

スマートグラスを用いた顔向きに基づく集中度推定手法の提案

山本朔叶^{*1}・金恵鎮^{*2}・金義鎮^{*1}
Email: kim@mail.tohoku-gakuin.ac.jp

*1: 東北学院大学工学研究科電気工学専攻

*2: 日本大学商学部

◎Key Words

スマートグラス, 集中度推定, 頭部姿勢推定, リアルタイム

1. はじめに

近年, 教育現場では個別最適な学びの実現に向けて, 学習者一人ひとりの理解状況や学習状態を把握し, 個々に応じた学習支援を行うことが求められている⁽¹⁾。そのため, 教員は授業進行と並行して全ての学習者の状態を継続的かつ客観的に把握し, 適切に指導する必要がある。例えば, 大人数授業では学習者の集中状態や注意状態を個別に把握することが容易ではなく, 教員による学習状況の把握を支援することが望ましい。

このような学習支援の一つの方法として, 視線情報, 顔表情, 身体姿勢, 頭部姿勢などを利用した注意状態推定手法が提案されている⁽²⁾⁻⁽⁶⁾。視線情報は高い推定精度が期待できるが, 専用の計測装置や高精度な撮影環境を必要とする場合が多い。そこで, Web カメラやスマートフォンカメラなどの一般的な RGB カメラで取得可能な頭部姿勢情報を利用した手法が注目されている。頭部姿勢は人物の注視方向と関連が深く, 集中状態を推定する有効な指標である。しかし, 従来研究の多くは教室前方や天井に設置した固定カメラを前提としており, 教員が教室内を移動しながら学習者の注意状態を推定する利用形態については十分に検討されていない。そのため, 教員の移動に追従しながらハンズフリーかつリアルタイムに映像を取得できる撮影手段が求められている。

この課題に対して, 装着者の視点から画像を取得できるスマートグラスは有効な手段である。スマートグラスを利用することで, 教員は学習者に視線を向けることで情報を取得でき, 固定カメラでは困難な柔軟な運用が可能となる。我々はこれまでに, スマートグラスを用いた顔認識システムを開発し, 学習者の識別や学習状況の確認を支援する手法を提案するとともに, スマートグラス環境における人物認識の実装可能性を確認している⁽⁶⁾。一方で, 学習者の集中状態を把握する機能は実現していない。

そこで本研究では, スマートグラスに搭載されたカメラ映像から顔向きを推定し, その結果に基づいて集中度を推定する手法を提案する。提案手法では, ML Kit Face Detection および MediaPipe Face Landmarker を用いて顔情報を取得し, 顔向きに基づいて集中度を推定する。推定された集中度はスマートグラス上にリアルタイム表示され, 授業中における教員による学習者の状態把握を支援する。実験では, ML Kit Face Detection と MediaPipe Face Landmarker の両方式を実装し, スマートグラス環境における顔向き推定性能および集中度推定への適用可能性を定量的に比較評価する。

2. 関連研究

本章では, 頭部姿勢推定, 注意・集中度推定, およびスマートグラスを用いた研究について述べる。

2.1 頭部姿勢推定に関する研究

頭部姿勢推定は, 顔画像から頭部の向きを推定する研究分野であり, 従来から多くの手法が提案されている。頭部姿勢推定手法としては, 顔ランドマークを用いて頭部姿勢や表情を推定する OpenFace などの実装も広く利用されている⁽⁷⁾。総説研究において, 頭部姿勢推定は制約の少ない実環境において困難な課題であり, 顔ランドマークを用いる手法や深層学習を用いる手法が発展してきたと報告されている。これらの研究は, 頭部姿勢が人物の向きや注視傾向を表す有効な指標であることを示している。

頭部姿勢は, 人物の注意方向を推定するための代理指標としても利用されている。しかし, 既存研究の多くは頭部姿勢の推定精度向上を目的としており, 推定結果を教育現場における集中度推定へ応用する検討は十分ではない。そのため, 頭部姿勢情報を利用した集中度推定手法の実用化に向けた検討が必要である。

2.2 注意・集中度推定に関する研究

注意状態や視覚的焦点を推定する研究では, 頭部姿勢が重要な入力情報として利用されている。会議環境を対象とした研究では, 頭部姿勢から Visual Focus of Attention を推定する手法が提案されており, 頭部の向きと注視対象の関係が分析されている。また, 教室環境を対象とした研究では, コンピュータビジョンを用いて学生の注意状態を監視する試みが報告されている。これらの研究は, 頭部姿勢が学習場面における注意推定へ有効な指標として応用可能であることを示している。

