

Web 会議における全方位 Web カメラ映像の 顔向き推定に基づく集中度の推定

松田 薫[†] 阿倍 博信[†]

東京電機大学[†]

1. はじめに

新型コロナウイルスの感染が拡大して以降、リモート環境下でもより対面に近い環境での作業ができるように、全方位 Web カメラを使用した Web 会議の機会が増加している。しかし、参加者の状況を把握することが難しいという問題があげられる。そこで、本研究では、全方位 Web カメラを用いた Web 会議において、会議映像を解析し、参加者の集中度を推定する方式を提案する。映像解析による集中度の推定は、顔の表情や動き、体の姿勢や動きを利用したものが一般的だが、本研究では、顔向きに着目し、参加者がマスクを着用していた場合でも、その推定結果を用いて集中度を推定する。本論文では、全方位 Web カメラが出力した映像を、3 次元顔姿勢推定ライブラリである img2pose を用いて顔姿勢推定を行い、その結果を元に集中度の推定を行った結果について報告する。

2. 関連研究

一般的に集中度の推定は、姿勢や顔の表情、動きを推定し、推定する研究が多い。株式会社富士通研究所のオンライン学習・営業活動など参加者の集中状態の定量化を行う研究では、顔の各部位の動作単位である ActionUnit(AU)を活用し、AU 毎の動きの変化を時間単位で最適化し捉えることで集中度を推定している[1]。他にも、藤白らは子供の授業中での集中度合いは子供の姿勢から推定できるとしている[2]。授業中の子供の姿勢推定を行い、両肩を頂点とした法線を体の向きと定義し、子供の集中状態を保育経験者に判断してもらうことにより、姿勢を用いた集中度の推定を行っている。

3. 集中度の推定方式の提案

関連研究の結果より、本研究では、顔向きに着目し、参加者がマスクを着用していた場合で

も、その推定結果を用いて集中度を推定する。

3.1 顔向きの推定方式

顔向きの推定は、画像内の顔を検出し、顔の特徴点をランドマークとして抽出することで姿勢を推定する方式が多く用いられている。HyperFace[3]では、CNN を用いて顔検出、ランドマーク抽出を行い、顔の姿勢推定を行っている。2 つの ResNet-101 モデルを改良したもので構築されており、アルゴリズムの高速化やパフォーマンスの向上に寄与している。また、3 次元顔姿勢推定ライブラリである img2pose[4]では、事前の顔検出やランドマーク抽出を必要とせず、Faster R-CNN ベースのモデルを用いたアルゴリズムで顔の剛体変換の推定を行い、リアルタイムに顔の 6 つの自由度を回帰することで、3 次元の顔の姿勢推定を行っている。

3.2 集中度の推定方式

集中度を推定するために、img2pose を用いて被写体の顔向き推定を行い、頭部の Yaw, Pitch の情報から被写体がどこを向いているかを判定する。また、同じ方向を向いている時間にも着目し、頻繁に顔向きが変化している状態や長時間同じ方向を向いている状態などから集中度の推定を行う。

25 フレームごとに顔向きを D 、一定時間ごとの顔向きの動きを T として、集中していないと判定された場合にそれぞれから決められた値を引いて行くことで集中度を推定する。

式(1)に、提案した集中度の推定式を示す。

$$C_t = C_{t-1} + D + T \quad (1)$$

この集中度が一定の値を下回った場合に集中力が切れていると判断し、会議の参加者にフィードバックを行う。その後、集中度の値をリセットし、再度集中度の推定を行う。

4. 評価実験

4.1 実験準備

本研究では、Zoom を用いた Web 会議の映像を

Concentration estimation based on face orientation estimation of omnidirectional web camera video in web conference

[†] MATSUDA KAORU, ABE HIRONOBU, Tokyo Denki University

全方位 Web カメラである Kandao Meeting Pro を使用し撮影する。Kandao Meeting Pro は Zoom などを用いた Web 会議において、会議室を画面の上下 180° ずつの画角で 360° 撮影することができる。以下に、Kandao Meeting Pro で Zoom を用いた Web 会議を撮影した例を図 1 に示す。

次に、カメラと被験者の配置を図 2 のように設定した。カメラの正面に座る被写体 B、E はそれぞれカメラから 70cm、その隣の A、C、D、F はそれぞれ 135cm の位置に座っている。



