

オンライン会議中の参加状態推定と 音声変換提示による参加促進手法の提案

城谷知葵^{†1}

九州大学大学院システム情報科学研究院

中村優吾^{†3}

九州大学大学院システム情報科学研究院

難波洸也^{†2}

九州大学大学院システム情報科学研究院

荒川豊^{†4}

九州大学大学院システム情報科学研究院

1. はじめに

会議においては発言やリアクションなど場への参加度が高いことが望ましい。発言やリアクションなどの情報によって相手の状況を理解することができ、会議を円滑に進めることができる。これは近年盛んであるオンライン会議においても同じである。オンライン会議は場所を気にすることなく参加できるため、従来の対面会議に比べて会議への参加が容易になった。オンライン会議において参加度を上げるために効果的な方法はカメラとマイクを ON にすることである。カメラを ON にすることで表情や傾きなどの仕草を確認することができ、相手の反応を理解することが出来る。つまり参加者の情報を対面会議同様に得ることができる。

しかしオンライン会議中にカメラとマイクを ON にすることは参加者に測定されている・監視されているなどの印象を与えてしまう。そのため対面会議よりも心理的な負荷が大きくなる。さらに話し手・聞き手ともに無意識のうちに自然な振る舞いが阻害されることや心理的な緊張感による疲れも懸念される。会議などのコミュニケーションにおいて心理的な安全性を確保することは会議を活発にする上でも会議の質を高める上でも大切である。

上記の問題を解決するにはカメラとマイクを OFF にする方法が考えられる。カメラを OFF にすることによって自分と相手の顔が見れなくなり、心理的負担を減らすことができる。しかし、カメラやマイクを OFF にするとノンバーバル情報(身振り・手振り、息遣い、視線、布ズレ音、椅子に座り直す音など)が欠落し、他の参加者の反応を理解することが困難になる。その結果会議を円滑に進めることが難しくなり、会議の質が低下してしまう。

カメラやマイクを OFF でもリアクションを伝えるために

Teams, Zoom, Discord, Slack, hangout, WebEx などの各種コミュニケーションツールはチャットやエモート、他者のチャットへのいいねなどの機能が存在している。しかしこれらの機能は参加者自らが用いる必要があり、対面会議時のリアクションのように自然に発生するものとは異なる。そのため参加者に心理的負担を与えることになる。よって参加者のリアクションを自動でフィードバックするシステムがあれば、情報開示のような負荷やリスクの影響なく、会議を円滑に進めることができる。

しかしカメラ OFF でのオンライン会議においてはリアクションを伝えることよりもオンライン会議特有の問題を先に解決すべきである。1つ目はカメラを OFF にすることによって参加者の緊張感が減ることである。対面会議ではオンライン会議のカメラほどではないが、他の参加者に見られているかという緊張感を少なからず与えることができる。しかし、オンライン会議でカメラを OFF にすると他者の視線が完全にカットされるため、会議への集中力が減ってしまうことがある。2つ目は認知資源が他のことに奪われやすいことである。オンライン会議に使用しているデバイス(PCやタブレット、スマートフォンなど)を前に置いていることによってポップアップ通知やその他の機能に気が散ってしまう。また個々の作業空間によってはデバイスの外にも気が散るものが存在し、注意が別の方向に向いてしまうことがある。

以上のことからリアクションを自動で伝えるシステムの開発より先にオンライン会議中に他所にそれる注意を引き戻すシステムが必要である。

Arakawa らは、ビデオベースの学習において非集中状態を音声変化によって再び集中状態に戻す手法を提案している [1]。Mindless Attractor は実験参加者の視線によってビデオへの集中力を測定しており、PC から視線が外れている(スマホなどを見ている)場合には、ビデオ音声のピッチを変化させることで集中状態に引き戻している。しかし上記の手法ではオンライン会議への参加と PC 上での内職(ネットサーフィンや動画視聴)を識別することが困難である。ま

