

カメラ画像を用いた頭部の動きの判定

土江田織枝*, 篠田裕人**, 林 裕樹*, 山田昌尚*, 峯 恭子***, 香山瑞恵****

Estimation of Head Motion using Camera Images

Orie DOEDA*, Hiroto SHINODA**, Hiroki HAYASHI*, Masanao YAMADA*, Kyoko MINE***,
and Mizue KAYAMA****

Abstract

In recent years, online classes have been increasingly used to prevent infectious diseases. In this form of teaching, the teacher has a disadvantage that it is difficult to understand how the students are looking and reacting. Therefore, we have developed a system that uses information on the direction of the students' faces obtained from the webcam images to detect information such as nodding and head-tilting from the webcam image. This system can be used to estimate the level of understanding of the students to grasp the students' level of understanding of the class. In this paper, we describe a basic study of a method for judging motion from face direction and evaluation experiments.

Key word: face direction, nodding and leaning, Judgment of operation

1. はじめに

近年, 新型コロナウイルス感染の拡大を予防するため, 従来は対面で行われていた授業や会議などを, インターネットを介したオンラインで実施することが増えた. このような形態で行うことで, 参加者の安全が確保でき, お互いが離れていても機会を作れるという利点があるものの, 対面で実施した場合と比較すると, 相手の様子や反応などが分かりにくいという難点がある. そのような中で, 対面でなくては実施が特に難しいと言われている実技指導についても, オンデマンドや同時双方向といった実施形態に依らず, 授業が行われている [1][2]. このような同時双方向型授業の場合は, 教員が見本を示している間に学生の様子を確認することが特に難しい.

そこで筆者らは, ウェブカメラの映像から得た顔の向いている方向 (以後, 顔方向と呼ぶ) と頭部の動きの情報から, 特定の動作を判定しその回数を記録することで, 授業に対する相手の理解度などを把握できないかと考えた.

顔の情報を理解度の把握に用いる研究には, e-Learning で学習中の心理状態の時系列変化を推定するもの [3] や, e-Learning で学習中の学習者の表情や視線, 頭部姿勢といった顔動作を複数のカメラで観測することで, 教材の難しさに対する学習者の主観的印象

を推定するもの [4] があり, このように個人学習に取り組む時の学生の様子を, 顔の情報などから把握する研究が行われている. 本研究は頭部の動きを, 「理解している」や「理解していない」などの意味づけをすることで, 授業に対する学生の理解度などを確認するシステムを開発することを目標とし, 本稿では, その基礎研究として, 顔方向や頭部の動きから特定の動作を判定する方法の検討や評価実験について述べる.

2. システムの概要

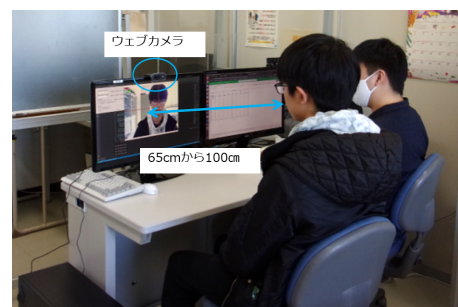


図 1: システムを使用している様子

本システムは, ディスプレイの中央上部に設置した 1280 × 720 画素で総画素数が 120 万画素程度のウェブカメラを使用して動作解析対象者の顔の画像を取得

* 釧路工業高等専門学校創造工学科

** 北見工業大学

*** 大阪大谷大学

**** 信州大学

する。カメラのレンズと顔の距離は 65cm から 100cm 程度とし、顔全体がカメラの画面内に納まるようにする。実際にシステムを使用している様子を図 1 に示す。顔の特徴点を正しく検出できるように、システムの使用中はマスクを着用せず、口はできるだけ閉じた状態とする。本システムの処理は Python3 で実装しており、OpenCV3 のライブラリである dlib より顔の特徴点を検出し [5]、顔方向を取得する。また、リアルタイムで処理を行うため、特徴点の検出は計算負荷が比較的小さい回帰ツリー分析を用いている。対象とする動作は頭部を上下と左右に動かす 2 種類とした。

3. 顔方向の取得と動作の判定

3.1 特徴点からの顔方向の取得

dlib で検出した 68 点の顔の特徴点を図 2 に示す。十字記号で示している両眉の内側・両目の内側・鼻頂点・両小鼻の外側・口の両外側・唇の下・顎の先の 11 点の特徴点の位置情報を顔方向の取得に用いる。11 点の特徴点に関して、取得した画像上での 2 次元座標からカメラの射影行列に基づいて鼻頂点を原点とする仮想的なワールド座標系における 3 次元座標を求め、これらの特徴点群から頭部の剛体モデルのワールド座標系上の回転と並進を求める [6]。図 2 の鼻の先端の矢印記号が、頭部剛体モデルの正面方向を表すベクトルを示しており、この方向を顔方向とする。

