

頷きのリアルタイムフィードバックによる ビデオ会議支援手法の提案

徳原 耕亮^{1,a)} 荒川 豊¹ 石田 繁巳^{2,3}

概要:

本論文では、ビデオ会議における円滑なコミュニケーションスキル習得支援を目的として、会議中に表出する無意識の仕草を定量的に評価し、数値として参加者にフィードバックするシステムの提案、実装、評価を行う。近年の働き方改革やコロナ禍の影響により、テレワークやオンライン会議の頻度が高まっている。オンライン会議の場合、カメラがオフであったり、オンであっても顔がアップされた状態であることが多く、対面会議では感じ取られる阿吽の呼吸や場の雰囲気をつらいつらいつらという課題がある。また、誰かがずっと話していたり、一度も発話しない参加者が出てくることも多い。ビデオ会議を対面会議に近づけるために、参加者自身がこれまで無意識に行っていた、非言語コミュニケーション（ノンバーバル・コミュニケーション）を意識し、能動的に情報を発信していくことが大事ではないかと考えている。そこで我々は、頷きや笑いなどミーティング中に表出する仕草を認識するシステムの開発を行っている。本稿では、仕草の中でも発現頻度が最も高い、頷きにフォーカスし、参加者全員の頷きをリアルタイムに認識した上で、リアルタイムにフィードバックするシステムを開発し、リアルタイムフィードバックが参加者に与える影響について調査した結果について報告する。

1. はじめに

近年の働き方改革やコロナ禍の影響により、様々な企業でビデオ会議が広がっている。企業だけではなく、大学などの教育現場においてもビデオ会議が広く用いられるようになってきている。それにより、ビデオ会議のメリットである、通勤時間の削減やオフィスコストの削減などが見直されおり、コロナ禍収束後も、ビデオ会議の活用の流れは止まらないと考えられる。

しかし、Skype や Teams, WebEX といったビデオ会議システム自体は、コロナ禍よりも前から存在していた。それにも関わらず、広まっていなかったのはなぜだろうか？そこには、対面会議との体験の差があるのではないかと考えられる。例えば、ビデオ会議環境では参加者側で任意にカメラやマイクをミュートすることができ、ミュートにしてしまうと話を聞いているのか、そこに人がいるのかが相手に伝わらず、話者が話しづらいといった問題がある。ま

た、ビデオ会議システムでは、全員の様子を確認しづらい、画面共有を行っている間に他の参加者の様子を確認しづらいといった問題もある。プライバシーや帯域の問題を理由に、ビデオをオフにして講義に参加されると、学生が本当に学んでいるのか一切わからないというのは、全教員が体験していると思われる。一方、対面会議においては、全体を俯瞰できるため、誰が誰を見ていて、誰がどういう理解を示しているのかといった場の状況がわかる。アイコンタクトや微妙な仕草によって、意見を求められているのかが伝わったり、割り込んでいくタイミングを掴むことが可能である。参加者側の視点では、対面の会議よりも会議への集中力が低下してしまったり、頷きや腕組みのような無意識下で行っている動作への意識の低下による参加態度に生じる問題が考えられる。

これまで、このような対面会議と遠隔会議のギャップを埋めるために、さまざまな研究が行われてきている。代表的なものとしては、テレプレゼンスロボットを用いた手法 [1,2] が上げられる。テレプレゼンスロボットは、Double シリーズ^{*1}や Temi^{*2}など、すでに商用ロボットが市販されており、筆者らの研究室でも利用している。確かに、その場にいるような体験となり、遠隔会議のデメリットをなく

¹ 九州大学大学院システム情報科学研究院

ISEE, Kyushu University

² 公立はこだて未来大学システム情報科学部

Sch. Systems Information Science, Future Univ. Hakodate

³ 九州大学システム LSI 研究センター

SLRC, Kyushu University

a) tokuhara.kosuke@arakawa-lab.com

^{*1} <https://www.doublerobotics.com/>

^{*2} <https://www.robotemi.jp/>

すことが可能であるが、高価なロボットを配備する必要があり、さまざまな場所で利用できる手法とは言えない。他の研究としては、視線認識装置と眼球ディスプレイを組み合わせ、アイコンタクトを物理的に実現する手法 [3,4] や、誰が喋りそうかを推定して提示することで発話衝突を防ぐ手法 [5,6] など提案されているが、いずれも何らかの装置や高性能な計算環境が必要となる。