^{†1} TOMOKI JOYA, Kyushu University

^{†2} KOYA NAMBA, Kyushu University

^{†3} YUGO NAKAMURA, Kyushu University

^{†4} YUTAKA ARAKAWA, Kyushu University

た音声変化に用いているのはピッチだけであり、音声変化が1種類だけであると馴化してしまう可能性がある。

そこで本研究では、オンライン会議において参加者のPC情報からオンライン会議への集中状態を推定し、その状態に応じて音声を変化させ、参加者をオンライン会議へと注意を引き戻す手法を提案する。オンライン会議への集中状態を推定するために用いるのはPCのプロセス情報だけであるため、カメラやマイクなどによる測定感を軽減することができ、参加者に不快感を与えることがなくなる。さらにプロセス情報によってPC上での内職を識別することもできる。また、非集中状態と推定された参加者に対しては、オンライン会議の音声に対し、ピッチやリバーブといった音変化をさせることにより、参加者に強制感を与えずに会議への注意を引き戻すことができる。

オンライン会議中、非集中状態と推定された参加者に対し、ピッチやリバーブといった音声変換提示をすることで、参加者の注意を会議へ引き戻し、集中状態へ戻る効果があるかを明らかにするための評価実験を実施した。7名の実験協力者に対し、普段オンライン授業に参加するときと同じ状況下において、ビデオ講演を聞いてもらった。音声変換ありとなしの場合の集中状態時間を比較した結果、音声変換を体験した実験参加者5名中4名に対し、ビデオ講演への集中時間割合が増加した。また、実験参加者は、ピッチやリバーブといった音声変換機能の詳細まで意識的に気づかなくとも、集中状態への復帰行動を取ることが確認できた。

2. 関連研究

本章ではユーザの状態をセンシング技術によって取得する研究と、聞き手を集中状態に引き戻す研究について述べる。

2.1. ユーザの状態を認識する研究

ユーザの集中状態を外部センサを用いて取得する研究が多く行われている [2,3]。文献 [2] では、被験者が装着したウェアラブルデバイスから生理学的なデータを取得することで被験者の参加状態をセンシングしている。また文献 [3] では外部センサをPCに装着することで視線情報を取得している。しかし、外部センサを使用するためにはコストがかかる上に、会議のためにセンサを装着することは現実的ではない。

オンライン会議で用いるPCに付属しているカメラでユーザの状態をセンシングする研究も行われている [4-7]。文献 [4] ではオンライン講義を受講している学生の表情をPCのカメラで撮影し感情を分析している。文献 [5] では表情に加えて顔や眉の動きをセンシングしている。PCのカ

メラを用いる手法は外部センサを新しく導入する必要がないため手軽に実施できるが、カメラの画角に顔を収める必要がある。また、オンライン会議ではカメラをオフにすることもあり、その場合は自分がカメラの画角内に収まっているか確認することができない。

そこで本研究ではCPU使用率やマウスの挙動などのユーザのPC情報に着目した。PC情報を用いることで画角を気にすることなくデータを収集することが可能となる。本研究ではActivityWatch [8] を参考にユーザのPC情報を取得し参加度を推定する。

2.2. 非集中状態のユーザを集中状態に戻す研究

次に参加度が低い聞き手に対してのフィードバック手法について述べる。文献 [9] では視線情報から参加度を推定し、参加度が低い聞き手に対して試聴している動画を止めて画面上にアラートを表示する。文献 [10] ではプレゼン資料上に四角形の枠を表示して聞き手に注意を促している。このようなPCの画面に提示する手法では聞き手がPCの画面を見ている必要があるが、別の画面を見ている場合には有効ではない。

そこで本研究では音による提示によって聞き手へフィードバックする。文献 [1] では、聞き手が聞く音声のピッチを変換することで無意識的に会議に集中するよう促す手法を提案している。本研究でも音声変換をすることにより参加度が低い聞き手の参加度を向上させることを目標とする。

3. 提案システム

本章では、オンライン会議中において、注意が会議に向けられていない非集中状態にいる参加者を、会議へと注意が向いている集中状態に引き戻すシステムの概要について記す。まずはじめに、3.1節でオンライン会議における参加者の状態や遷移について定義し、ターゲットとなる参加者の状態と目指すべき参加者の状態について述べる。次に、3.2節でターゲットとなる参加者の状態のセンシング推定手法について述べ、3.3節でセンシング結果に基づいてシステムが参加者へどういった介入をすべきかについて述べる。最後に、3.4節で上記要件を満たしたプロトタイプ開発の詳細について述べる。

