

3人1組でのビデオ通話における対話相手への直視および視線回避の強調表示による参加者の状況伝達手法の実装

芦沢 優介^{1,a)} 川口 一画²

概要: 複数人で会話を行うときに重要な役割を果たす視線から得られる非言語情報として、相互注視 (mutual gaze) および視線回避 (gaze aversion) がある。ビデオ通話では、対面形式での会話とは異なり、対話相手の映るビデオウィンドウのみから、対話相手の視線状況、すなわち対話相手の直視および視線回避を理解することは困難である。この課題に対して、先行研究として、著者らは2人1組でのビデオ通話において対話相手の直視および視線回避を強調表示する手法を提案し、その有効性を調査した。一方、参与役割がより分かりにくい3人以上でのビデオ通話に先行研究の提案手法を適用すると、話者交替の円滑さ、対話相手の社会的存在感、およびビデオ通話の質などの向上に大きく貢献できると考えられる。そのため、本研究では、先行研究の提案手法を3人1組用に拡張した。拡張したシステムを用いて、3人1組の状況で提案手法の有効性および改善点を調査するための試行的評価を実施した。

Implementation of a Method for Communicating Participant's Status by Highlighting Direct Gaze and Gaze Aversion with Interlocutors in Video Calls with Groups of Three

1. はじめに

複数人で会話を行うときに重要な役割を果たす非言語情報として、視線があげられる [1]。この視線から得られる非言語情報として、会話において、お互いの視線を合わせることである相互注視 (mutual gaze) および視線を対話相手からそらすことである視線回避 (gaze aversion) がある。相互注視は発話権の譲渡や発言の終了を対話相手に伝える役割を果たす [2]。また、視線回避は発言内容を考えるため [2], [3], [4], [5]、会話全体の親密さを調整するため [2], [6]、および会話の場を保持するため [2] に行われる。

しかし、ビデオ通話では、対面形式での会話とは異なり、対話相手の顔映像を参加者側に表示するビデオウィンドウ (以降、参加者側のビデオウィンドウ) のみから、対話相手の視線状況を理解することは困難である。そのため、話者交替の円滑さ、対話相手の社会的存在感、およびビデオ通話の質が低下する [7], [8]。

この課題を解決するために、先行研究 [9] において、著

者らは2人1組でのビデオ通話において対話相手の直視および視線回避を強調表示する手法 (以降、先行手法) を提案し、その有効性を調査した。先行手法では、この非言語情報をビデオウィンドウの大きさの変更またはビデオウィンドウの枠の色の変更により伝達した。例として、ビデオウィンドウの大きさを変更する場合の先行手法の概要を図1に示す。ビデオウィンドウの大きさを変更する場合の先行手法は、ビデオ通話において対話相手が画面を直視しているときに、参加者側のビデオウィンドウを大きくする。一方、ビデオウィンドウの枠の色を変更する場合の先行手法では、先述した手法における参加者側のビデオウィンドウの大きさの変化が参加者側のビデオウィンドウの枠の色の不透明度の変化に置き換わる。先行研究 [9] において、対話相手の直視および視線回避の強調表示は対話相手の視線の変化の気づきやすさおよび話者交替の円滑さの向上に対して効果的であることが示された。一方で、対話相手の社会的存在感およびビデオ通話の質に関しては、通常のビデオ通話と比較して、改善は見られなかった。

先行研究 [9] において着目した2人1組でのビデオ通話では、参加者は発話を行う話し手または発話を受ける受け

¹ 筑波大学 情報理工学位プログラム

² 筑波大学 システム情報系

^{a)} ashizawa@iplab.cs.tsukuba.ac.jp

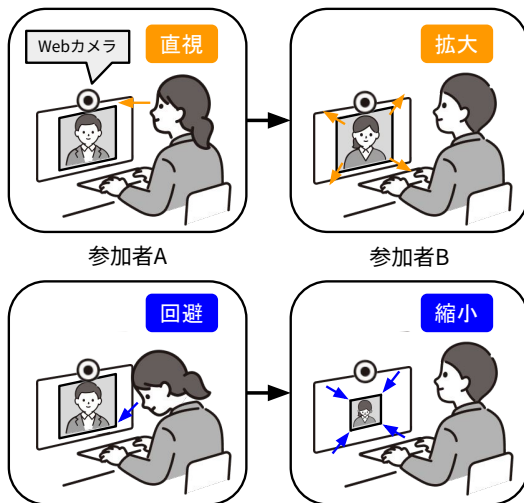


図 1 ビデオウィンドウの大きさを変更する場合の先行手法の概要 [9]

Fig. 1 Overview of conventional method when changing the size of the video window [9]

手のどちらかであるため、会話中の各参加者の役割（以降、参与役割）が比較的分かりやすい [10]。一方、3人以上でのビデオ通話では、話し手および受け手だけではなく、会話に参加しているが発話を受けていない傍参与者および会話を聞いているだけの傍観者も存在するため、参与役割が分かりにくい [10], [11]。そのため、先行手法を拡張し3人以上のビデオ通話において用いることにより、2人1組で用いる場合よりも、話者交替の円滑化、対話相手の社会的存在感の向上、およびビデオ通話の質の向上などに大きく貢献できると考えられる。そのため、本研究では、先行手法を3人1組でのビデオ通話に拡張する。