それに対して、筆者らは、装置で常時支援するのではなく、気づきを与え、自分自身のノンバーバル・コミュニケーションを振り返ったり、他者と比較することで、人のスキルとして身につける方が良く考えている。そのための要件としては、1) 専用のパソコンやソフトウェアではなく、広く使われているビデオ会議システムと共存できるシステムであること、2) 専用のハードウェアや高性能な計算機を必要とせず、一般的なユーザがビデオ会議に利用するパソコンで、ビデオ会議に支障なく使えること、3) それでいて、分析とフィードバックをリアルタイムに実行できること、という3点が必要であると考えた。しかしながら、現状、これらの条件を満たすビデオ会議支援システムは存在しない。

そこで、本稿では、ブラウザ環境で実行できるビデオ会議システム（例えば、Google Meet など）に対して、リアルタイムな映像解析とフィードバックを実現する、ブラウザ拡張システムを設計開発した。開発システムは、Chrome Extension として開発し、chromium をコアエンジンとする、Microsoft Edge や Brave 等のブラウザでも動作する。実装には、Web ブラウザで動作する軽量顔検出フレームワーク MediaPipe Face Mesh を用いている。今回、複数人で実施するビデオ会議において、表出頻度の高い“頷き”に特化し、参加者それぞれの頷き数をリアルタイムにカウントして、ビデオ上に重畳表示する機能を実装した。そして、その効用について、被験者実験を通じて明らかにした。今回の実験では、ビデオありの状態を前提としているが、今回開発したシステムを用いれば、顔画像を消し、統計情報（頷き以外には、発話量やカットイン率、笑顔時間、腕組回数など様々なものが考えられる）に置き換えると言った映像操作もリアルタイム実行可能であり、ビデオオフ状態と同等の環境を作ることも可能である。

今回、4名の被験者に協力してもらい、システムを利用しないビデオ会議と、システムを利用したビデオ会議での体験の違いについて評価した。その結果、提案システムによってリアルタイム認識とフィードバックを実現できたことを確認するとともに、その機能によって、参加者自身の振り返り、気づきに役立ったこと、さらにその後の仕草やコミュニケーションに影響を与えることを明らかにした。

2. 関連研究

遠隔会議や会議支援に関連する研究として、テレプレゼ

ンスを示すための代理人ロボットに関する研究 [1,2] や遠隔会議中の視線情報の伝達に関する研究 [3,4,7-10]、ビデオ会議における発話衝突の回避に向けた研究 [5,6,11]、会議時の360度カメラ映像を用いたミーティング分析に関する研究 [12,13] などが報告されている。

遠隔会議では会議参加者がディスプレイ上に表示される形式が一般的であるが、「その場にいる」ことの重要性は古くから注目されており、あたかも「その場にいる」という感覚を与えるためのテレプレゼンス技術が研究されている。代理人ロボットを用いる手法 [1,2] はその最たる例であり、遠隔会議の支援としても有効であると考えられる。

遠隔会議においては参加者がどこを見ているかという非言語情報が重要であることが報告されている [14]。このため、遠隔会議中の視線情報を伝達する技術の研究が進められている。Ohno はカメラ画像から取得した頭部の向きとアイトラッカを用いて取得したディスプレイ上の視線位置を共有してアイコンタクトを実現するシステムを示している [7]。小峯らは、代理人ロボットを利用した遠隔会議システムを想定し、遠隔会議システム上での視線を取得してロボットの視線によりアイコンタクトを実現する手法を示している [8]。大槻らは、アイトラッカで取得した視線情報を表示するデバイスとして、眼球ディスプレイ ThirdEye を提案している [3,4]。この手法では遠隔会議用のディスプレイ上部などに眼球ディスプレイを具備させ、相手の視線に基づいて目の動きを再現することで、視線を交わしながらの遠隔会議を実現している。また、Tausif らは、顔の向き及び視線情報に基づいてカメラ位置を調整可能な対面式遠隔会議システムを提案している [10]。

しかしながら、代理人ロボットを用いる手法、視線情報を提示しながらの遠隔会議はロボットやアイトラッカなどの特殊なデバイスを必要とするため、場所の制約も少なく手軽に参加可能という遠隔会議の優位性が制限される。