また、本研究で提案するシステムの適応対象となるオンライン会議の想定環境について下記に記す。

- ・会議形式：先生（話し手）と生徒（聞き手）といったような役割が分かれている会議
- ・発表資料：あり。話し手は画面共有をしている
- ・ハードウェア：PCやスマートフォンといった視覚と聴覚機能を有するデバイス
- ・ソフトウェア：TeamsやZoomといったオンライン会

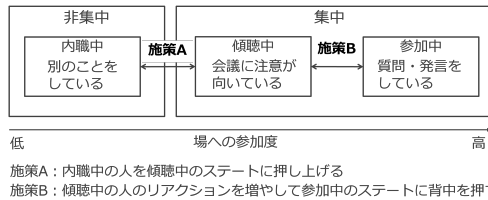


図1 オンライン会議中における参加者の状態遷移

議システム

- ・参加者のマイクの状態：問わない
- ・参加者のカメラプレビュー：OFF

3.1. オンライン会議における参加者の状態遷移

図1に示すように、オンライン会議における参加者の状態は、非集中状態と集中状態に大別できる。

非集中状態とは、参加者の注意がオンライン会議から外れており、心身的な理由から集中度が低い・乱れている状態、または会議システムを実行している機器の別のアプリケーションに認知資源を割いている状態、会議システムを実行している機器の外（参加者が置かれている環境）に認知資源を割いている状態を示す。代表的な行動としては、会議には直接関係のないアプリケーションを使用する（メールチェックやWeb検索）、会議に参加せずにスマートフォンを操作するなどが挙げられる。

集中状態とは、オンライン会議に注意を向け、高いパフォーマンスが発揮されている状態である。集中状態は参加者の状態を傾聴中・参加中の二つに分類することができる。傾聴中はその名の通り、話に耳を傾けている状態を示す。参加中の代表的な行動としては、発話をする・チャットに書き込む・エモートを押すなどが挙げられる。

図1に示したように、左側に行くほど場への参加度が低くなり会議への貢献が低下し、反対に右側へ行くほど場への参加度が高まり、会議への貢献が向上すると言える。1章でも触れた通り、我々は会議の参加者の場への参加度を高めるため、会議の参加者をこの一番右側の参加中の状態へと促すことを目指したい。しかし先に挙げた様々な要因から参加者の集中が保たれず、参加者の注意がオンライン会議から逸れて、非集中状態に落ちてしまうことがある。

そこで、オンライン会議の参加者の場への参加度を高めるためには、図2のようにまず参加者の状態が集中・非集中状態のどちらにあるのかを検出し、参加者の状態に応じて、“施策A：内職中の人を傾聴中のステートに押し上げる”と、“施策B：傾聴中の人リアクションを増やして参加中の状態に背中を押す”を使い分ける必要がある。これらの施策のうち、本研究ではオンライン会議中に注意が逸れてしまう参加者の注意を引きもどすシステムを提案する。また本論

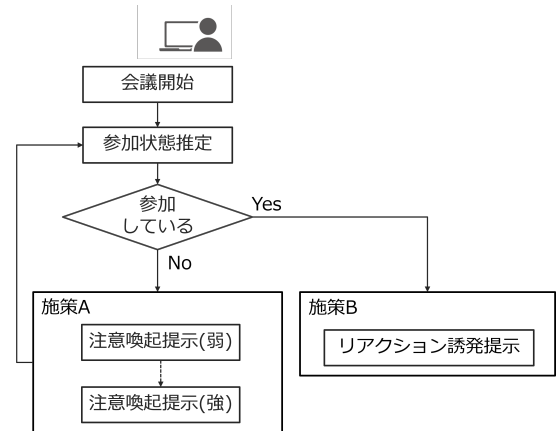


図2 参加度を高めるための全体フロー

文では、デバイスは普段使用しているノートPCを想定し、代表的な非集中状態の行動として、内職作業を取り上げる。

3.2. 参加状態推定ブロック

参加状態推定ブロックではユーザの参加状態を推定する。図3に参加状態推定ブロックの概要を示す。

まずユーザのPCからPC情報を取得する。PC情報はCPU使用率・Active Window（PC上で作業中や選択中の画面）・AFK時間（マウスやキーボードを触っていない秒数）で構成されている。PC情報を取得するプログラムはpythonで記述しユーザのPCで動作させる。プログラムは1秒毎にCPU情報を取得しcsvファイルに書き込む。