2. 関連研究

本節では、まず、本研究において着目する非言語情報である対話相手の直視および視線回避に関連する非言語情報として、相互注視および視線回避が会話に与える影響を示す。また、本研究および関連研究の差分を示すために、ビデオ通話の参加者の視線方向を伝達する研究および著者らの先行研究を示す。

2.1 相互注視および視線回避が会話に与える影響

相互注視および視線回避は会話において重要な役割を果たす。相互注視とは、会話において、お互いの視線を合わせることである。会話においてお互いの視線を合わせることである相互注視は、対話相手との物理的な距離、話題の親密さ、および微笑みの量と同じく、対話相手との距離感に影響を与える [12]。また、相互注視は自分自身の発言を終了するときに行われる [2]。一方、視線を対話相手からそらすことである視線回避は会話における3つの主要な機能である認知的機能、親密さの調整機能、および場の管理

機能を果たすために用いられる [13]。また、視線回避は発言内容の熟考および伝達への集中 [2], [3], [4], [5] をするとき、会話全体の親密さの調整 [2], [6] をするとき、ならびに発言の意思表示 [2] をするときに行われる。

本研究では、ビデオ通話において、対話相手の直視および視線回避をビデオウィンドウの大きさまたはビデオウィンドウの枠の色に反映する。これにより、ビデオ通話における理解が困難である、対話相手の視線状況を伝達する。

2.2 ビデオ通話の参加者の視線方向の伝達

ビデオ通話の参加者の視線方向を伝達するための研究が複数ある。Kumar ら [14] は、各参加者の顔映像が表示されているビデオウィンドウ（以降、参加者のビデオウィンドウ）の右上にその参加者を直視している参加者を表す円形の泡を表示することにより、各参加者の視線方向を伝達する手法を示した。この研究において、各参加者のビデオ通話への関心度が高くなることが示された。Iitsuka ら [15] は、各参加者のビデオウィンドウからその参加者が直視している参加者のビデオウィンドウへの矢印の表示または注目している人数の多い参加者のビデオウィンドウの拡大により、各参加者の視線方向を伝達する手法を示した。この研究において、参加者間の話者交替の円滑化および発話数の増加が示された。

本研究では、これらの研究のような、対話相手が「どこを見ているか」の伝達ではなく、対話相手の直視および視線回避の度合いの伝達を行う。また、その限定的な情報伝達がビデオ通話に与える影響を明らかにすることを目的とする。

2.3 著者らの先行研究

先行研究 [9] では、著者らは2人1組でのビデオ通話において対話相手の直視および視線回避を強調表示する手法を提案した。具体的には、対話相手の直視および視線回避の度合いという限定的な情報を基に、ビデオウィンドウの大きさを変更する手法およびビデオウィンドウの枠の色を変更する手法を実装した。また、これらの手法の評価実験を行い、対話相手の直視および視線回避の強調表示が対話相手の視線の変化の気づきやすさおよび話者交替の円滑さの向上に寄与することを示した。一方、この強調表示が対話相手の社会的存在感およびビデオ通話の質の向上に寄与することは示されなかった。

本研究では、先行研究 [9] において提案した、2人1組でのビデオ通話における対話相手の直視および視線回避の強調表示手法を3人1組でのビデオ通話に拡張する。

3. 実装

本研究では、2人1組でのビデオ通話に着目した先行手法を3人1組でのビデオ通話に拡張した。本節では、提案

手法であるビデオウィンドウの枠の色を変更する場合の手法（以降、FrameChange 手法）およびビデオウィンドウの大きさを変更する場合の手法（以降、SizeChange 手法）を導入したビデオ通話システム（以降、本システム）の構成、参加者側のビデオウィンドウの枠の色または大きさの変更方法、および先行手法の実装内容との変更点を示す。

3.1 システム構成

本研究では、本システムを React, TypeScript, HTML, および CSS から構成される Web アプリケーションとして実装した。このアプリケーションは映像、音声、および対話相手側のラップトップ PC において推定した参加者側のビデオウィンドウの大きさを表すデータ（以降、ウィンドウサイズに変換されたデータ）を送受信する機能を持つ。本システムでは、映像および音声の取得、映像、音声、およびウィンドウサイズに変換されたデータの送受信、ならびに映像および音声の表示を行うために、リアルタイムコミュニケーションを実現するマルチプラットフォーム SDK である SkyWay [16] の JavaScript SDK を用いた。また、頭部方向を推定するために、オープンソースのフレームワークである MediaPipe [17] を用いた。MediaPipe を用いることにより、Web カメラから取得した対話相手の顔映像から、顔の特徴点の座標を顔映像上の座標として取得できる。本システムでは、このように取得した顔の特徴点を用いることにより、Web カメラのある方向および頭部方向間の角度を推定する。

3.2 参加者側のビデオウィンドウの枠の色または大きさの変更方法

本研究では、参加者側のビデオウィンドウの枠の色または大きさを変更するために、飯塚 [11] の実装を参考にした。具体的には、まず、MediaPipe の Face Mesh [18] からリアルタイムに取得した顔上の 478 点 $f_g(n)$ ($1 \leq n \leq 478$) の座標を用いて、顔の中心点 g_c の座標を計算する。 g_c の座標の計算式は以下のとおりである。

$$g_c = \left(\sum_{1 \leq k \leq 478} f_g(k) \right) / 478$$

その後、顔の中心点 g_c から、取得した 478 点のうち、正面を向いたときの顔の中心点である鼻の先の点 g_n へのベクトル（以降、顔の中心点から鼻の先の点へのベクトル）の長さを計算する。このベクトルが Web カメラのある方向および頭部方向間の角度に対応する。また、顔の中心点から鼻の先の点へのベクトルの計算後、参加者側のビデオウィンドウの大きさまたは枠の色の不透明度をこのベクトルの長さに反比例するように変更する。これにより、SizeChange 手法では、対話相手側の Web カメラのある方向および対話相手の頭部方向間の角度が小さいときほど、参加者側のビデオウィンドウが大きくなる。このとき、ビ