遠隔会議では限られた情報から相手の状況を推定する必要があり、通信遅延の影響も受けるため、発話が衝突しやすいという問題も発生する。このような発話衝突の解決に向けた研究として、玉木らは、遠隔会議において発話衝突を防ぐために発話欲求を推定して提示する手法を提案している [5]。参加者が誰かに対して発言をしようとするときに発生する予備動作に着目し、参加者が予備動作を検出して提示することで発話衝突を防いでいる。敷田らは、複数人のいる拠点同士を接続して行う遠隔会議において、拠点内での議論に他拠点から割り込みづら問題に着目し、発話要求や発話終了のジェスチャを検出して話者交代に対する気づきを与える手法を提案している [9,11]。山田らは、遠隔会議において発話直前の予備動作に着目し、遠隔会議映像を分析して発話欲求度を推定して提示することで発話衝突を防ぐ手法を提案している [6]。

このような話者交代に関する技術に対し、筆者らは「話

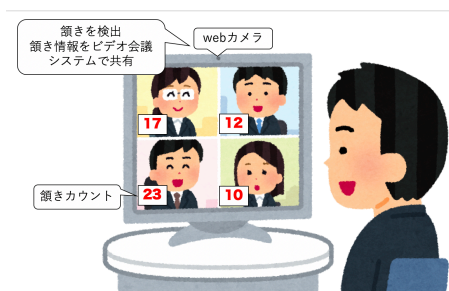


図 1 頷きのリアルタイムフィードバックシステム概要

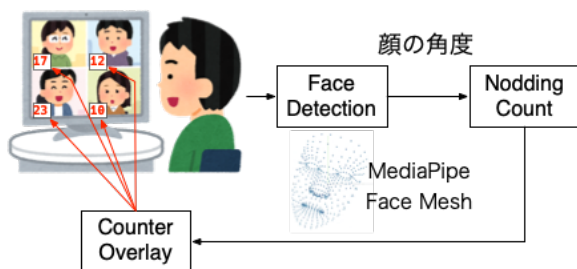


図 2 リアルタイム頷き認識の概要

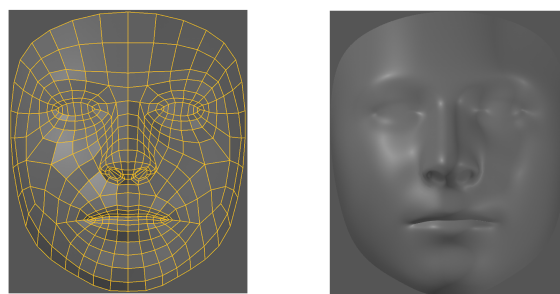
者の話しづらさ」に着目した。プレゼンテーション画面を共有しながらの遠隔会議などにおいて、話者は全参加者の顔を確認しながら参加することは難しく、聞き手の情報を一部しか確認できない。このため、自分の話に対して相手がどのような反応を示しながら聞いているかを確認することができないという問題がある。

筆者らはこれまで、会議の「質」を評価するシステムの実現に向けて会議映像から取得した情報を分析する手法を提案してきた [12,13]。しかしながら、これらの手法はリアルタイムな分析を行うことができず、遠隔会議を行っている間の支援技術として不十分である。

3. 頷きのリアルタイムフィードバックシステムの提案

図 1 に頷きのリアルタイムフィードバックシステムの概要を示す。ビデオ会議中に参加者それぞれのパソコンにおいてカメラ映像からリアルタイムに頷きを認識し、その頷き情報をビデオ会議システムで共有する。ビデオ会議システムでは他者から共有された頷き情報を「頷きカウント」としてビデオ会議システム上に表示する。発話者は「頷きカウント」を見ながら発話することで、会議参加者の反応を間接的に得ることが可能となる。

この実現に向けてはリアルタイムな頷きの認識が必須である。図 2 にリアルタイム頷き認識の概要を示す。リアルタイム頷き認識は、顔検出ブロック、頷きカウントブロック、頷きカウンタオーバーレイ表示ブロックの 4 つのブロックで構成される。まず、顔検出ブロックにおいてビデオ会議映像からリアルタイムに顔画像を検出する。次に、



(a) (b)

図 3 Face Mesh における (a) ランドマークによるメッシュドポロジ及び、それに基づく (b) 3次元細分割曲面 (文献 [15] より抜粋)