次に取得したデータをそれぞれ前処理ブロックに入力する。CPU使用率データに対しては、全て値が0であるCPUプロセスを削除した後を使用したWeb会議システムのCPU使用率で正規化する。Active Windowデータに対しては、使用したWeb会議システムかそれ以外かの2値に分類する。AFKデータに対しては、閾値を設定して2値化する。本研究では閾値は3秒とする。

特徴量抽出ブロックでは前処理したデータを窓で区切り、それぞれ特徴量を抽出する。本研究では窓サイズは4サンプル、スライド幅は2サンプルとする。CPU使用率データからは最大値と最小値の差・標準偏差を計算する。Active

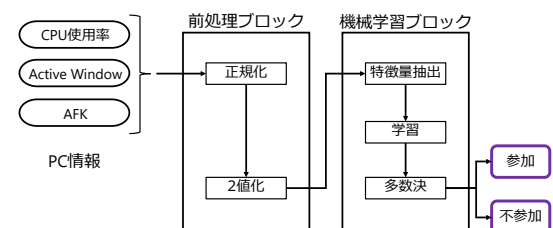


図3 参加状態推定ブロックの全体像

Window と AFK データからは合計値を計算する。

機械学習ブロックでは教師あり学習を用いて特徴量からユーザの参加状態を推定する。正解ラベルは参加状態と不参加状態の 2 値である。機械学習モデルは限定しないが、本研究では lightGBM を用いた。

最後に推定されたラベルについて数サンプル毎に多数決をとり、最も個数が多かった状態を最終状態として出力する。本研究では 5 サンプル毎に多数決をとる。

3.3. 施策 A: どのように参加者の注意を引き戻すのか

施策 A の目的は、オンライン会議の非集中状態における参加者に対し、再びオンライン会議へ注意を向けさせることである。この際、我々が重要視するのは、参加者に対して強制感を与えることなく、感情や心理に訴えかけてさりげなく行動を促したいということである。よって、システムの介入によって参加者が作業を中断する必要がある・不快な思いをするようなことは避けたい。また、オンライン会議では、非集中状態にある参加者は常に PC 画面を見ているとは限らず、PC 画面を見ていたとしても打ち合わせアプリケーションではない他のアプリケーションを閲覧している状況も考えられる。こういった状況下で、参加者に対して有効な提示をするには、視覚的な提示よりも聴覚提示の方が適していると考えた。先にも述べたが、文献 [1] では、聞き手が聞く音声のピッチを変換することで無意識的に会議に集中するよう促す手法を提案している。しかし、ピッチ変換は音声変換の開始と終了時に注意を惹くことはできるが、変換中は馴化してしまい、効果としては不十分であると予測されたため、音声のピッチ変換だけでなく、音声変換中ずっと効果の持続性が期待できると考えられるリバーブ変換に着目した。

3.4. 施策 A: プロトタイプ実装

図 4 に提案手法の一連の流れを示す。我々は、非集中状態と判定された場合には、会議の音声を変換するプロトタイプ実装を行った。音声変換は 3 種類（ピッチ、リバーブ弱、リバーブ強）用意した。非集中時間の継続時間に応じて、ピッチ→リバーブ弱→リバーブ強と段階的に刺激の強度が変化するようにし、参加者が非集中状態になるごとに、繰り返されるようにした。オンライン会議システムに入力された話し手の音声を、仮想マイクを介すことで聞き手の PC にインストールされたプログラムに入力して音声変換を施す。

使用 OS は macOS のみを対象としており、音声変換は pedalboard, CPU 情報取得には psutil を用いた。PC 情報は 1 秒間に 1 回取得し、音声変換は非集中状態の場合は 15 秒に 1 回施した。

4. 実験

4.1. 参加度推定性能

参加度推定の機械学習モデルを作成するために、教師データ取得実験と評価用データ取得実験を行った。いずれの実験でも、実験参加者のパソコンに「CPU 使用率」・「Active Window」・「AFK」を取得するソースコードをインストールし、コードを実行した状態で、Microsoft Teams を使用したオンライン会議を行った。

まず教師データ取得実験のため、オンライン会議に入った 5 人の実験参加者に指定した行動をとってもらった。参加している状態では「Teams の画面を真面目に見る」・「議事録をとる」・「Teams のチャットに質問を書く」の行動をしてもらった。不参加の状態では「ネットサーフィンする」・「資料を作成する」・「YouTube 動画を見る」などの行動をしてもらった。なお、パソコンを使用する以外の不参加状態はここでは取得していない。