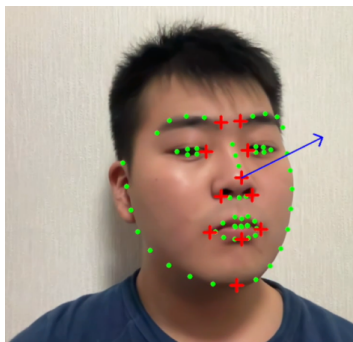


図 2: 顔の特徴点と顔方向の取得

3.2 動作の判定

「うなずく」を判定するためには、頭部剛体モデルの回転から求まる Pitch を用い (図 3)、「首を振る」や「首を傾げる」といった動作は頭部全体の動きを鼻の頂点の横方向の動作の変化を検出して判定する。それぞれの動作が行われたかどうかは、頭部剛体モデルの Pitch 角や鼻の頂点の位置が予め設定した閾値の範囲から外れたかどうかによって判定する。

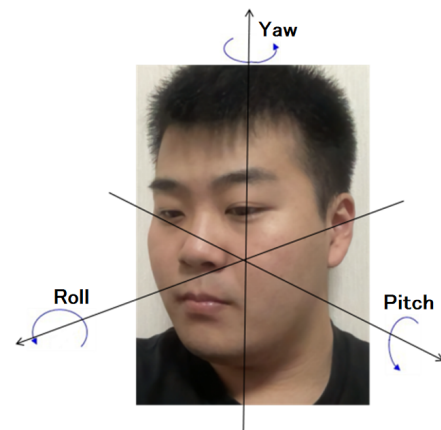


図 3: 剛体モデルの回転軸

3.3 正面方向の調整

本システムを使用する際は、「うなずく」や首を動かす動作をしていない時はなるべくカメラに対し、頭の位置も動かさないようにする必要があるが、本人にそのつもりがない場合でも、ある程度顔方向が変化してしまったり、頭部の位置がずれてしまったりする。また、システムを使用中に使用者が正面であると感じている方向が微妙に変化することもある。このような使用者の主観とシステムの想定する正面方向がずれた状態では、動作を正しく判定できなくなる可能性がある。そのため、本システムでは、動作の開始時に推定した顔方向や頭部の位置を、継続的に更新する処理を行う。これらの初期値は、システムの動作開始時に、使用者にカメラに向かって静止した状態を 100 フレームの間保ってもらい、その間の画像から推定した頭部剛体モデルの顔方向と位置の平均値とする。以後は、直近 100 フレームにおける平均値を動作判定の基準として用いる。

3.4 動作の個人差への対応

動作の大きさには個人差があるため、システムが想定するよりも小さい動作の場合は、閾値による動作判定ができなくなってしまう。そのため、本システムでは、動作の大きさによって閾値の幅を変更できるようにしている。本システムでは、図 4 に示す設定画面から、「うなずく」動作と頭部の位置変化について、0 から 6 を選択することで閾値を設定する。特に設定を行わない場合は、3 の閾値となっている。図 5 の緑色の横の線が「うなずく」動作の、縦の線が頭部の位置変化の動作の判定を行う閾値となっている。図 5 の左の図は閾値を変更していない状態で、右の図は頭部の位置変化の動作を判定する閾値を 6 に変更した状態である。