頷きカウントブロックで顔画像の変化から頷きを検出し、その回数をカウントする。最後に、カウント結果を頷きカウンタオーバーレイ表示ブロックを用いてビデオ会議映像にリアルタイムに重畳表示する。

本研究では Web ブラウザで参加可能なビデオ会議システムの利用を想定し、リアルタイム頷き認識をブラウザの拡張システムとして実現する。各会議参加者には拡張システムをインストールした Web ブラウザを用いてビデオ会議に参加してもらい、会議参加者自身の Web カメラ映像を自身の Web ブラウザ上で解析して頷きを検出・カウントする。そして、カウント結果を共有することで図 1 に示した頷きのリアルタイムフィードバックシステムを実現する。

以降では、各ブロックの詳細について述べる。

3.1 顔検出ブロック

顔検出ブロックでは、Web ブラウザ上で動作する軽量な顔検出フレームワークである MediaPipe Face Mesh^{*3}を用いてビデオ会議映像から顔を検出し、顔の角度を算出する。

Face Mesh は顔検出手法である BlazeFace [16] を用いて 2次元映像から顔を検出し、検出した顔について 468 個の顔ランドマーク (顔の特徴点) を 3次元でトラッキングするフレームワークである [15]。図 3 は、文献 [15] で定義されている顔ランドマークのメッシュドポロジを示している。

Face Mesh で推定した顔のランドマーク点群を取得し、顔の輪郭を取り出して顔の輪郭点群の座標から顔の向きを表す回転行列を算出する。まず、平均的な目の間隔が 6 cm 程度であることを利用して顔ランドマーク点群の座標を正規化する。顔点群座標の集合を \mathbf{F} 、左右の目の外側の座標点を eye_{left} , eye_{right} とすると、スケールを正規化した顔点群座標の集合 $\bar{\mathbf{F}}$ の要素 $\bar{f} \in \bar{\mathbf{F}}$ は、 $f \in \mathbf{F}$ を用いて以下のように計算できる。

$$\bar{f} = \frac{f}{|eye_{left} - eye_{right}|} \times 0.06 \quad (1)$$

*3 https://google.github.io/mediapipe/solutions/face_mesh.html

次に、正規化した顔座標点群の中心位置を求め、顔の中心が座標 (0, 0, 0) となるように移動した正規化顔座標点群 $\tilde{f} \in \tilde{\mathbf{F}}$ を計算する。

$$\tilde{f} = \bar{f} - \frac{1}{|\tilde{\mathbf{F}}|} \sum_{f \in \tilde{\mathbf{F}}} f \quad (2)$$

このように正規化された顔点群の中から、同一平面上の4点を取り出して直交する2つのベクトルを作成する。ここでは簡単のため、顔の上下左右の輪郭点など x 座標が同じ2点 X_1, X_2 , y 座標が同じ2点 X_3, X_4 をそれぞれ取り出して直交する2つのベクトル $\overrightarrow{X_1X_2}, \overrightarrow{X_3X_4}$ を作成する。回転行列は作成した2つの直交ベクトルを正規化したベクトルと、その外積を計算して作成した3つの正規直交ベクトルとを用い、3つの正規直交ベクトルが作る正規直交基底への回転として算出できる。具体的には、 $X_n = (x_n \ y_n \ z_n)^T$, 回転行列 R の第 n 行を R_n とすると、 R は3行3列の正方行列であり、

$$R_1 = \frac{1}{|X_1 - X_2|} \begin{pmatrix} 0 & y_1 - y_2 & z_1 - z_2 \end{pmatrix} \quad (3)$$

$$R_2 = \frac{1}{|X_3 - X_4|} \begin{pmatrix} x_3 - x_4 & 0 & z_3 - z_4 \end{pmatrix} \quad (4)$$

$$R_3 = R_1 \times R_2 \quad (5)$$

と求められる。ここで、 $()^T$ はベクトルの転置を、 \times はベクトルの外積演算を表している。

最後に、算出された回転行列から顔の角度を求める。回転行列の i 行 j 列の要素を r_{ij} , とすると、 x, y, z 軸に対する角度 $\theta_x, \theta_y, \theta_z$ は以下のように計算できる。

$$\theta_x = \arctan \frac{r_{32}}{r_{33}} \quad (6)$$

$$\theta_y = \arctan \frac{-r_{31}}{\sqrt{r_{11}^2 + r_{21}^2}} \quad (7)$$

$$\theta_z = \arctan \frac{r_{21}}{r_{11}} \quad (8)$$

3.2 顔きカウントブロック

顔きカウントブロックでは、顔検出ブロックで算出した顔の角度 $\theta_x, \theta_y, \theta_z$ の時間変化に基づいて顔きを検出する。まず、 $\theta_x, \theta_y, \theta_z$ の時系列データを固定幅のウィンドウで区切り、そこから以下の特徴量を抽出する。

