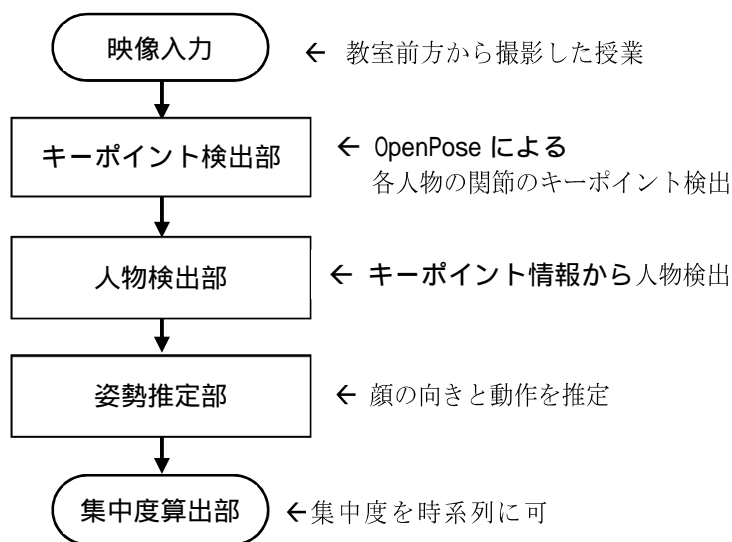


図1：提案システム

(1)キーポイント検出部では、入力映像を一定時間毎に量子化を行い映像のフレーム数を減らし、OpenPose により特徴量を抽出する。OpenPose は、深層学習を用いて単眼カメラの映像から複数人の関節のキーポイント情報をリアルタイムに抽出する姿勢推定ライブラリである。キーポイントの抽出は動画に含まれる人物の身体のみならず、顔や手、足などの各関節部位についても抽出可能である。そして、検出した各キーポイント間を線で結び、人体モデルを出力した画像を得ることができる。また、キーポイントの座標データを CSV や JSON などのデータフォーマットでファイルとして出力することができ、映像から生徒一人一人を検出することが可能である。

(2)人物検出部では、OpenPose により得られる 18 種類の座標データを用い、人物を検出する。具体的には、首の座標を利用し、 Δx と Δy の候補領域(Candidate Location of Audiences, CLAs)を設定する。したがって、人数分の領域が生成され、領域内に存在する首の座標を、人物の中心として判断する。このように、検出した人物 CLAs を、教室の前列から割り当てていく作業を行う。前列の人物と重なってしまい、上手く検出できていない人物や、聴講者以外の人物については排除処理を行う。

(3)姿勢推定部では、顔の動きと動作の推定を行う。具体的には、CLAs の画像に対し、人物の顔向きを6つの向き(左, 右, 上, 下, 正面, その他)にラベリング(数値化)を行い、時系列のディスプレイット信号(動作譜)を生成する。本研究では、左, 右, 上, 下, 正面, その他に対し、それぞれ、5, 4, 3, 2, 1, 0 の値を与える。例えば、DCNN による結果が、左-左-正-正-正-下-下のように得られたとすると、その動作譜は 5-5-1-1-1-2-2 の時系列信号となる。

(4)集中度算出部では、生徒の集中度を厳密に特定することができないが、授業担当の教師による主観的な評価に基づき集中度のパターンを定義する。また、動作検出部で得られた全ての動作譜から相関関係を求める。これにより、学習集団と異なる姿勢や動きを発見することができ、単に正面を向いている生徒を見つけることができる。最終的に集中度のパターンと相関関係を総合し、集中状態(話者の方を向き、話を聞こうとしている姿勢であり、そして学習集団との相関を持っている状態)、曖昧状態(集中状態か退屈状態か判断できない状態)、退屈状態(話者の方を向いていない、話を聞こうとしていない姿勢であり、そして学習集団との相関を持っていない状態)のいずれかに判別する。

4. 研究成果

本研究で提案するシステムは、教育現場で利用されることを想定しており、入力となる動画に映る人物は複数存在し、全ての人物を検出する必要がある。また、各人物の集中度を可視化することが求められている。本研究における具体的な成果は、以下の通りである。

(1) 先行研究では、教室内に複数設置した固定カメラで授業映像を撮影し、その映像から OpenCV を用いて人物を検出し、顔の向きを判定した。しかし、OpenCV を用いた人物検出器では、人物以外の物体を人物であると誤検出してしまうことが多く、顔の向きを正確に判定することが困難であった。本研究では、OpenPose を導入し、授業動画から各フレームを 1 枚の画像として切り取り、その画像について k-means 法を用いキーポイントの位置を推定し、人体モデルを描画した画像を出力することができた。また、入力の映像に複数人数が映っていたとしても、それぞれの人物に対してキーポイントの検出が可能である（図 2）。