デオウィンドウの大きさまたは枠の色の急激な変化を防ぐために、参加者側のビデオウィンドウの大きさまたは枠の色の不透明度の変更に対して、本研究では 10 点移動平均を用いている。1 点あたりの処理時間は SizeChange 手法および FrameChange 手法ともに平均 3.5 ミリ秒である。

なお、FrameChange 手法では、Zoom [19] を用いたビデオ通話において発話者の映るビデオウィンドウの枠の色が緑色に変化することを基に、ビデオウィンドウの枠の色の RGB 値を $(r, g, b) = (83, 253, 49)$ に固定している。また、ビデオウィンドウの枠の太さを 10 px、枠の色の不透明度の最小値を 0.0、およびこの最大値を 1.0 としている。

3.3 先行手法の実装内容との変更点

3.1 節および 3.2 節に示した実装内容は先行手法においても適用されている。そのため、提案手法の実装内容は基本的に先行手法の実装内容と同じである。

先行手法の実装内容との変更点は以下のとおりである。FrameChange 手法における緑色の枠の変化を強調するために、提案手法では、図 2 に示すように、各参加者の映るビデオウィンドウを黒色背景の前面に配置する。具体的には、自分自身の顔映像を画面左上に、他者の顔映像を画面右上および画面下部中央の計 3 箇所配置する。画面下部中央のウィンドウは、画面を左右に 2 分割かつ上下に 4 分割したときに得られる交点のうち、最も下の点を中心とする。これら 3 つのウィンドウの枠の色または大きさは、各領域に収まる範囲内で変化する。なお、自分自身の顔映像を表示する理由は、先行研究 [9] において行った評価実験（以降、先行実験）において、「普段のビデオ通話と同じく、自分の顔も表示してほしい」という意見が得られたためである。

4. 試行的評価

本研究では、提案手法の試行的評価（以降、パイロットスタディ）も行った。本節では、パイロットスタディの概要および結果を示す。

4.1 概要

パイロットスタディの実験条件は先行実験と同じく以下の 3 つである。

- C1. ビデオウィンドウの大きさおよび枠の色を固定する条件（以降、Baseline 条件）
- C2. ビデオウィンドウの枠の色が対話相手の頭部方向に合わせて変化する条件（以降、FrameChange 条件）
- C3. ビデオウィンドウの大きさが対話相手の頭部方向に合わせて変化する条件（以降、SizeChange 条件）

提案手法のうちの 1 つである SizeChange 手法では、各参加者が対話相手の顔映像をつねに認識できるようにするために、各参加者の映るビデオウィンドウの大きさの最小

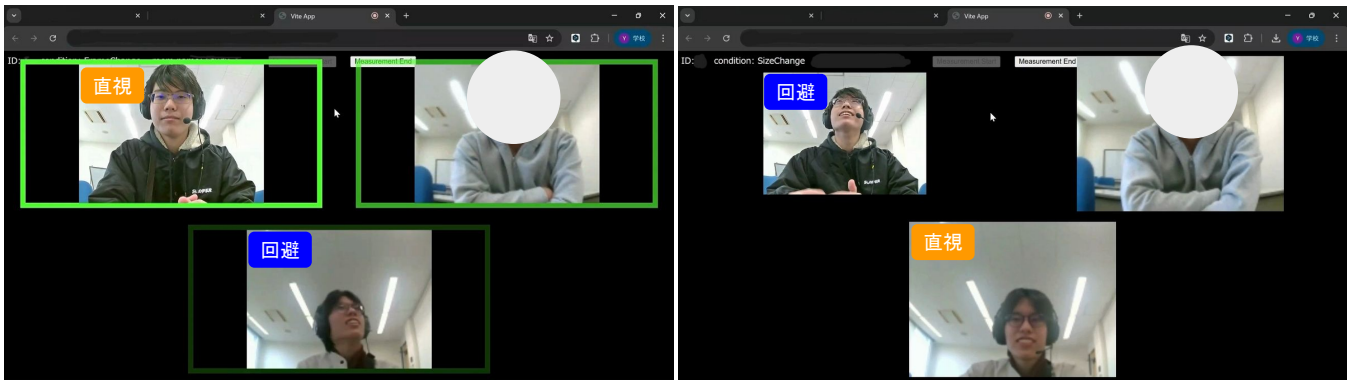


図 2 提案手法を用いたときのラップトップ PC の画面. 左側: FrameChange 手法を用いたときの画面. 右側: SizeChange 手法を用いたときの画面. 個人情報保護の観点より, 一部参加者の顔画像には匿名化処理を施している.

Fig. 2 Displays of laptop PCs when using proposed method. Left: When using FrameChange method. Right: When using SizeChange method. From a personal information protection standpoint, anonymization processing has been applied to the facial images of some participants.