- 角度 $\theta_x, \theta_y, \theta_z$ の平均
- 角度 $\theta_x, \theta_y, \theta_z$ の分散

顔きとは角度の変化の仕方から検出できることから、顔きしていない状態の基準を平均値によって抽出するとともに、分散によって変化の有無を抽出する。

次に、抽出した特徴量を用いて顔きの有無を教師あり学習により推定する。本研究では顔きの検出に用いるアルゴリズムは限定しない。二値分類が可能なアルゴリズムであれば顔きカウントブロックで利用できる。Webブラウザ上で動作させることを考慮して軽量の推定アルゴリズムを選

定する。

最後に、検出された顔きに基づいて顔きの回数をカウントする。

顔き検出に用いる教師あり学習のモデルは、事前に収集したデータで学習して構築する。本研究では、Webブラウザに導入した拡張システムを初めて使う際に連続して顔きしてもらいなどしてキャリブレーションを行い、使用者の環境で学習に必要なデータを取得することを想定している。

3.3 顔きカウンタオーバーレイ表示ブロック

顔きカウンタオーバーレイ表示ブロックでは、顔きカウントブロックでカウントしたカウント値を自身のWebカメラ映像に画像として重畳する。Webブラウザを用いるビデオ会議システムではWebカメラの映像もブラウザ拡張システムで取得できることから、このカメラ映像にカウント値を重畳した映像を作成し、ビデオ会議システムに対してカメラ映像として入力する。

本システムでは顔きの結果をWebカメラ映像として渡していることから、原理的にはWebカメラ映像ではなくカウント値だけをWebカメラ映像として渡すことも可能である。このため、カメラOFFのまま参加者の顔き情報だけを伝送することも可能である。

現在の実装ではカウント値をテキストで表示しているが、顔きの履歴や顔き以外の情報の表示など、様々な拡張を考えている。将来的には、ビデオ会議をサポートする情報の検討、フィードバック方法の検討を行い、図1に示した提案システムを拡張していくことを考えている。

4. 実験

本稿では、実装した顔きのリアルタイムフィードバックシステムを利用した場合と利用していない時にビデオ会議中のコミュニケーションがどのように変化したのかをアンケートを通じて定性的に評価を行う。なお今回の実験では初期的実験として、顔きの判定モデルは単純化し、顔の角度の時間変化が一定のしきい値を超えた場合に顔きとして認識する「しきい値方式」により実装した。アンケートの内容を表1に示す。

表1 顔きカウンタ使用後のアンケート

設問	回答形式	質問
1	自由記述	顔きカウンタによって、何かこれまでに気づいていなかった仕草に気づきましたか？
2	自由記述	顔きカウンタによって、自分自身の発話や仕草に影響を与えましたか？
3	自由記述	顔きカウンタによって、参加者全体のコミュニケーションに変化がありましたか？
4	4択	顔きカウンタはあった方がよい？

設問4の選択肢は、良い、どちらかといえば良い、ど

らかといえば無くても良い、無いほうがいい、の4つとした。ビデオ会議は4人で接続し、自分自身を含め4人の様子が見えるグリッドビュー状態で行う。ただし、互いの姿は、直接的には、見えないように各被験者を配置した。

4人の被験者に内容は特に指定せずに5分間の雑談ビデオ会議システムを通じて行ってもらう。その後、顔のリアルタイムフィードバックシステムを利用し、さらに5分間の雑談を行い、アンケートを通じて定性的に評価を行う。顔のリアルタイムフィードバックシステム利用中のビデオ会議中の画面の様子を図4に示す。この際、被験者には利用するシステムについての説明は行わず、表示される数字が顔のカウンタであることは伝えずに実験を行う。今回の被験者は、九州大学システム情報科学府の20代学生4名である。なお、本実験は、九州大学倫理審査委員会の承認（承認番号：シス情認2020-06）を受け、実施している。