(a)入力映像

(b)人物の検出映像

図 2： キーポイントによる複数人物の検出

(2) 人物の検出のみならず、動作の推定として「顔に手を持ってくる動作」、「髪を触る動作」の場合に限定し、両目、両肩の座標に手首座標が近づいた時、ほお杖をつく、もしくは髪を触る動作が発生したと判定を可能にした（図 3）。

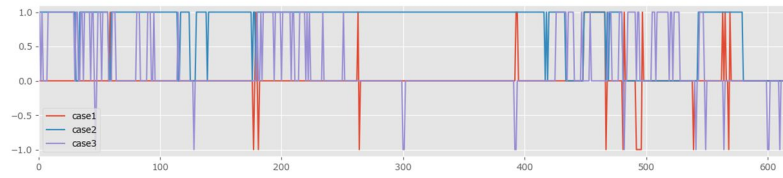


図 3： 顔に手を持ってくる動作推定一例

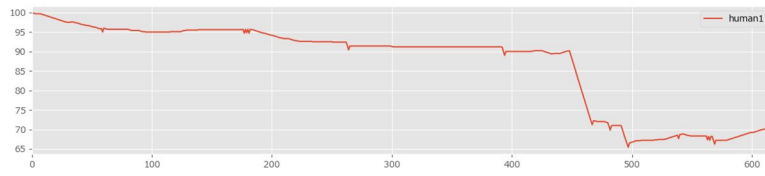
(3) 顔向きと動作をもとに各人物の動作譜を生成し、時刻 t における各人物の総合的な集中度を（式 1）により可視化を実現した。

$$C(t) = C(t-1) + M(t) + G(t) + D(t) \quad (\text{式 1})$$

具体的には、動作の検出においては、人物の過半数以上が行っている動作を正解であると仮定し、正解値と各人物個人の動作が一致しているかどうか比較を行う。大きく動作が発生したタイミングを M で表し、正解でない動作が発生した場合、一定の値を引いていく。また、髪を触る動きやほお杖をつく仕草を G で表し、こちらも集中度が低い動作であるとみなす。逆に集中度が低い仕草があった状態から、仕草がない状態へと変化した場合、動作に合わせて値の加算を行う。顔の向きを D で表し、正面を向いている場合は一定値を加算し、正面から顔の向きの変化が起こった場合は一定値を減算する（図 4）。



(a)動作譜



(b)集中度スコアの推移

図4：集中度の算出一例

<引用文献>

Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields" arXiv:1611.08050 [cs.CV] 14 April 2017.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 1件 / うちオープンアクセス 0件）

1. 著者名 笹澤吉明、姜東植、小林稔	4. 巻 15
2. 論文標題 沖縄の児童における睡眠教育の実践とその効果	5. 発行年 2020年
3. 雑誌名 日本睡眠環境学会雑誌	6. 最初と最後の頁 54-60
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Zahra Nazari, Masooma Nazari, Dongshik Kang	4. 巻 15
2. 論文標題 A Bottom-up Hierarchical Clustering Algorithm with Intersection Points	5. 発行年 2019年
3. 雑誌名 International Journal of Innovative Computing, Information and Control	6. 最初と最後の頁 291-304
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

〔学会発表〕 計5件（うち招待講演 1件 / うち国際学会 2件）

1. 発表者名 姜東植、笹澤吉明
2. 発表標題 Visualization of comprehension by concentration estimation
3. 学会等名 2021 Biomedical Interface Workshop
4. 発表年 2021年

1. 発表者名 姜東植、長嶺澄、笹澤吉明
2. 発表標題 映像の姿勢分析に基づいた集中度の可視化
3. 学会等名 2020 Biomedical Interface Workshop
4. 発表年 2020年

1. 発表者名 姜東植, 桑江和子, 小林稔, 笹澤吉明
2. 発表標題 映像分析による集中度の可視化
3. 学会等名 2019 Biomedical Interface Workshop
4. 発表年 2019年

1. 発表者名 Zahra Nazari, Dongshik Kang
2. 発表標題 A New Hierarchical Clustering Algorithm with Intersection Points
3. 学会等名 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 Zahra Nazari, Seong-Mi Yu, Dongshik Kang, Yousuke Kawachi
2. 発表標題 Comparative Study of Outlier Detection Algorithms for Machine Learning
3. 学会等名 Proceedings of 2018 2nd International Conference on Deep Learning Technologies (国際学会)
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	笹澤 吉明 (SASAZAWA Yosiaki) (50292587)	琉球大学・教育学部・准教授 (18001)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	小林 稔 (KOBAYASHI Minoru) (70336353)	琉球大学・教育学研究科・教授 (18001)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関