注意推定の精度向上には, 視線, 顔表情, 身体姿勢などの複数情報を組み合わせる手法が有効である。しかし, 視線計測や複数センサを利用する構成は専用機器を必要とするため, 教育現場への導入コストが高くなるという課題がある。そのため, 単一の RGB カメラのみを利用して集中度を推定できる手法が求められる。

2.3 スマートグラスを用いた研究

スマートグラスを用いた研究では, 装着者の視覚情報や注意情報を利用したインタラクション手法が提案されている。iGaze は, スマートグラスを用いて装着者の視覚的注意を取得し, 関心対象との接続に反映する仕組みを

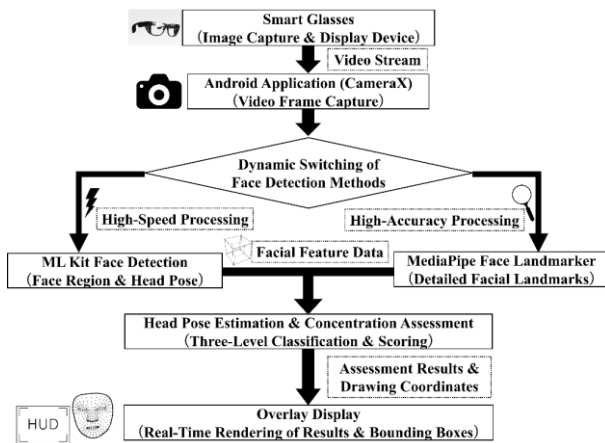


図1. システムの概要図

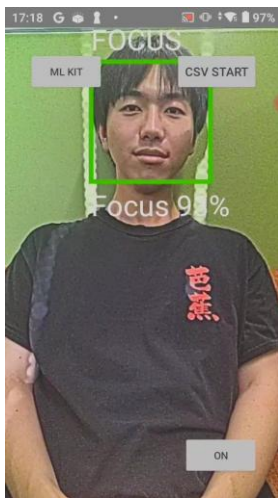


図2. ML Kit 動作画面



図3. MediaPipe 動作画面

提案している⁸⁾。この研究は、スマートグラスが表示端末としてだけでなく、人物の注意状態を扱う入力デバイスとしても利用可能であることを示している。

一方で、既存研究ではスマートグラスを用いて学習者の集中状態を把握する研究と、教員がリアルタイムに学習者の状態を把握する仕組みについては十分な検討が行われていない。そのため、スマートグラス上で顔向きを利用した集中度推定を実現する手法の検討が必要である。

3. 提案手法

前章で示した課題を解決するため、本研究では顔向き情報を利用した集中度推定システムを提案する。提案システムは、頭部姿勢情報を利用した集中度推定、RGBカメラのみを用いた低コストな構成、およびスマートグラス上でのリアルタイム表示を特徴とする。本章では、提案システムの構成および集中度推定手法について述べる。

3.1 システム構成

図1に提案システムの概要を示す。本研究では、スマートグラス上で動作するAndroidアプリを用いて、対象者の顔向きを推定し、集中度を推定するシステムを構成した。システムは、カメラ映像取得部、顔検出部、顔向き推定部、集中度推定部、表示部の5つで構成される。カメラ映像はCameraXで取得し、検出結果は画面上のオーバーレイに重ねて表示する。

本システムは、スマートグラス上でのリアルタイム動

作を考慮し、撮影から表示までを端末内で完結する構成とした。これにより外部サーバとの通信を必要とせず、ネットワーク環境に依存しない運用が可能である。また、教員が学習者へ視線を向けるだけで推定結果を確認できるため、授業中の利用を想定した低遅延な情報提示を実現している。顔検出方式は、ML Kit Face DetectionとMediaPipe Face Landmarkerの2方式を切り替えて利用できるようにした。ML Kit Face Detectionを用いる方式では、顔領域と頭部回転角度を取得する。MediaPipe Face Landmarkerを用いる方式では、顔ランドマークから両目と鼻の位置関係を利用して顔向きを推定する。ML Kit Face Detectionは、顔の位置情報とheadEulerAngleX, headEulerAngleY, headEulerAngleZを提供し、MediaPipe Face Landmarkerは3次元顔ランドマークを出力する⁹⁾⁽¹⁰⁾。