値を幅 414 px および高さ 233 px にしている (図 2 右側). また, 参加者側のビデオウィンドウ全体が画面に収まるようにするために, その最大値を幅 526 px および高さ 296 px にしている (図 2 右側). そのため, 3 つの実験条件のうち, Baseline 条件および FrameChange 条件における各参加者の映るビデオウィンドウの大きさを幅 470 px および高さ 264.4 px にした (図 2 左側). 想定するディスプレイの解像度は 1920 × 1080 である. また, 提案手法のうちの 1 つである FrameChange 手法では, 各参加者の映るビデオウィンドウの枠が実際に表示されているビデオウィンドウよりも外側に表示されている.

パイロットスタディの参加者はボランティアとして参加した研究室内の大学生・大学院生 1 組 3 人 (以降, P13–P15) であった. 彼らの構成は男性 3 人, 平均年齢 22.7 歳 ($SD = 1.15$) であった. なお, P13–P15 のうち, P13 および P15 は本実験にも参加していた.

パイロットスタディにおいて用いたラップトップ PC は Gen 11 (Intel (R) Core (TM) i7–1195G7, Intel (R) Iris (R) Xe Graphics, 32 GB, 14 型 16:9, 1920 × 1080) である. パイロットスタディでは, 外付けの小型 Web カメラ [20] の配置されたラップトップ PC を用いて, 互いの音声がかえらないように別々の部屋に案内した各参加者に実験タスクを行ってもらった. このとき, 各参加者の音声の個別取得によりデータ分析を容易にできるようにするために, 各参加者にヘッドホンを装着してもらった.

パイロットスタディの実験タスクは, 先行実験と同じく, アイデアを出す会議である創造会議である. パイロットスタディでは, お題の難易度をできる限り同じにするために, 会議の題材を靴下, ヘルメット, および箸とした. このとき, 関連研究 [21] および先行実験において出された「鍋

というお題が他のお題よりも難しかった」という 2 人の参加者からの意見を参考にした.

パイロットスタディの評価指標は以下の 11 項目である.

- 会話分析から得られる指標
 - 会議中に視線をそらした状態から対話相手を直視した状態に切り替えた回数
 - 会議中に相互注視していた時間の割合
 - 会議中に対話相手を直視していた時間の割合
 - 会議中の発話数
 - 会議中の沈黙時間の割合
 - 会議中に出されたアイデアの数
- アンケートから得られる指標
 - 会議中における対話相手の視線の変化の気づきやすさを測定するためのアンケート
 - 会議中における話者交替の円滑さを測定するためのアンケート
 - 会議中の社会的存在感を測定するためのアンケート
 - 会議に対する印象を測定するためのアンケート
 - 各条件の好みの順位を測定するためのアンケート

パイロットスタディでは, 先行実験において用いた 10 項目に加えて, 各参加者のヘッドホン装着により取得しやすくなった会議中の発話数を会話分析から得られる指標として追加した. 話者交替が円滑になると, 自分自身の発話するべきタイミングが明確になるため, 会議中の各参加者の発話が増加すると考えた. なお, データの分析時には, パイロットスタディ中に撮影したビデオ映像から会議中の発話数を集計した.

4.2 結果

本節では, パイロットスタディにおいて得られた結果を

会話分析から得られた結果およびアンケートから得られた結果に分けて示す。なお、今回のパイロットスタディの参加者は1組3人と少なかったため、統計的な解析を行わず、分析結果のみを示す。

まず、会話分析から得られた結果を示す。今回のパイロットスタディでは、参加者3人のうち、1人以上を見つめているときに、各参加者は直視しているとみなしている。また、会話分析では、会議開始から最大5秒間のログデータが欠損していたため、会議開始から5秒後以降から会議終了までの時間を分析対象としている。会議中に視線をそらした状態から対話相手を直視した状態に切り替えた回数(図3(A))に関して、SizeChange条件では他の2条件よりも会議中に視線をそらした状態から対話相手を直視した状態に切り替えた回数が多いことが示唆された。会議中に相互注視していた時間の割合(図3(B))に関して、SizeChange条件では他の2条件よりも会議中に相互注視していた時間の割合が高いことが示唆された。なお、今回のパイロットスタディでは、3人の参加者のうち、2人以上の参加者が他の参加者を直視しているときに相互注視していたと定義している。会議中に対話相手を直視していた時間の割合(図3(C))に関して、SizeChange条件では他の2条件よりも会議中に対話相手を直視していた時間の割合が高いことが示唆された。会議中の発話数(図3(D))に関して、各条件における会議中の発話数は同程度であることが示唆された。会議中の沈黙時間の割合に関して、Baseline条件では44.9%、FrameChange条件では33.7%、SizeChange条件では44.0%であったため、FrameChange条件では他の2条件より会議中の沈黙時間の割合が小さいことが示唆された。会議中に出されたアイデアの数に関して、Baseline条件では17個、FrameChange条件では13個、SizeChange条件では11個であったため、Baseline条件、FrameChange条件、SizeChange条件の順に会議中に提出されたアイデアの数が多いことが示唆された。

次に、アンケートから得られた結果を示す。会議中における対話相手の視線の変化の気づきやすさを測定するためのアンケートの結果(図4)に関して、Q1.「会話相手の見ている方向の変化に気づいた」において、FrameChange条件では他の2条件よりもスコアが高いことが示唆された。また、Q2.「会話相手が自分の方を見ている時、それが容易に分かった」およびQ3.「会話相手が自分の方を見ている時、目が合っているように感じた」において、FrameChange条件およびSizeChange条件ではBaseline条件よりもスコアが高いことが示唆された。会議中における話者交替の円滑さを測定するためのアンケートの結果(図5)に関して、Q4.「話者が変わる時、話始めが会話相手とぶつかった」、Q5.「話者が変わる時、話始めが会話相手とぶつかったことがストレスに感じた」、Q6.「話者が変わる時、話始めることなく、沈黙が発生した」、およびQ7.「話者が変わる時、