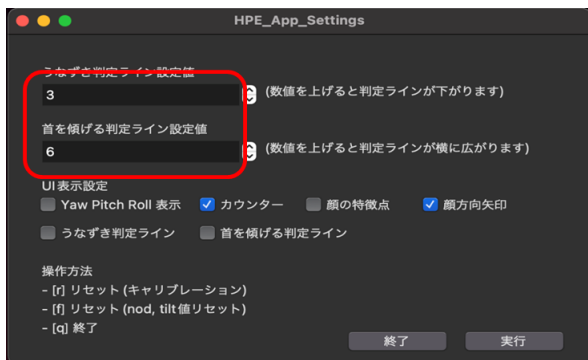


図 4: 各種の設定画面

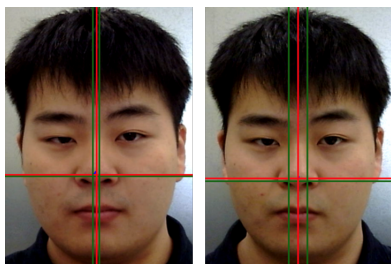


図 5: 閾値の範囲を変更した状態

4. 評価実験

本システムの動作の判定精度について、被験者 21 名を対象とした評価実験を実施した。被験者には「うなずく」と頭部の位置変化を伴う「首を傾げる」というそれぞれの動作に対して、「浅く短く」・「深く短く」・「浅く長く」・「深く長く」の 4 つのパターンの動作を意識して各 5 回ずつ行ってもらった。各動作の大きさの程度はとくに指示はしていない。また、本実験は閾値の設定は変更せずに行った。評価は、実際の回数とシステムが判定した回数を比較することで、システムの動作の判定の精度とした。実験結果を図 6 から図 9 に示す。図 6 と図 7 のグラフの縦軸は動作の検出回数、横軸は被験者 21 名の各動作を示し、それぞれの被験者について左から順に「浅く短く」、「深く短く」、「浅く長く」、「深く長く」の結果を表している。各動作は 5 回ずつ行ってもらっているため、図 6 と図 7 の図中に横線で示した 5 よりもグラフが高ければ検出過多、低ければ検出過少を示す。図 6 の「うなずく」動作の結果では、「浅く短く」は約 47% は正しく判定できているものの、約 23% が実際の回数の半分以下の検出である。「深く長く」は、約 76% が正確に検出できており、それ以外についても 1 回程度の判定ミスだったことから、「うなずく」動作は、「深く長く」行うことで判定精度が高いことが確認できた。次に、図 7 の「首を傾げる」動作の結果から、被験者によって精度が大きく異なり、

特に「深く長く」は実際の回数よりも多く検知することを確認した。これは、「首を傾げる」動作を「深く長く」行くと、左右どちらかに振った状態から戻したときに若干反対側に顔が振れてしまうことが多く、本システムは左右どちらに振っても判定ができるようにしていることで、1 回の動作が左右両方で検知され、重複してカウントすることが原因だと考えられる。

図 8 と図 9 は、各動作に対する評価結果を、4 つのパターンの動作ごとに検出の状態をまとめた結果である。図の横軸は動作の検出の状態を示し、左から順に検出過少、正解、検出過多となっている。また縦軸は 4 つのパターンの動作を示しており、その動作に対する検出の状態ごとの数を円で表し、円が大きいほどその状態の検出が多かったことになる。「うなずく」動作について図 8 からは、正解の 5 回の検出が多く、検出過多は少ない。しかし、検出過少が多いことから、特徴点の変化が小さく変化の時間が短い動作は検出が難しいことがわかる。また、「首を傾げる」動作について図 9 からは、正解の回数の検出は「うなずく」動作よりも低く、検出過少は少ないものの、検出過多での検出が非常に多いことがわかる。

まとめ

顔方向の情報から「うなずく」と「首を傾げる」動作の判定を行った。評価実験では「うなずく」動作は、深く長く行うことで正しい判定ができたが、「首を傾げる」動作については、短時間に左右両方の振りを検知した場合には動作を無効にするなどの対応をすることで、正しい判定ができるように改良する。それぞれの動作について、「うなずく」動作は「理解できている」や「実現できる」、また、「首を傾げる」動作は「理解できていない」、「実現できない」などと意味付けをし、授業の画像と各動作が生じたときの時間情報や動作の情報を結びつける機能を実装する。このような機能により、特に同時双方向型の授業中などで、学生が教員に自分の状況を伝えることができないときには、これらの動作をすることで、教員はどの段階で学生が理解できなかったか、また理解できていたかということを確認できるようにする。特に、楽器演奏の同時双方向型のオンライン授業への活用を目指してシステムの開発を進める予定である。

参考文献

- [1] 西田順一, 木内敦詞, 中山正剛, 難波秀行, 園部豊, 西脇雅人, 平工志穂, 小林雄志, 西垣景太, 中田征克, 田原亮二, 新型コロナウイルス感染症第 1 波の流行直後における大学体育授業の学修成果: 遠隔授

業による主観的恩恵と身体活動に焦点をあてた検証, 大学体育スポーツ学研究 18, pp.2-20(2021).

- [2] 山真理, 酒井国作, 藤原一子, 森田千智, 山本馨栄子, 保育者養成教育としてのピアノレッスンにおけるオンライン授業実践の省察, 東海学園大学教育研究紀要, 第5巻, pp.99-114,(2021).
- [3] 黒川 智哉, 野洲 潔, 山崎 清之, 顔画像および生体信号のテンプレートを用いた e-Learning 学習者の心理状態時系列変化推定の検討, 映像情報メディア学会, Vol. 61, No.12, pp.1779-1784(2007).
- [4] 中村 和晃, 角所 考, 村上 正行, 美濃 導彦, e-learning における学習者の顔動作観測に基づく主観的難易度の推定, 電子情報通信学会論文誌 D Vol. J93D, No. 5, pp. 568578(2010).
- [5] Face landmark detection in a video, https://docs.opencv.org/3.4/d8/d3c/tutorial_face_landmark_detection_in_video.html (参照 2023.1.10).
- [6] Head Pose Estimation using OpenCV and Dlib, <https://learnopencv.com/head-pose-estimation-using-opencv-and-dlib/> (参照 2022.01.10).