5. アンケート結果と考察

顔カウンタ使用後のアンケートの結果とそれぞれに対する考察を以下にまとめる。

設問1：顔カウンタによって、何かこれまでに気づいていない仕草に気づきましたか？

設問1は、顔というこれまで意識していなかった動作が数値として表示されることによって、新しい気づきを得られるのではないかとという仮説に基づいている。ただし、被験者4人共、この1年間、コロナ禍により、オンラインミーティングを高頻度に行っているため、すでに自身の仕草を意識している可能性もあった。

アンケート結果は、4名中3名が「これまでに気づいてなかった仕草に気付かされた」と回答し、もう1名は「自分で思っているより無表情でいる時間が長いことに気づいた」という回答であった。この結果からわかることは、顔カウンタを使用した場合には、使用していない場合よりも会話中の仕草や表情を強く意識するようになったということである。意外な結果としては、無表情であるという、顔の仕草について言及があった点である。このことから考えられるのは、顔カウンタを表示することによって、ミーティング参加者の意識がより画面に集中している、特に、数字が表示される自身の顔に対して意識的に確認をしているということである。

設問2：顔カウンタによって、自分自身の発言や仕草に影響を与えましたか？

設問2は、フィードバックを受けた後、自分自身の仕草に変化が出るのではないかとという仮説に基づいている。これは、以前、実際のディスカッション中に、無意識に鼻を

触っていることを指摘したら、その後、その人が意識的に触らなくなったという実体験に基づいている。

結果は、ニュアンスを伝えるため、記述された回答文をそのまま掲載する。

被験者1 少し意識した

被験者2 とてもそう思う

被験者3 実験中には特に気づかなかったが、実験後の数値をみて、もう少し他人の会話にリアクションをとるべきだなと思った

被験者4 はい

被験者1, 4, 2の順で、自分自身の表情や仕草に対する意識が強くなったと感じていることがわかる。これは、鏡の前で身だしなみを整えるのと同じように、顔カウンタを表示することにより、カウンタを表示しない通常のビデオ会議時よりも自身が他人に見られているという意識が強くなったのではないかと考えられる。

また、被験者3については、自身のリアクションの少なさを再認識している。これは、会議中は自分自身の表示に集中していたために気にならなかったが、数値化されることで、会議終了後に他者と比較できるようになり、自分自身の無反応さに気づいたのではないと思われる。このことから、個人の数値を表示するだけでなく、他の参加者の数値も表示することが、このような他者との違いを気づかせるきっかけになるという知見が得られた。

設問3：顔カウンタによって、参加者全体のコミュニケーションに変化がありましたか？

設問3は、通常のビデオ会議に、顔カウンタという新しい要素が入ったことで、会話が弾んだり、声が大きくなったり、笑が増えたり、といったコミュニケーション全体の雰囲気を与える影響があるのではないかと、という仮説に基づいている。

こちらの結果についても、ニュアンスを伝えるため、記述された回答文をそのまま掲載する。

被験者1 カウンタの話になることがあった

被験者2 とてもそう思う

被験者3 コミュニケーションの変化は特に感じなかったが、このシステムに慣れてきたら、変わってくるのかもしれない

被験者4 はい

結果としては、被験者2や4は、顔カウンタによって、全体のコミュニケーションに変化があったと報告している。被験者1の回答にもあるが、顔カウンタの仕組みや性能（首を横に振ってもカウントされないかなど）に興味を持ち、機能を試しながら話すという場面も見られ、話題の1つとなっていた。しかしながら、被験者3が指摘しているように、数字の表示が当たり前になってくると、