次に評価用データ取得実験を行った。2 名の実験参加者を参加している状態と不参加の状態に分け、オンライン会議室に入ってもらった。教師データ取得実験とは異なり、実験参加者には指定した行動ではなく自由な行動をとってもらった。実際は、参加状態では「Teams の画面を真面目に見る」、不参加状態では「ネットサーフィンする」の行動があった。

取得した教師データを用いて 3.2 章の手順で機械学習モデルを作成した。前処理とアンダーサンプリングの結果、機械学習モデルに使用したデータは参加状態と不参加状態いずれも 1770 サンプルであった。作成した機械学習モデルを評価データで検証したところ正解率は 0.776 であった。

4.2. 施策 A: 予備実験

ピッチやリバーブ音声変換によって、参加者の注意をオンライン会議に引き戻す効果があるかを検証するにあたり、まず参加者がピッチやリバーブ音声変換に気づくかどうか、またその機能はシステムの挙動として受け入れられる（受容性）のか、ピッチとリバーブ変換はどちらが効果的なのか

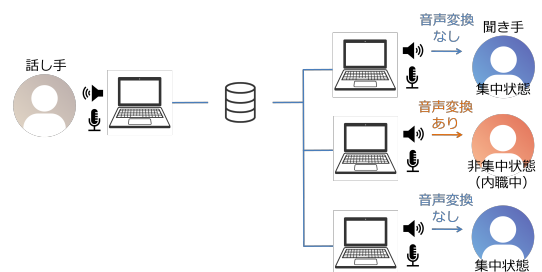


図 4 提案手法の一連の流れ

といった知見を得るための予備実験を行った。

4.2.1 実験概要

音声変換データを3種類(ピッチ、リバーブ弱、リバーブ強)用意した。14名の実験参加者に対し、3種類の音声変換データのうち1種類を体験するようにグループ分けした。内訳はピッチ変換(5名)、リバーブ弱変換(5名)、リバーブ強変換(4名)である。テスト中、実験参加者は変換された音声を聴きながら、手元のPCで文章を読む作業を行った。テスト終了後に音声内容理解の確認テストと作業負荷や音声に対する印象についてのアンケート調査を行った。

4.2.2 実験結果

アンケート調査によって、ピッチよりもリバーブ音声変換の方が気づきやすいという結果が得られた。また、テスト中に変換された音声を聞くことによる認知負荷については、変換されていない音声を聞くときと比較して大きな変化は見られなかった。変換されていない音声と変換された音声を聞く場合を比較して、音声内容の理解度や参加者の手元PC画面でのタスクパフォーマンスに大きな違いは見られなかった。

4.2.3 まとめと本実験に向けて

アンケート調査から明らかになった、ピッチよりもリバーブ変換の方が気づきやすいという結果から、音声変換した際に参加者に与える刺激の強度としてピッチ<リバーブの関係があると考えられる。また、テスト中に変換された音声を聞くことによる認知負荷については、変換されていない音声を聞くときと比較して変化はなかったことから、ピッチやリバーブによる音声変換提示が参加者の心理や感情を阻害することはないと言える。ピッチ変換は音変換の開始と終了時に注意を惹けるが、変換中は馴化してしまい、また元々の話し手の音声に依存する部分も重なって効果が不十分であることが明らかになった。一方でリバーブ変換にはその懸念がなく、音変換開始、終了時に注意を惹けるのはもちろんのこと、音声変換中ずっと効果の持続性が期待できると考えられる。よって、本実験では、非集中時間の継続時間に応じて、ピッチ→リバーブ弱→リバーブ強と段階的に刺激の強度が変化するようにし、参加者が非集中状態になるごとに、繰り返されるようにした。

4.3. 施策 A：本実験

4.3.1 目的

予備実験の結果および得られた知見を踏まえ、オンライン会議において参加者が別アプリケーションに認知資源を割いている状態において、音声変換提示が、参加者の注意を引き付けて別アプリケーション作業を中断させる効果があるか否かを明らかにするための実験を行った。