3.2 顔検出と顔向き推定

ML Kit Face Detectionを用いる場合、本システムは頭部姿勢情報として上下方向および左右方向の回転角度を利用する。headEulerAngleXは顔の上下方向の傾き、headEulerAngleYは左右方向の傾きを表す。本研究では顔向きの代表値として両角度の絶対値の大きい方を利用する。図2にML Kit Face Detectionを用いた動作画面を示す。ML Kit方式は顔検出と同時に頭部姿勢情報を取得できるため、追加の計算処理をほとんど必要としない。この特徴により、処理能力が限られるスマートグラス環境においても安定したリアルタイム動作が期待できる。

MediaPipe Face Landmarkerは、MediaPipe Face Meshで提案された顔ランドマーク推定技術を基盤としており⁽¹¹⁾、本研究では取得したランドマークから両目と鼻の位置関係を利用して顔向きを推定する。顔向き推定には、左目、右目および鼻先の3次元ランドマークを利用する。まず左右の目の中心座標を求め、その中心から鼻先へ向かうベクトルを算出する。顔の左右回転角度は、鼻先と目中心とのX方向成分およびZ方向成分を用いて逆正接関数atan2を適用して推定する。これにより、MediaPipe Face Landmarkerが出力する3次元情報を利用した顔向き推定を実現する。また、MediaPipe方式では顔ランドマーク群の外接矩形を算出し、顔位置の可視化にも利用している。MediaPipe Face Landmarkerは高密度な3次元顔情報を取得できるため、顔向きだけでなく、将来的には視線推定や表情解析などへの拡張も期待できる。

3.3 集中度推定手法

本研究では、顔向きが小さいほど対象者が正面方向を向いているとみなし、集中状態を推定する指標として利用する。ML Kit方式では、角度が10度以下のときにFOCUS、10～20度のときにWARNING、それ以外をLOSTとする3段階推定を行う。集中度は、式(1)により算出した値を0～100の範囲に補正して百分率として表示する。この設計では、顔向きの変化を定性的な状態だけでなく、定量的な値としても扱える。

$$Focus = 100 - 2 \times Angle \quad (1)$$

表示部は、顔領域の矩形、状態ラベル、集中度の3要素を重ねて表示する。FOCUSは緑、WARNINGは黄、LOS



図4. 実験に使用したマネキン

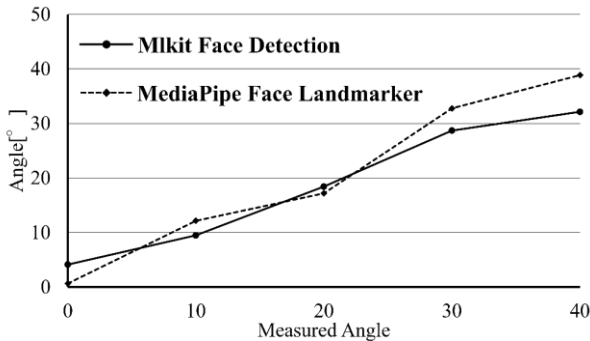


図5. ML Kit Face Detection および MediaPipe Face Landmarker の1mの結果

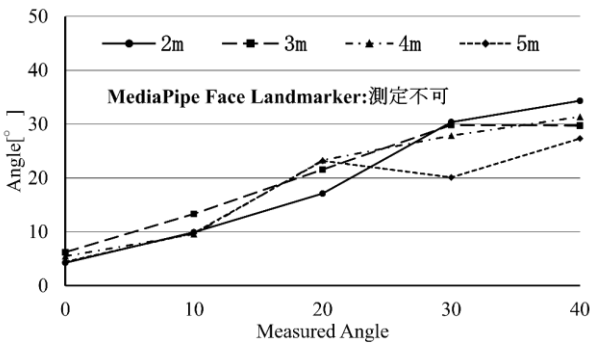


図6. 2~5mにおける ML Kit Face Detection の結果

Tは赤で示し、装着者が一目で状態を把握できるようにする。MediaPipe方式でも同様に、検出顔ごとに状態と集中度を描画する。また、本システムはスマートグラスでの表示位置ずれを考慮し、描画オフセットを補正する。この補正により、本システムはスマートグラス特有の表示環境においても利用時の見やすさを維持できる。本研究では予備実験において顔向き推定値を確認した結果に基づき、10°および20°を閾値として設定した。なお、本閾値は集中状態との対応を直接示すものではなく、正面注視からの逸脱度を表す指標として利用する。