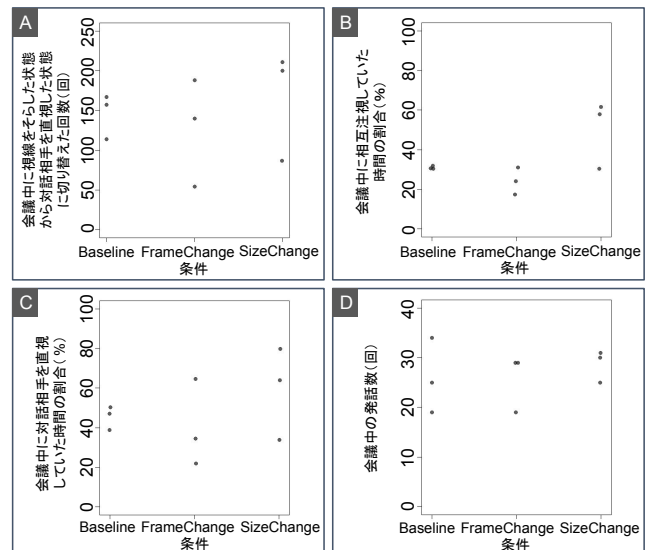


図3 会話分析から得られた一部の評価指標の結果。(A) 会議中に視線をそらした状態から対話相手を直視した状態に切り替えた回数。(B) 会議中に相互注視していた時間の割合。(C) 会議中に対話相手を直視していた時間の割合。(D) 会議中の発話数。

Fig. 3 Results of selected evaluation metrics derived from conversation analysis. (A) The number of times participants switched from averting gaze to gazing directly at their interlocutor during conference. (B) The percentage of time spent in mutual gaze during conference. (C) The percentage of time spent gazing directly at the interlocutor during conference. (D) The number of utterances during conference.

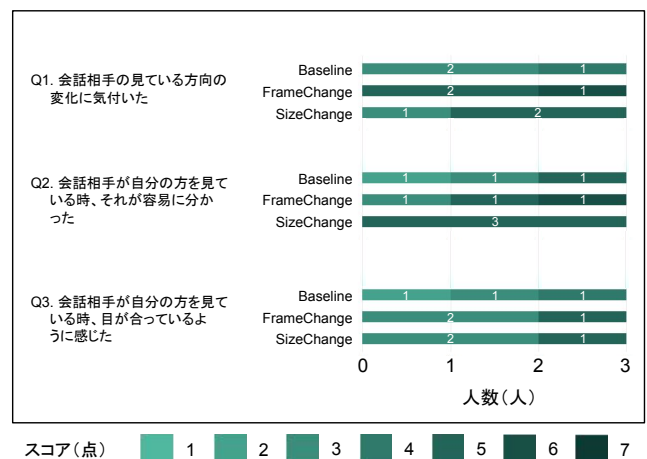


図4 会議中における対話相手の視線の変化の気づきやすさを測定するためのアンケートの結果。グラフ内の数字は度数を示す。

Fig. 4 The results of the questionnaire about noticing changes in the gaze of the interlocutor. The numbers within the graph represent frequencies.

話始めることなく、沈黙が発生したことがストレスに感じた」において、FrameChange条件では他の2条件よりもスコアが低いことが示唆された。また、Q8.「スムーズに議論が進んだ」およびQ9.「議論に集中できた」において、

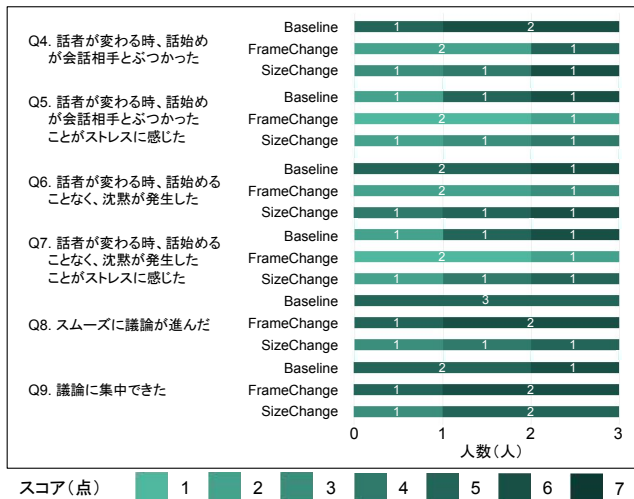


図 5 会議中における話者交替の円滑さを測定するためのアンケートの結果. グラフ内の数字は度数を示す.

Fig. 5 The results of the questionnaire about the smoothness of turn-taking during conference. The numbers within the graph represent frequencies.