4.3.2 タスク

実験参加者は、各自が普段使用しているノートPC内の打ち合わせアプリケーション(Teams)を通して、用意された音声ビデオを学習するように指示された。

4.3.3 素材

15分程度のUIST2019のプレゼンテーション発表動画を用意し、実験参加者に学習状況を再現させた。

4.3.4 実験参加者

実験はファシリテータ1名と実験参加者7名で行った。ファシリテータ1名と実験参加者7名は、共に九州大学の学生である。

4.3.5 実験手順

実験参加者は、椅子に座り、机の上にPCを置き、有線のイヤホンまたはヘッドフォンを装着してもらい、普段オンライン授業に参加するときと同じ状況下で実験を行った。音声変換機能がある場合とない場合での、違いを比較するため音声ビデオ学習を2回実施した。その際、順序効果をなくするため、最初に音声変換機能を体験するグループと2回目体験するグループに実験参加者を分けた。図5に実験の流れを示す。まずはじめに、実験参加者に、これから音声ビデオ学習を行ってもらう旨を伝えた(事前説明)。音声ビデオ学習は2回に分けて行い、ビデオ学習後に作業負荷や音声に対する印象についてのアンケート調査を行った。すべてのテスト終了後、事後インタビューを行った。

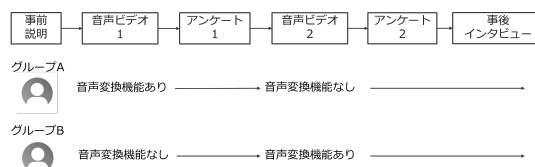


図5 実験の流れ

4.3.6 測定データ

音声ビデオ学習中は、参加者のPCのCPU使用率、ActiveWindow、AFKのログを採取した。学習後に行ったアンケートについては、ポジティブ/ネガティブ感情の変化や音声変換に気づいたか否かを調査した。すべてのテスト後の事後インタビューによって、参加者の感想を集めた。

4.3.7 実験結果

図6に、音声変換機能なしとありの場合の集中状態時間割合の変化を示す。7名の実験参加者のうち、複数回にわたって音声変換を体験したのは5人(ID:3,4,5,6,7)であった。図6のグレー色で示した2名(ID:1,2)はテスト中ずっと集中していたため、音声変換機能をほとんど体験しなかった。音声変換を体験した5人中4人が、音声変換提示をすることで、音声ビデオへの集中時間の割合が上がった。また、音声ビデオへの集中状態へ戻ったのは、リバーブ変換発動時

が最も多く、ピッチ変換だけでは音声ビデオへの集中状態に戻らなかった。また、テスト後の事後インタビューから、実験参加者 7 人中 2 人が、音声変換に気づき、残りの 5 名は気づかなかった。音声変換を体験した 5 人中、リバーブ変換機能に気づいた人は 1 人しかいなかった。実験参加者 7 人中 2 人が、音が変化した際に Teams 画面を確認したとコメントした。音声変換による、ポジティブ/ネガティブ感情の大きな変化は見られなかった。

参加者ID	音声変換なし(%)	音声変換あり(%)	変化率(%)
1	87.6	96.3	109.9
2	100	100	100.0
3	21.1	27.6	130.8
4	65.1	69.6	106.9
5	29.9	41.0	137.1
6	64.1	45.6	71.1
7	25.2	56.2	223.0
平均	56.1	62.3	111.0

図 6 実験参加者の集中時間割合の変化

4.3.8 まとめと今後の課題

音声変換を体験した 5 名中 4 名に対し、音声ビデオへの集中時間割合が増加する結果が得られたことから、本システムによる音声変換提示は、参加者の注意を惹き戻す効果があると言える。音声ビデオへの集中状態へ戻ったのは、リバーブ時が多かった理由としては、実験参加者が、そもそもピッチ変換に気づいていない可能性が高い。本実験で使ったプロトタイプは、ピッチ変換時の音声の変化量が小さかったために参加者に気づかれにくかったことや、実装の都合上、変換時にノイズ（ぷつぷつ途切れる音）が混ざってしまったことが原因と考えられる。音声変換を体験した 5 人中、リバーブに気づいた人は 1 人だけだったが、意識的に気づいてなくとも実際に集中状態への復帰行動は複数の実験参加者の間で観測されていた。最後に、音声変換によるポジティブ/ネガティブ感情の大きな変化は見られなかったことから、本システムによる音声変換に対する参加者の受容度（受け入れやすさ）に関しては問題ないレベルであったといえる。