4. 実験

本章では、提案システムの有効性を確認するために実施した実験方法および結果について述べる。

4.1 実験方法

本研究では、顔向き推定機能が実際の顔向き変化をどの程度反映できるかを確認するため、角度および距離を変化させた実験を行った。実験にはスマートグラス上で動作する試作アプリを用いた。評価対象には図4に示す

マネキンを使用し、カメラに対して正面(0°)、10°、20°、30°、40°の5条件で顔の向きを変化させた。また、マネキンとスマートグラス間の距離は1m、2m、3m、4m、5mの5条件に設定した。測定は室内環境で実施し、照明条件を大きく変化させないようにした。各角度条件では、マネキンの設置位置を固定したまま顔向きのみを変更した。本実験では、マネキンを用いることで人体の動きや姿勢変化による影響を排除し、頭部姿勢推定結果のみを評価できるようにした。

各条件について5回ずつ測定を行い、ML Kit Face DetectionおよびMediaPipe Face Landmarkerから取得した顔向き推定値を記録した。なお、MediaPipe Face Landmarkerについては、スマートグラス上での処理負荷が大きく、2m以上の距離では安定した顔検出が困難であったため、1m条件のみ評価対象とした。本研究では、取得した測定値の平均値を用いて顔向き角度と推定値の関係を評価した。複数回測定の平均値を利用することで、一時的な検出誤差やフレームごとのばらつきの影響を低減し、評価結果の信頼性を向上させた。

4.2 結果

本節では、ML Kit Face DetectionおよびMediaPipe Face Landmarkerによる実験結果について示す。

図5に1m条件におけるML Kit Face DetectionとMediaPipe Face Landmarkerの測定結果を示す。両方式とも、顔向き角度の増加に伴い推定値が増加する傾向が確認された。MediaPipe Face Landmarkerは実角度に近い推定値を示した一方で、ML Kit Face Detectionは推定値がやや小さくなる傾向が見られた。しかし、いずれの方式も顔向きの変化に応じて推定値が一貫して変化しており、顔向きの変化を反映できることが確認された。

図6に2~5m条件におけるML Kit Face Detectionの測定結果を示す。ML Kit Face Detectionは、全ての距離条件において顔検出および顔向き推定を継続して行うことができ、距離が増加しても顔向き角度の増加に伴う推定値の増加傾向が維持された。一方、MediaPipe Face Landmarkerは2m以上では顔ランドマークの検出が安定せず、継続的な測定結果を取得することができなかった。以上より、両方式とも1mでは顔向き推定が可能であったが、測定可能距離および処理性能には差が見られた。

5. 考察

本章では、実験結果に基づき、ML Kit Face DetectionおよびMediaPipe Face Landmarkerの特徴と実用性について比較しながら詳細に考察する。

5.1 ML Kit Face Detectionの評価

ML Kit Face Detectionでは、全ての距離条件において顔検出および顔向き推定を実施することができた。図5および図6より、顔向き角度の増加に伴い推定値も増加する傾向が確認できることから、頭部姿勢の変化を概ね適切に反映できていると考えられる。一方で、実際の顔向き角度と推定値の間には差が見られた。特に30°以上では推定値の増加が緩やかとなり、大きな顔向き変化に対して推定値が飽和する傾向が確認された。

しかし、本研究の目的は顔向きを厳密に計測することではなく、集中状態の変化を推定することである。そのた

め、実角度との完全な一致よりも、顔向き変化に応じて推定値が一貫して変化することが重要である。本実験結果から、ML Kit Face Detection は集中度推定に必要な顔向き情報を取得する手法として有効であると考えられる。また、5m 条件においても安定した顔検出が可能であったことから、教室環境のような比較的広い空間への適用可能性が示された。

5.2 MediaPipe Face Landmarker の評価

MediaPipe Face Landmarker では、1m 条件において顔向き角度の増加に伴い推定値も増加する傾向が確認された。ML Kit Face Detection と比較すると、実角度に近い推定値を示す傾向があり、顔向きの変化をより詳細に把握できる可能性が示された。また、468 点の顔ランドマークを利用するため、将来的には顔向き以外の特徴量を利用した応用も期待できる。