図 1. Kandao Meeting Pro での Web 会議の撮影例

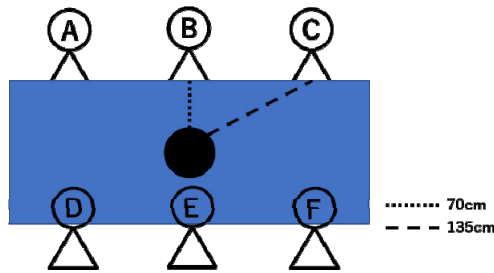


図 2. カメラと被験者の配置

4.2 img2pose による顔向き推定

img2pose では、6DoF という顔の 3 次元位置を 6 つの自由度として回帰し、顔の 3 次元姿勢推定を行っている。本研究ではまず、この 6DoF のうち、顔の左右の傾きを表す Yaw と顔の前後の傾きを表す Pitch に着目し、各被写体の位置において、img2pose を用いた顔向きの角度が何度まで検出することができるか評価を行った。

図 3 に被写体 B において約 1 分間会議を模し

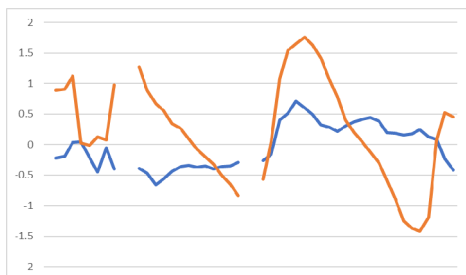


図 3. 被験者 B の Pitch (青), Yaw(橙)

た映像を撮影し、Yaw と Pitch を推定した結果を示す。

次に、測定した Yaw と Pitch から、各被写体の測定可能な顔向きの角度を表 1、表 2 に示す。

表 1. 推定可能な顔向きの角度(A~C)

	A	B	C
Yaw	-0.48 ~ 1.82	-1.41 ~ 1.76	-1.49 ~ 0.93
角度	68.4° ~ 171.9°	26.6° ~ 169.2°	23.0° ~ 131.9°
Pitch	-0.46 ~ 0.54	-0.65 ~ 0.70	-0.62 ~ 0.51
角度	-48.6° ~ 41.4°	-63.0° ~ 58.5°	-45.9° ~ 55.8°

表 2. 推定可能な顔向きの角度(D~F)

	D	E	F
Yaw	-1.50 ~ 0.87	-1.54 ~ 1.34	-0.49 ~ 1.99
角度	22.5° ~ 129.2°	20.7° ~ 154.4°	68.0° ~ 179.6°
Pitch	-0.61 ~ 0.92	-0.44 ~ 0.62	-0.42 ~ 0.69
角度	-82.8° ~ 54.9°	-55.8° ~ 39.6°	-62.1° ~ 37.8°

4. おわりに

本研究では、集中度の推定方式の提案と、Kandao Meeting Pro で撮影した全方位画像の各被写体に対して img2pose を用いた 3 次元顔姿勢推定の基礎評価を行った。

今後は、今回の結果をもとに、今回提案した顔向きや時間ごとの顔の動きにおける集中状態・非集中状態を元に集中度の推定方式について評価を行っていく。また、推定した個人の集中度をもとに場の集中度を推定し、フィードバックを行うことで会議を円滑化するシステムの構築を行っていく予定である。

参考文献

- [1] 人工知能研究所: 様々なタスクにおける人の集中度を定量的に推定可能な汎用 AI モデルを開発, 株式会社富士通研究所, <https://pr.fujitsu.com/jp/news/2021/03/1-1.html>, (参照 2022/12/05).
- [2] 藤白 智也, 浅利 恭美, 山田 徹志, 大森 隆司: 授業中の画像からの姿勢推定による集中度の評価, 第 34 回人工知能学会全国大会, 2C1-GS-12-03, 2020.
- [3] Rajeew Ranjan, Vishal M. Patel, Rama Chellappa: HyperFace: A Deep Multi-task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.41, No.1, pp.121-135, 2019.
- [4] Vitor Albiero, Xingyu Chen, Xi Yin, Guan Pang, Tal Hassner: img2pose: Face Alignment and Detection via 6DoF, Face Pose Estimation, IEEE CVPR2021, 2021.