今後の課題としては、まず今回の本実験におけるピッチ音声変換時にノイズが混ざってしまったことで、純粋にピッチ変換だけの効果ではなく、ピッチ+ノイズの効果になってしまった点が挙げられる。また、音声変換提示は、話し手の声の性質や聞き手との関係性（聞き手にとって面識のない人が発表する際はより気づきにくなる可能性）、参加者が使用しているヘッドフォンの性能、出力方法（ヘッドフォン or スピーカ）といった外部環境に注意する必要がある。

5. おわりに

本研究ではオンライン会議においてカメラを ON にすることなく参加者を集中状態にする手法を提案した。そのための第一歩として内職中の人を傾聴中のステートに押し上げる手法（施策 A）の有効性を検証した。提案手法では、まず参加者の PC から取得した CPU 使用率、Active Window、AFK 情報を機械学習モデルに入力し参加状態を推定する。次いで非集中状態と判定された参加者の音声に対してピッチシフト、リバーブの加工を施す。実験の結果より、音声変換を体験した実験参加者 5 名中 4 名に対し、ビデオ講演への集中時間割合が増加したこと、音声変換を施しても無意識的に実験参加者を集中状態へ復帰できることを確認した。今後は、傾聴中の参加者のリアクションを増やして参加中の状態に背中を押す手法（施策 B）について取り組む予定である。

謝辞

本稿の研究を行うに際し、ご助言いただいたソニーグループ株式会社川名氏、森氏、朽木氏に感謝の意を表する。

参考文献

- [1] Arakawa, R. and Yakura, H.: Mindless Attractor: A False-Positive Resistant Intervention for Drawing Attention Using Auditory Perturbation, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–15 (2021).
- [2] Di Lascio, E., Gashi, S. and Santini, S.: Unobtrusive Assessment of Students' Emotional Engagement during Lectures Using Electrodermal Activity Sensors, Vol. 2, No. 3 (online), 10.1145/3264913 (2018).
- [3] Yao, N., Brewer, J., D'Angelo, S., Horn, M. and Gergle, D.: Visualizing Gaze Information from Multiple Students to Support Remote Instruction, *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, New York, NY, USA, Association for Computing Machinery, p. 1–6 (online), 10.1145/3170427.3188453 (2018).
- [4] Yao, N., Brewer, J., D'Angelo, S., Horn, M. and Gergle, D.: Visualizing Gaze Information from Multiple Students to Support Remote Instruction, *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, New York, NY, USA, Association for Computing Machinery, p. 1–6 (online), 10.1145/3170427.3188453 (2018).
- [5] Murali, P., Hernandez, J., McDuff, D., Rowan, K., Suh, J. and Czerwinski, M.: AffectiveSpotlight: Facilitating the Communication of Affective Responses from Audience Members during Online Presentations, *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY,

- USA, Association for Computing Machinery, (online), 10.1145/3411764.3445235 (2021).
- [6] Watanabe, K., Soneda, Y., Matsuda, Y., Nakamura, Y., Arakawa, Y., Dengel, A. and Ishimaru, S.: Discaas: Micro behavior analysis on discussion by camera as a sensor, *Sensors*, Vol. 21, No. 17, p. 5719 (2021).
- [7] Chen, C., Arakawa, Y., Watanabe, K. and Ishimaru, S.: Quantitative Evaluation System for Online Meetings Based on Multimodal Microbehavior Analysis, *Sensors and Materials*, Vol. 34, No. 8, pp. 3017–3027 (2022).
- [8] : ActivityWatch, <https://github.com/ActivityWatch/activitywatch>.
- [9] D’Mello, S., Olney, A., Williams, C. and Hays, P.: Gaze tutor: A gaze-reactive intelligent tutoring system, *International Journal of human-computer studies*, Vol. 70, No. 5, pp. 377–398 (2012).
- [10] Sharma, K., Alavi, H. S., Jermann, P. and Dillenbourg, P.: A Gaze-Based Learning Analytics Model: In-Video Visual Feedback to Improve Learner’s Attention in MOOCs, *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge*, LAK ’16, New York, NY, USA, Association for Computing Machinery, p. 417–421 (online), 10.1145/2883851.2883902 (2016).