

オンライン講演者の快適性向上のための視聴者の心的状態推定と伝達

菊地 怜† 小林 美月†† 辻 愛里††† 藤波 香織†††

† 東京農工大学 工学部 情報工学科

†† 東京農工大学 大学院 生物システム応用科学府 生物機能システム科学専攻

††† 東京農工大学 大学院 工学研究院 先端情報科学部門

1 はじめに

COVID-19 ウイルスの蔓延により、会議や授業、学会発表など以前は対面で行われていたものがオンラインに移行した。オンライン化の利点としては、遠隔地の相手が容易に参加できる点や場所の確保が不要な点などが挙げられ、COVID-19 ウイルス収束後も利用され続けると予想される。一方で、オンラインは対面と比べて視聴者の反応が得づらいことが大きな課題である [1]。カメラやウェアラブルデバイスを用いて検出・推定した視聴者の反応を可視化することでオンライン発表を支援する研究は盛んに行われている。その多くは理解度や集中度など発表者に有益な情報を推定し提示することに成功しているが、発表者が指標の悪化を認知すると緊張や不安を感じ快適に発表できないことが課題となっている。本稿では、複数デバイスを用いて視聴者の反応を詳細に検出しつつ、オンライン講演者の快適性を向上する可視化方法を提案する。

2 視聴者の反応のフィードバックシステム

2.1 システム概要

本システムは、オンライン発表中の視聴者の反応をリアルタイムで検出し、発表者に提示する。提示方法として顔のアイコンを提案する。顔のアイコンに対応した色や形を設定でき、頭部動作に対応した動きを付けることができる。また、個人を特定できる表示方法では視聴者のプライバシーを侵害する可能性があるが、アイコンでの表示は匿名化も可能である。図1に本システムの構成図を示す。本システムは視聴者の反応検出部、サーバー部、情報提示部の3つで構成されている。視聴者の反応検出部ではオンライン発表でよく利用されるイヤホンとウェブカメラを用いて、視聴者の頭部動作と表情を検出する。イヤホンはIMUセンサー内蔵のeSense[2]を使用する。各視聴者のPC上で検出された反応をサーバーに集約し、アプリケーション上で可視化することで、発表者にリアルタイムで視聴者の反応をフィードバックする。以下に各機能の説明を述べる。

2.2 視聴者の反応検出部

2.2.1 頭部動作検出

eSense から頭部動作の加速度データと角速度データをサンプリング周波数 100Hz で取得する。取得したeSenseの生データのウィンドウ処理（ウィンドウサイ

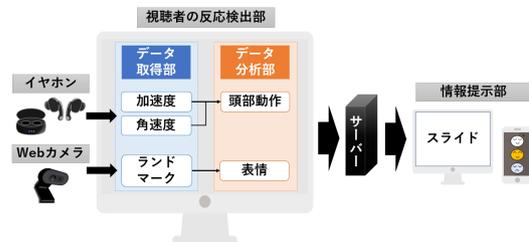


図1: システム構成

ズ:32, オーバーラップ率:50%)や合成加速度の計算等の前処理を行い、各ウィンドウに対して行動判定を行う。発表者が見たい視聴者の反応として、うなずきと首振りが上位に挙げられている [3]。そこで本システムでは、頭部動作としてこの2種類の行動を判定する。時間領域 (69次元) と周波数領域 (84次元) の計153次元の特徴量を用いて、事前に収集した学習データを訓練データとし、リアルタイムに取得されるデータにランダムフォレストを適用することで判定する。判定された行動をサーバー部へ送信する。事前収集データについて層化10分割交差検証を行った結果、マクロ平均のF値が0.91を示した。

2.2.2 表情検出

各視聴者PCのウェブカメラ映像を取得し、取得した映像にPAZ[4]を適用することでリアルタイム表情判定を行う。判定する表情は喜び、悲しみ、怒り、驚き、恐れ、嫌悪、無表情の7種類であり、各表情の可能性を数値化し、最も数値の高い表情を選択する。判定した表情とその値をサーバー部へ送信する。

2.3 サーバー部

サーバー部では、各視聴者PCから送信される判定結果を集約する。発表者デバイスとは一定時間ごとに接続が確立され、その間の各視聴者の頭部動作と表情を決定し、発表者デバイスに送信する。