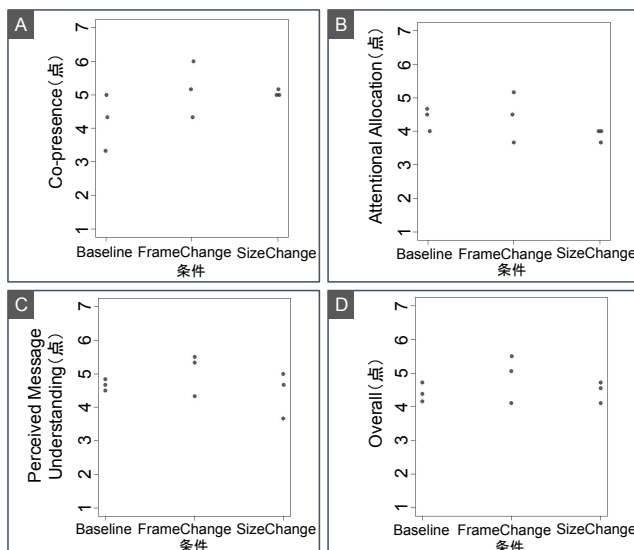


図 6 会議中の社会的存在感を測定するためのアンケートの結果. (A) Co-presence. (B) Attentional Allocation. (C) Perceived Message Understanding. (D) Overall.

Fig. 6 The results of the questionnaire about social presence. (A) Co-presence. (B) Attentional Allocation. (C) Perceived Message Understanding. (D) Overall.

FrameChange 条件および Baseline 条件では SizeChange 条件よりもスコアが高いことが示唆された. 会議中の社会的存在感を測定するためのアンケートの結果 (図 6) に関して, 各項目において, FrameChange 条件では他の 2 条件よりもスコアが高いことが示唆された. 会議に対する印象を測定するためのアンケートの結果 (図 7) に関して, 各項目において, FrameChange 条件および Baseline 条件では SizeChange 条件よりもスコアが高いことが示唆された. 各条件の好みの順位を測定するためのアンケートの結果

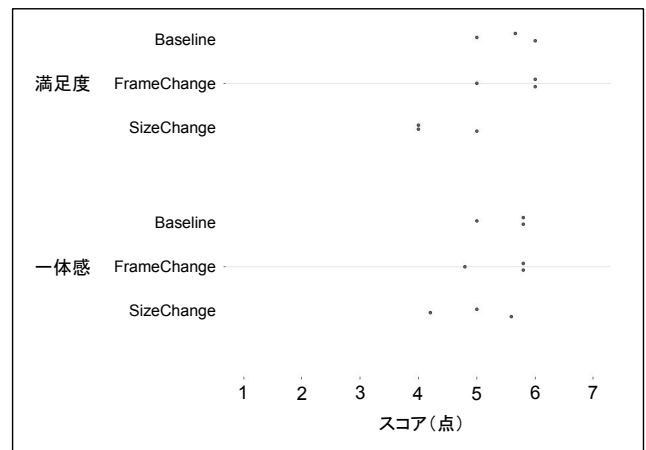


図 7 会議に対する印象を測定するためのアンケートの結果. Fig. 7 The results of the questionnaire about the impression of the conference.

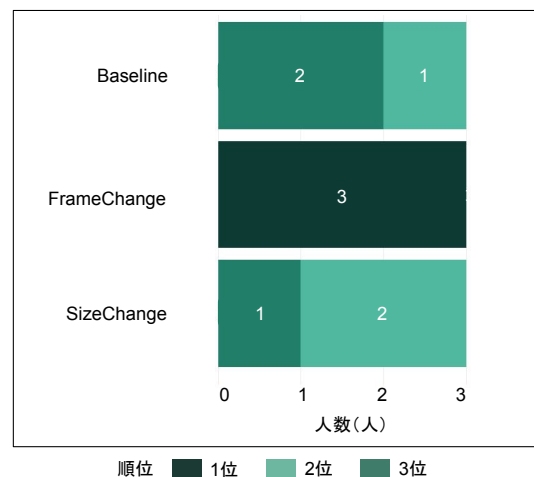


図 8 各条件の好みの順位を測定するためのアンケートの結果. グラフ内の数字は度数を示す.

Fig. 8 The results of the questionnaire about preference ranking. The numbers within the graph represent frequencies.

(図 8) に関して, パイロットスタディの参加者 3 人全員が FrameChange 条件を最も好ましい条件と判断したことが示唆された.

5. 考察

本節では, パイロットスタディの結果より, 対話相手の視線の変化の気づきやすさ, 話者交替の円滑さ, 対話相手の社会的存在感, およびビデオ通話の質に関して考察する.

5.1 対話相手の視線の変化の気づきやすさに関する考察

会議中に視線をそらした状態から対話相手を直視した状態に切り替えた回数 (図 3 (A)) より, SizeChange 条件において他の 2 条件よりもこの回数が多いことが示唆された. 会議中に相互注視していた時間の割合および会議中に対話相手を直視していた時間の割合の結果 (図 3

(B), 図3(C))より, SizeChange条件において他の2条件よりもこれらの割合が高いことが示唆された. 会議中における対話相手の視線の変化の気づきやすさを測定するためのアンケートのQ1.「会話相手の見ている方向の変化に気づいた」, Q2.「会話相手が自分の方を見ている時、それが容易に分かった」, およびQ3.「会話相手が自分の方を見ている時、目が合っているように感じた」の結果(図4)より, FrameChange条件およびSizeChange条件においてBaseline条件よりも各スコアが高いことが示唆された. これらに関して, 「FrameChange条件では, 完全にこちらを向いているときと向いていないときの違いが分かりやすかった」(P13)および「SizeChange条件では, 相手の視線が変化した瞬間に気づきやすかった」(P13)などの意見が得られた. これらの結果より, SizeChange条件, FrameChange条件, Baseline条件の順に対話相手の視線の変化の気づきやすさに寄与する可能性が示唆された. 一方, 先行実験では, FrameChange条件, SizeChange条件, Baseline条件の順に対話相手の視線の変化の気づきやすさに寄与することが示された. そのため, 提案手法では, 先行手法よりも, 対話相手の視線の変化への気づきやすさが改善される可能性がある.