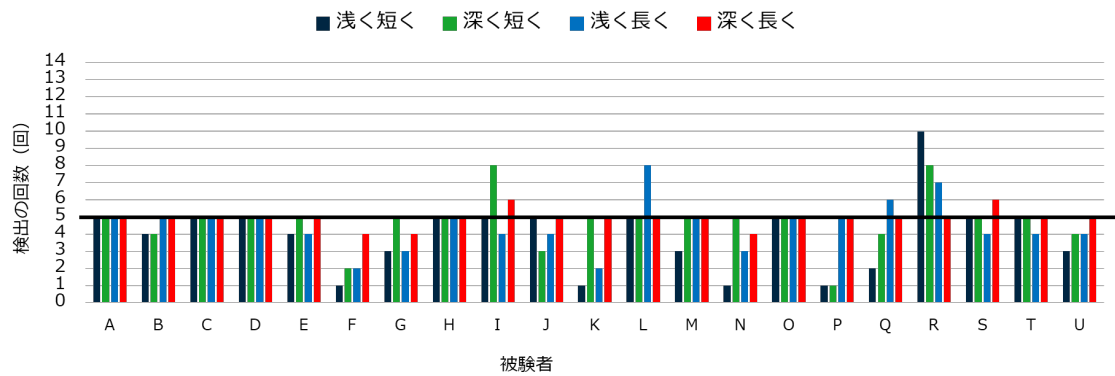


図 6: うなづく動作の評価実験の結果

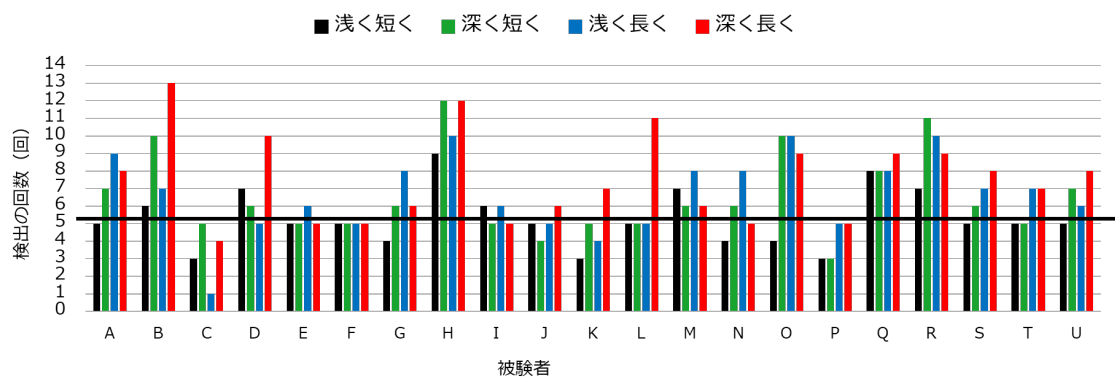


図 7: 首を傾げる動作の評価実験の結果

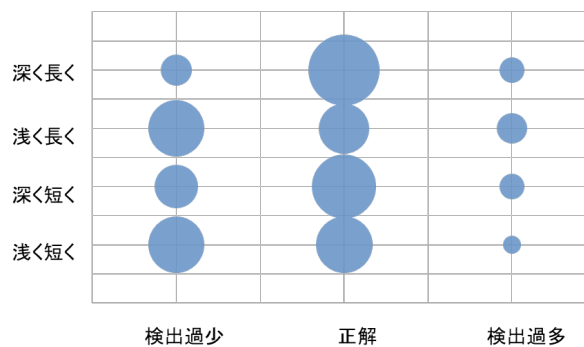


図 8: うなづく動作の 4 つのパターン動作ごとの評価結果のまとめ

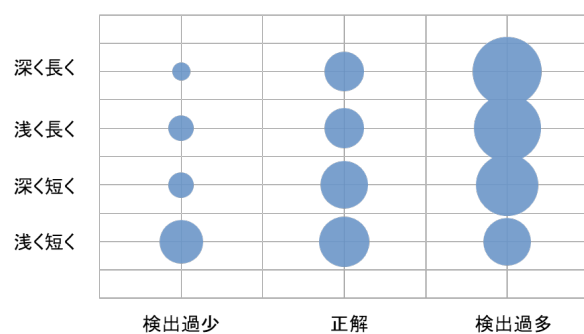


図 9: 首を傾げる動作の 4 つのパターン動作ごとの評価結果のまとめ