一方で、本研究で使用したスマートグラス環境では処理負荷が大きく、リアルタイム処理時に描画の遅延やフレームレートの低下が確認された。また、2m 以上では顔ランドマークの検出が安定せず、十分な測定結果を取得できなかった。以上より、MediaPipe Face Landmarker は推定精度の面では有望であるものの、本研究のようなスマートグラス上でのリアルタイム集中度推定においては、処理速度および測定可能距離の面で課題が残る。そのため、現段階では ML Kit Face Detection の方が実運用に適した方式であると考えられる。この結果は、顔向き推定精度とリアルタイム性との間にトレードオフが存在することを示している。

6. おわりに

本研究では、スマートグラス上で動作する Android アプリを用いて、対象者の顔向きを推定し、集中度を推定するシステムを提案した。本システムは、ML Kit Face Detection と MediaPipe Face Landmarker を切り替えて利用し、顔領域、顔向き、集中状態をリアルタイムに表示する。また、顔向きの変化に応じて集中度を数値および 3 段階の状態として提示することで、学習者の状態把握を支援するための基盤システムを実現した。

実験の結果、1m 条件では両方式とも顔向きの変化に応じた推定値の増加が確認された。一方、2m 以上では ML Kit Face Detection は安定して顔検出および顔向き推定を行えたのに対し、MediaPipe Face Landmarker では安定した測定結果を取得できなかった。以上より、本研究で対象とするスマートグラス上でのリアルタイム集中度推定には、ML Kit Face Detection が有力な候補であることが示された。今後は、測定数を増やした評価実験を実施し、顔向きと実際の集中状態との関係を検証する。また、顔表情や視線情報との統合により集中度推定の高精度化を図るとともに、実際の教室環境において複数の学習者を対象とした評価を行い、教育支援システムとしての実用性および有効性を検証していく。

参考文献

- (1) 文部科学省 : "「個別最適な学び」と「協働的な学び」の一体的な充実", 文部科学省, https://www.mext.go.jp/a_menu/shotou/new-cs/senseioun/mext_01317.html.

- (2) Erik Murphy-Chutorian, Mohan M. Trivedi : "Head Pose Estimation in Computer Vision: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.31, No. 4, pp.607-626 (2009).
- (3) Khalil Khan, Rehanullah Khan, R. Leonardi, P. Migliorati, S. Benini : "Head Pose Estimation: A Survey of the Last Ten Years", Signal Processing: Image Communication, Vol.99, pp.116479 (2021).
- (4) Sileye O. Ba, Jean-Marc Odobez : "A Study on Visual Focus of Attention Recognition from Head Pose in a Meeting Room", Lecture Notes in Computer Science, Vol.4299, pp. 75-87 (2006).
- (5) Daniel Canedo, Alina Trifan, António J. R. Neves : "Monitoring Students' Attention in a Classroom Through Computer Vision", Communications in Computer and Information Science, Vol.725, pp.371-378 (2018).
- (6) 山本朔叶, 金恵鎮, 金義鎮 : "スマートグラスに適する異なる深層学習モデルの評価", 画像電子学会年次大会, paper_57 (2025).
- (7) Tadas Baltrušaitis, Peter Robinson, Louis-Philippe Morency : "OpenFace 2.0: Facial Behavior Analysis Toolkit", IEEE International Conference on Automatic Face and Gesture Recognition (FG), pp.59-66 (2018).
- (8) Lan Zhang, Xiang-Yang Li, Wenchao Huang, Kebin Liu, Shuwei Zong, Xuesi Jian, Puchun Feng, Taeho Jung, Yunhao Liu : "It Starts with iGaze: Visual Attention Driven Networking with Smart Glasses", Proceedings of the 20th Annual International Conference on Mobile Computing and Networking, pp.91-102 (2014).
- (9) Google : "ML Kit Face Detection for Android", Google Developers, <https://developers.google.com/ml-kit/vision/face-detection/android>
- (10) Google : "Face Landmarker Guide for Android", Google AI Edge, https://ai.google.dev/edge/mediapipe/solutions/vision/face_landmarker/android
- (11) Yury Kartynnik, Artsiom Ablavatski, Ivan Grishchenko, Matthias Grundmann : "Real-time Facial Surface Geometry from Monocular Video on Mobile GPUs", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops) (2019).