5.2 話者交替の円滑さに関する考察

会議中の沈黙時間の割合に関して, FrameChange条件では, Baseline条件よりも, その割合が小さいことが示唆された. 会議中における話者交替の円滑さを測定するためのアンケートのQ4.「話者が変わる時、話始めが会話相手とぶつかった」, Q5.「話者が変わる時、話始めが会話相手とぶつかったことがストレスに感じた」, Q6.「話者が変わる時、話始めることなく、沈黙が発生した」, およびQ7.「話者が変わる時、話始めることなく、沈黙が発生したことがストレスに感じた」の結果(図5)より, FrameChange条件において他の2条件よりも各スコアが低いことが示唆された. これらに関して, 「FrameChange条件では, 話者交替をスムーズにかつストレスなくできた. つねに誰かが話している状態を維持できたと思う」(P15)という意見が得られた. 一方, 同アンケートのQ4.およびQ5.の結果より, SizeChange条件においてFrameChange条件よりも各スコアが高いことが示唆された. 同アンケートのQ8.「スムーズに議論が進んだ」およびQ9.「議論に集中できた」の結果より, SizeChange条件において他の2条件よりも各スコアが低いことが示唆された. これらに関して, 「SizeChange条件では, 大きさが頻繁に変わって集中できなかったため, 会話内容を一部聞き逃している気がする」(P14)および「SizeChange条件では, 相手も集中していないように感じて, 自分も集中できなかったように感じた」(P15)などの意見が得られた. 各条件における会議中の発話数の結果(図3(D))より, 各条件におけるこの数は同程度である

ことが示唆された. これらの結果より, FrameChange条件はBaseline条件よりも話者交替の円滑さに大きく寄与する可能性が示唆された. 加えて, SizeChange条件は他の2条件よりもこの項目に寄与できない可能性も示唆された. 一方, 先行実験では, SizeChange条件はBaseline条件よりも話者交替の円滑さに寄与することのみが示された. そのため, 提案手法では, 先行手法と同様に, 話者交替の円滑さが改善される可能性がある. また, 提案手法の1つであるSizeChange手法を話者交替の円滑さの向上につなげるために, その改善指針として, 閾値による離散変化の導入を検討している.

5.3 対話相手の社会的存在感に関する考察

会議中の社会的存在感を測定するためのアンケートのCo-presenceの結果(図6(A))より, FrameChange条件およびSizeChange条件においてBaseline条件よりも各スコアが高いことが示唆された. これに関して, 「FrameChange条件では, 自分が話している時に相手の枠が薄くなると, ちょっと悲しい気分になった」(P15)および「SizeChange条件では, 相手が片手間に会議に参加している印象を受けた」(P15)などの意見が得られた. 一方, 同アンケートの各項目の結果より, FrameChange条件およびSizeChange条件の各スコアは同程度であることが示唆された. これらの結果より, FrameChange条件およびSizeChange条件はBaseline条件よりも対話相手の社会的存在感の向上に寄与する可能性が示唆された. 加えて, FrameChange条件およびSizeChange条件における対話相手の社会的存在感の向上への寄与度合いは同程度である可能性も示唆された. 一方, 先行実験では, 対話相手の社会的存在感の向上への寄与度合いにおいて差は示されなかった. そのため, 提案手法では, 先行手法よりも, 対話相手の社会的存在感が改善される可能性がある.

5.4 ビデオ通話の質に関する考察

各条件の好みの順位を測定するためのアンケートの結果(図8)より, パイロットスタディの参加者3人全員がFrameChange条件を最も好ましい条件と判断した. 一方, 会議に対する印象を測定するためのアンケートの各項目の結果(図7)より, FrameChange条件およびBaseline条件において各スコアが同程度であることならびにSizeChange条件においてFrameChange条件よりも各スコアが低いことが示唆された. また, 会議中に出されたアイデアの数はBaseline条件, FrameChange条件, SizeChange条件の順に多い. これらの結果より, FrameChange条件およびBaseline条件におけるビデオ通話の質の向上への寄与度合いは同程度である可能性が示唆された. 加えて, SizeChange条件は他の2条件よりもこの項目に寄与できない可能性も示唆された. 一方, 先行実験では, ビデオ通話の質の向上へ

の寄与度合いにおいて差は示されなかった。そのため、提案手法では、先行手法と同じく、ビデオ通話の質が改善される可能性を現時点では期待できない。

5.5 考察のまとめ

ここまで示した考察より、提案手法は、先行手法よりも、対話相手の視線の変化の気づきやすさおよび対話相手の社会的存在感の向上に大きく寄与する可能性が示唆された。また、先行手法と同様に、話者交替の円滑さに寄与する可能性も示唆された。一方、先行手法と同じく、提案手法がビデオ通話の質の向上に寄与する可能性を現時点では期待できないことも示唆された。なお、今回のパイロットスタディの参加者は1組3人と少なかったため、この結果は人数不足であり参考程度の結果である。そのため、より多くの参加者に対する提案手法の評価実験において上記結果を確認することが今後必要である。