2.4 情報提示部

情報提示部では、サーバー部から受信した視聴者の反応を可視化する。ただし、先行研究 [5]によると周辺表示法では、情報提示によるメインタスクへの妨害感の低減が課題となっており、情報の閲覧を強制することはユーザの不快感を高めることが明らかになっている。フィードバック内容を発表画面に表示する場合、常に発表者の視界に入るため強制的な情報の閲覧につながる。そこで、本研究における可視化はスマートフォンで行う。スマートフォンを用いることで、視聴者の好みの場所に設置できるとともに、発表画面を圧迫しないため、発表者の不快感を低減できると考えられる。

図2に、本研究のフィードバックに用いる各表情のアイコンを示す。フィードバック画面には、視聴者ごとの反応がアイコンとして匿名化された状態で表示され

Estimating and Communicating Audience Psychological States to Improve Online Speaker Comfort

† Rei Kikuchi †† Mizuki KOBAYASHI ††† Airi TSUJI ††† Kaori FUJINAMI

† Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology

†† Department of Bio-Functions and Systems Science, Tokyo University of Agriculture and Technology

††† Division of Advanced Information Technology and Computer Science, Tokyo University of Agriculture and Technology

る。頭部動作は、各アイコンを上下左右に動かすことで表現する。アイコンの更新間隔は、事前調査により7秒間隔が最も反応が見やすく発表にも集中できるとわかった。そのため、情報提示部では7秒ごとにサーバー部との通信を行い、アイコンを更新する。リアルタイム頭部動作検出における動作開始から判定までのラグは平均1.2秒であり、表情判定は1秒間に平均15.9回行われる。これは情報提示間隔の7秒を大きく下回るため、反応の検出速度が情報提示に与える影響は小さい。

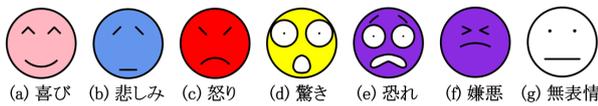


図2: フィードバックアイコン

3 ユーザ評価と考察

3.1 概要

オンライン発表にて、発表者の快適性を向上させる視聴者の反応の提示方法の調査を目的として、20代の12名(男性:8名, 女性:4名)を対象に評価実験を行った。発表資料として、被験者自身の趣味に関する5分程度のスライドを作成させた。被験者には、発表タスクと視聴タスクを行わせた。発表タスクで行う4条件を表1に示す。視聴タスクでは、eSense(左耳のみ)を装着し、割り当てられた発表者の発表条件に対応した状態(カメラやシステムの起動の有無)で発表を視聴する。発表者はカメラの共有を必須とした。

表1: 発表条件

	カメラの共有	システム	反応
条件1	×	×	
条件2	○	×	
条件3	×	○	全て
条件4	×	○	ポジティブのみ

一般的なオンライン発表方法と比較するため、条件1(カメラの共有なし)と条件2(カメラの共有あり)を行う。条件3は、検出した視聴者の全ての反応をフィードバックする方法である。条件4は、検出した視聴者のポジティブな反応のみをフィードバックする方法である。ポジティブな反応とは、頭部動作ではうなずき、表情では喜び(図2:a)を指す。条件3, 4では、無表情(図2:g)を視聴者のニュートラルな状態として表示する。驚きがポジティブかネガティブかの位置づけは曖昧であり、研究によって見解が異なるため[6]、本システムではポジティブな反応から除外する。

各提示方法の実施後、発表者に対し集中度や快適度、フィードバック内容などを5段階リッカート尺度で評価させ、全ての質問について、回答理由も記述させた。

3.2 実験結果

システムを使用した場合の各フィードバック内容の割合を表2に示す。ハイフンは仕様上存在しない項目である。また、提示方法の評価結果を表3に示す。数値は大きいほど良い評価を表している。なお、条件4ではシステムの不備により喜びが一切出現しなかったため、うなずきが発表者に与える影響について評価する。

3.3 考察

実験結果から、条件1はすべての項目で最低の評価であり、反応が得られないことが発表者にとってストレスであることがわかる。条件2から4を比較すると、

表2: システム使用時のフィードバック内容の割合(%)