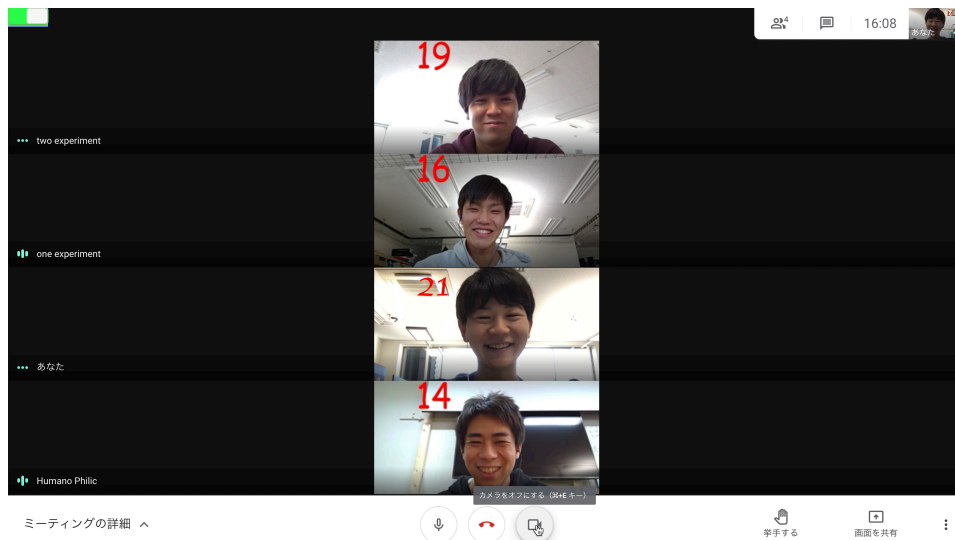


図 4 領きカウンタ利用中のビデオ会議画面

その数字の有無によって、コミュニケーションに変化が生まれなくなる可能性も十分に高いと感じた。

リアルタイムフィードバックという観点では、会議を邪魔しない範囲で、いかに表記のバリエーションを増やしたり、可視化する仕草の種類を増やすかが重要であると感じた。また、設問2の回答と合わせて考えると、リアルタイムフィードバックではなく、会議終了まで何も提示せず、会議終了後にレポートとして表示する形式も検討の余地があると考えた。

設問 4：領きカウンタはあった方がいい？

設問4は、本論文で提案、開発した領きカウンターが有用であったか否かを問うものである。アンケート結果としては、1名の被験者が「あったほうが良い」、2名の被験者が「どちらかといえばあったほうが良い」と回答し、残り1名の回答者が「どちらかといえば無くても良い」という回答であった。

より多くの被験者から、意見を募る必要があるが、領き回数の表示程度であれば、特にミーティングを邪魔するものでないため、不要であるという強い意見が出なかったのではないかと考えられる。

今回は、初めて利用して、今まで意識していなかった自身の仕草、他人の仕草というものが可視化されたことに対して、興味、関心が出たと考えられるが、何度か利用することで、気づきが減り、機能は不要になるという可能性も十分に有り得ると考えている。我々のシステムは、ブラウザの拡張機能として実装しており、拡張機能をインストールしていない人には数字は表示されない。また、トグルスイッチで On/Off できることから、このようなリアルタイムフィードバックを表示したくない人は、Off にするという形の利用が想定される。

6. おわりに

本稿では、領きのリアルタイムフィードバックによるミーティング支援を目的とし、MediaPipe Facemesh を用いて顔の角度を推定し領きの認識を行い、Chrome の拡張機能によるリアルタイムフィードバック手法を提案した。

提案システムを用いた実験を通じ、領きカウンターによって、自分や他人の仕草に意識が行き、これまで気づかなかった自分の仕草に気づくことを明らかにした。また、定量化された数字で表示されることで他人との比較が容易になり、そこから気づきが得られることもわかった。また、機能の受容性は高く、コミュニケーションを盛り上げる、話題の1つとしても導入できると感じられた。

今後は領きカウンタを適宜リセットする機能の追加や、数値ではなくグラフとして可視化するなどの拡張を行い実験をすることで追加検証する必要がある。コミュニケーションの変化に関しては、初期実験として雑談の内容は制限せず、また表示される数字が領きのカウントであることは伝えずに行われたことによりカウンタの話題になることがあった。今後は領きカウンタであることを伝えた場合との比較や、領きカウンタ使用時と非使用時の会議中の声量を比較するなどしてを定量的に評価する必要がある。領きカウンタの有用性についても追加で実験を行いデータを増やすことで検証する必要がある。

謝辞 本研究は、大阪大学が Society 5.0 実現化研究拠点支援事業 (JPMXP0518071489) のもとと実施するグランドチャレンジ研究、および、JST CREST 「経験サプリメントによる行動変容と創造的協働」(JPMJCR16E1) の支援のもと実施されている。