6. おわりに

本研究では、3人以上の状況を想定して、2人1組でのビデオ通話用の先行手法を3人1組でのビデオ通話用に拡張した。この提案手法のパイロットスタディを行った結果、提案手法は先行手法よりも対話相手の視線の変化の気づきやすさおよび対話相手の社会的存在感に大きく寄与する可能性が示唆された。加えて、先行手法と同様に、話者交替の円滑さに寄与する可能性も示唆された。今後、パイロットスタディにおいて得られた結果を基に、ビデオウィンドウの大きさを変更する提案手法の改善および3人1組でのビデオ通話における評価実験を行う。

参考文献

- [1] Emery, N. J.: The Eyes Have It: The Neuroethology, Function and Evolution of Social Gaze, *Neuroscience & Biobehavioral Reviews, International Behavioral Neuroscience Society*, Vol. 24, No. 6, pp. 581–604 (2000).
- [2] Kendon, A.: Some Functions of Gaze-Direction in Social Interaction, *Acta Psychologica, Elsevier B.V.*, Vol. 26, pp. 22–63 (1967).
- [3] Argyle, M. and Cook, M.: *Gaze and Mutual Gaze*, Cambridge University Press (1976).
- [4] Abeles, D. and Yuval-Greenberg, S.: Just look away: Gaze aversions as an overt attentional disengagement mechanism, *Cognition, Elsevier B.V.*, Vol. 168, pp. 99–109 (2017).
- [5] Glenberg, A. M., Schroeder, J. L. and Robertson, D. A.: Averting the gaze disengages the environment and facilitates remembering, *Memory & Cognition, The Psychonomic Society*, Vol. 26, No. 4, pp. 651–658 (1998).
- [6] Kendon, A.: Functions of Gaze in Social Interaction: Communication and Monitoring, *Journal of Nonverbal Behavior, Springer Nature*, Vol. 10, No. 2, pp. 83–101 (1986).
- [7] Regenbrecht, H. and Langlotz, T.: Mutual Gaze Support in Videoconferencing Reviewed, *Communications of the Association for Information Systems, AIS Electronic Library*, Vol. 37, No. 1, pp. 965–989 (2015).
- [8] Isaacs, E. A. and Tang, J. C.: What Video Can and Cannot Do for Collaboration: A Case Study, *Multimedia Systems, Springer Nature*, Vol. 2, No. 2, pp. 63–73 (1994).
- [9] 芦沢優介, 川口一画: オンライン会議における会話相手が自分自身を直視しているか否かのみでの伝達が参加者に与える影響の調査, 情報処理学会 情報処理学会論文誌, Vol. 67, No. 1, pp. 1–14 (2026).
- [10] 坊農真弓, 鈴木紀子, 片桐恭弘: 多人数会話における参与構造分析, 日本認知科学会 認知科学, Vol. 11, No. 3, pp. 214–227 (2004).
- [11] 飯塚陸斗: 多人数ビデオ会議における話者交替のための参与役割提示手法 (修士論文, 筑波大学大学院システム情報工学研究群, 2022).
- [12] Argyle, M. and Dean, J.: Eye-Contact, Distance and Affiliation, *Sociometry, American Sociological Association, Sage Publications, Incorporated.*, Vol. 28, No. 3, pp. 289–304 (1965).
- [13] Andrist, S., Tan, X. Z., Gleicher, M. and Mutlu, B.: Conversational Gaze Aversion for Humanlike Robots, *Proceedings of the 2014 Conference on Human-Robot Interaction, HRI '14, Association for Computing Machinery*, pp. 25–32 (2014).
- [14] Kumar, C., Saini, B. K. and Staab, S.: Enhancing Online Meeting Experience through Shared Gaze-Attention, *Proceedings of the 2024 Conference on Human Factors in Computing Systems-Extended Abstracts, CHI EA '24, Association for Computing Machinery*, pp. 128:1–128:6 (2024).
- [15] Iitsuka, R., Kawaguchi, I., Shizuki, B. and Takahashi, S.: Multi-party Video Conferencing System with Gaze Cues Representation for Turn-Taking, *Proceedings of the 2021 Conference on Collaboration Technologies and Social Computing, Springer-Verlag*, pp. 101–108 (2021).
- [16] docomo Business, N.: Skyway – NTT ドコモビジネスが提供する WebRTC アプリ開発者向けマルチプラットフォーム SDK&API, NTT docomo Business (オンライン), 入手先 (<https://skyway.ntt.com/>) (参照 2026-01-30).
- [17] Google: MediaPipe Solutions guide | Google AI Edge | Google AI for Developer, Google (オンライン), 入手先 (<https://ai.google.dev/edge/mediapipe/>) (参照 2026-01-30).
- [18] Google: Face landmark detection guide | Google AI Edge | Google AI for Developer, Google (オンライン), 入手先 (https://ai.google.dev/edge/mediapipe/solutions/vision/face_landmarker) (参照 2026-01-30).
- [19] Zoom: Meet face-to-face from any device | Zoom, Zoom Video Communications (オンライン), 入手先 (<https://explore.zoom.us/en/products/meetings/>) (参照 2026-01-30).
- [20] COMMUNICATIONS, P.: スマカメ 就活リモート会議カメラ「好感度 WEB カム」, PLANEX COMMUNICATIONS (オンライン), 入手先 (<https://www.planex.co.jp/products/usb-cam02/>) (参照 2026-01-30).
- [21] Meadow, A. and Parnes, S. J.: Evaluation of training in creative problem solving., *Journal of Applied Psychology, American Psychological Association*, Vol. 43, pp. 189–194 (1959).