	頭部動作			表情						
	うなずき	首振り	その他	喜び	悲しみ	怒り	驚き	恐れ	嫌悪	無表情
条件3	17.8	2.0	80.2	28.2	18.7	26.8	0.2	1.0	0.0	25.1
条件4	12.0	-	88.0	0.0	-	-	-	-	-	100.0

表3: 各提示方法の評価結果[平均値(標準偏差)]

質問項目	条件1	条件2	条件3	条件4
A. 反応を見たいと思ったか		4.42 (0.95)	3.92 (0.86)	4.17 (1.07)
B. 反応を見るのは簡単か		3.08 (1.44)	3.42 (1.04)	3.67 (1.11)
C. 反応に違和感がなかったか		3.58 (1.55)	2.58 (1.04)	3.50 (1.12)
D. 発表に満足しているか	3.42 (1.04)	3.92 (0.76)	3.83 (0.90)	3.67 (0.85)
E. 集中して発表できたか	3.92 (1.11)	4.42 (0.64)	3.92 (0.86)	4.00 (0.82)
F. 緊張せずに発表できたか	3.67 (1.18)	4.00 (1.00)	4.17 (1.07)	3.83 (1.07)
G. 不安を感じずに発表できたか	2.75 (1.09)	4.00 (1.00)	3.50 (1.19)	3.42 (1.04)
H. 視聴者との繋がりを感じたか	1.58 (0.49)	3.83 (1.21)	3.67 (1.11)	3.50 (0.87)
I. 快適に発表できたか	3.33 (1.11)	4.25 (0.92)	4.17 (0.80)	4.08 (0.76)

反応の見やすさ(項目B)と不安感(項目G)の結果が反比例している。反応の見やすさが不安感を高める原因とは考えにくいので、フィードバック内容に問題があったと考えられる。表2を見ると、条件3では複数の反応が満遍なく出現しているのに対し、条件4では発表者は視聴者のうなずきしか得ることができず不安を感じたと考えられる。フィードバック内容に対する違和感(項目C)は条件3が他の2つを大きく下回っており、ネガティブな反応の多さが理由として上げられていた。実際に、表2より条件3での表情の約半数がネガティブな反応であったが、映像を確認すると無表情をネガティブな反応に誤検出しているものがほとんどであった。結果として、快適度(項目I)について条件2が最高の値を示した。システム使用時(条件3,4)では、低評価の理由として表情に関する不満が多く挙げられ、高評価の理由としてうなずきが得られたことが多く挙げられていたことから、表情の検出精度が本システムの大きな課題であると考えられる。

4 おわりに

本稿では、オンライン講演者の快適性を向上する視聴者の反応の可視化方法を提案した。イヤホンとカメラを用いて視聴者の頭部動作と表情を検出し、発表者に対して視聴者の反応を可視化することで発表の快適度を高めることを目指した。実験の結果より、システムが反応の見やすさに大きく貢献したが、表情の検出に課題が多く、カメラ映像によるフィードバックが最も快適であると示された。今後は表情の検出方法を改善すると共に多様な可視化方法について検討する予定である。

謝辞

本研究はJSPS 科研費(19H01719)の支援を受けた。

参考文献

- 国立情報学研究所. 遠隔授業に関するアンケート調査の概要, 2020. https://www.nii.ac.jp/event/upload/20200914_Report.pdf 最終アクセス 2021/12/16.
- Fahim Kawsar. Earables for personal-scale behavior analytics. *IEEE Pervasive Computing*, Vol. 17, No. 3, pp. 83–89, 2018.
- Prasanth Murali. Affectivespotlight: Facilitating the communication of affective responses from audience members during online presentations. In *In Proc. of CHI2021*, pp. 1–13, 2021.
- Octavio Arriaga. Perception for autonomous systems (paz). *arXiv preprint arXiv:2010.14541*, 2020.
- 藤原康宏ほか. コンピュータ利用時の不快感を利用した警告インタフェースの提案. *情報処理学会論文誌*, Vol. 52, No. 1, pp. 77–89, 2011.
- 西川一. 「驚き」はニュートラル感情なのか? 驚きの状態と特性の PANAS-X による比較分析. *心理学叢誌*, Vol. 7, pp. 1–9, 2012.