参考文献

- [1] Adalgeirsson, S. O. and Breazeal, C.: MeBot: A Robotic Platform for Socially Embodied Presence, *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction, HRI '10*, Osaka, Japan, IEEE Press, pp. 15–22 (2010).
- [2] Kawanobe, H., Aosaki, Y., Kuzuoka, H. and Suzuki, Y.: iRIS: A Remote Surrogate for Mutual Reference, *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction, HRI '13*, Tokyo, Japan, IEEE Press, pp. 403–404 (2013).
- [3] Otsuki, M., Kawano, T., Maruyama, K., Kuzuoka, H. and Suzuki, Y.: Representing Gaze Direction in Video Communication Using Eye-Shaped Display, *Proceedings of the 29th Annual Symposium on User Interface Software and Technology, UIST '16 Adjunct*, New York, NY, USA, Association for Computing Machinery, pp. 65–67 (2016).
- [4] 大槻麻衣, 丸山啓太, 葛岡英明, 鈴木雄介: ThirdEye: 遠隔対話システムにおける対話者の視線を提示する眼球ディスプレイの開発と評価, *ヒューマンインタフェース学会論文誌*, Vol. 19, No. 2, pp. 211–218 (2017).
- [5] 玉木秀和, 東野 豪, 小林 稔, 井原雅行: 発話がぶつからない Web 会議を実現するための発話欲求伝達手法, *情報処理学会論文誌*, Vol. 54, No. 1, pp. 275–283 (2013).
- [6] 山田楓也, 白石 陽, 石田繁巳: Web 会議における予備動作を用いた発話欲求推定手法の提案, *情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO)* (2021). (will appear).
- [7] Ohno, T.: Weak Gaze Awareness in Video-Mediated Communication, *CHI '05 Extended Abstracts on Human Factors in Computing Systems, CHI EA '05*, New York, NY, USA, Association for Computing Machinery, pp. 1709–1712 (2005).
- [8] 小峯隆宏, 勝本道哲, 丹 康雄: 遠隔会議でのアイコンタクト実現手法の提案と評価, *情報処理学会研究報告マルチメディア通信と分散処理 (DPS)*, Vol. 2005, No. 33(2004-DPS-122), pp. 139–144 (2005).
- [9] 敷田幹文, アルニーラティカン: 人数が不均衡な遠隔テレビ会議における弱い光を用いた視線アウェアネス, *情報処理学会論文誌*, Vol. 58, No. 1, pp. 166–175 (2017).
- [10] Tausif, M. T., Weaver, R. and Lee, S. W.: Towards Enabling Eye Contact and Perspective Control in Video Conference, *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology, UIST '20 Adjunct*, New York, NY, USA, Association for Computing Machinery, pp. 96–98 (2020).
- [11] 敷田幹文, 増田雄亮: 分散環境における話者交替のアウェアネス支援, *情報処理学会論文誌*, Vol. 56, No. 1, pp. 126–136 (2015).
- [12] 曾根田悠介, 中村優吾, 松田裕貴, 荒川 豊, 安本慶一: ミーティング映像からの発話およびマイクロ動作識別手法, *情報処理学会研究報告ユビキタスコンピューティングシステム (UBI)*, Vol. 2020-UBI-65, pp. 1–8 (2020).
- [13] 徳原耕亮, Dawton, B., 荒川 豊, 石田繁巳, 曾根田悠介, 松田裕貴: グループミーティング動画からの発話量抽出手法の検討, *情報処理学会全国大会講演論文集*, pp. 3:311–3:312 (2020). 5X-07.
- [14] Vertegaal, R.: The GAZE Groupware System: Mediating Joint Attention in Multiparty Communication and Collaboration, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '99*, New York, NY, USA, Association for Computing Machinery, pp. 294–301 (1999).
- [15] Kartynnik, Y., Ablavatski, A., Grishchenko, I. and Grundmann, M.: Real-Time Facial Surface Geometry from Monocular Video on Mobile GPUs, *arXiv:1907.06724 [cs]* (2019). Comment: 4 pages, 4 figures; CVPR Workshop on Computer Vision for Augmented and Virtual Reality, Long Beach, CA, USA, 2019.
- [16] Bazarevsky, V., Kartynnik, Y., Vakunov, A., Raveendran, K. and Grundmann, M.: BlazeFace: Sub-Millisecond Neural Face Detection on Mobile GPUs, *CVPR Workshop on Computer Vision for Augmented and Virtual Reality*, Long Beach, CA, pp. 1–4 (2019).