小樽商科大学 学士論文

オンラインコミュニケーションにおける 視線を用いた話者交代を円滑にする手法の検討

2018101 笠原 有真

商学部 社会情報学科

令和3年度提出



目次

1章	はじめに	1
2章	関連研究	2
2.1	円滑なオンラインコミュニケーション	2
2.2	視線を利用した通話システム	2
3章	実験 1: 疑似的な視線を用いたシステム	4
3.1	システム概要	4
3.2	目的	5
3.3	実験の方法	5
3.4	実験参加者	7
3.5	結果	7
4章	実験 2: 視線推定を用いたシステム	10
4.1	システム概要	10
4.2	目的	11
4.3	方法	11
4.4	実験参加者	12
4.5	結果	13
5章	考察	15
5.1	主観的評価	15
5.2	発話権の衝突回数	15
5.3	視線の推定	16
5.4	場の共有	16
6章	おわりに	18
参老 文献		19

はじめに

コミュニケーションには「場の共有」が必要である。場の共有とは、同じ時間、同じ場所、同じ景色、同じ空気を共有することから得られる現象のことである。場の共有には、会話内容や声以外に、非言語情報を利用している。例えば、二人以上で道を歩きながら会話をするときには、歩く速さ、歩幅、変化する景色、環境音、話の内容、発話のタイミングなどが無意識のうちに共有される。それらの情報により引き込みが発生し、円滑なコミュニケーションが行わていると考えられる。[7]. 円滑な話者交代は、コミュニケーションにおける重要な要素であり、同じ場を共有することで発生する高度な機能である[11]. 一方で最近のオンラインによる遠隔でのコミュニケーションでは、得られる情報が削られることから、場の共有が発生せず話者交代が困難となっている。つまり、話し手の音声と顔の表情という限られた情報から、発話のタイミングを考えることになるため、同じタイミングで発話したり、無言が続いてしまう場合がある。その結果として円滑な話者交代が困難となる。

オンラインで共有できる情報には音や映像などがある。それら以外の情報として重要な要素に「視線の認識」がある。現在のオンラインコミュニケーションでは、互いの顔が写ったカメラ映像を共有しているものの、誰が何を見ているのか、視線を認識することができない。しかし視線の認識は、コミュニケーションにおいて話題に対する興味の有無、あるいは視線が集まることによる発話の誘発などの情報が含まれており、重要な情報のひとつである[4][14].

そこで、本研究では、オンラインのコミュニケーションに視線情報を付与することにより、「適切な話者交代」を行うことで、場の共有をつくりだすことを目的とする.

関連研究

2.1 円滑なオンラインコミュニケーション

オンラインコミュニケーションシステムにおいて、円滑にコミュニケーションを行うための研究はさまざまな観点から行われてきた。徳差らは、オンラインコミュニケーションにおいて、一番の阻害要因はネットワークの遅延であるとし、超低遅延の通話システムの設計を行っている [13]. また、Martin らは、ビデオ付きの通話について、複数地点からの映像を組み合わせることで、会話の主観的な評価が向上することを明らかにした [2]. 非言語情報を用いた研究として、玉木らはWeb 会議において、参加者の動作から発話欲求を推定し、ほかの参加者に伝達することによって円滑な話者交代を行う手法についての研究を行っている [9]. また、渡辺らはバーチャルアクターの非言語情報を作り出し、仮想空間上での円滑なコミュニケーションについての研究を行なっている [12]. 円滑なコミュニケーションを推進するための商用サービスとしては、SpatialChat*1やGather.Town*2といった、参加者が仮想空間を自由に移動し会話時の物理的な距離を再現するサービスも存在する.

2.2 視線を利用した通話システム

非言語情報の中で、特に視線に着目した研究も行われている。Sellen は、対面の会話、複数台のカメラとモニタを用いて視線情報を伝達できるビデオ通話システム(Hydra)を使った会話、ビデオとマイクのみのビデオ通話システムを使った会話の比較実験を行った。その結果、話者交代にかかる時間が、対面よりもビデオ通話で有意に長くなることが確認されているが、Hydra システムと通常のビデオ通話システムの間で有意な差は見られないことを明らかにしている[6]。一方で、実験参加者の主観的な評価においては、3分の2が Hydra システムを好むという結果が得られている。西村らは、視線を伝達できるテレビ会議システムを用いることで話者交代がスムーズに実現できることを確認している[8]。しかし、これらのシステムは複数台のカメラや視線検出のため複雑

^{*1} SpatialChat(https://spatial.chat/)

^{*2} Gather.Town(https://gather.town/)

なシステムとなり、一般的な用途で再現することは難しい.

特別な外部システムを用いずに視線情報を伝達する試みとして、石井らの、音声情報によるアバターの視線制御がある [15]. また、Zhenyi らは視線の動きのみを伝達可能な通話システム (GazeChat) を開発し、プライバシーを保護しながら効率的な会議を実現する手法について研究している [3]. しかし、この手法では視線情報以外の非言語情報が遮断されてしまう. 単一の Web カメラから視線を推定可能にする研究も行われているが、認識精度にはまだ課題がある [5][10].

本研究では、特別な外部システムを必要とせず簡単に視線を再現する手法としてマウスポインタを用いる. さらに、この視線情報をビデオ通話システムに組み込んだ際に、どのような影響があるのか明らかにする. また、単一の Web カメラから視線を推定できる WebGazer.js というライブラリを利用したシステムを用いた実験も行う.

実験 1: 疑似的な視線を用いたシステム

本研究では、一般的なユーザが再現可能な手法を用いて視線情報を伝達した場合、円滑な話者交 代が生まれるかどうかを明らかにするため、まずはマウスポインタを疑似的な視線とし、それを相 互に伝達するシステムを用いた実験を行った。疑似的な視線を用いる理由としては、現在の視線推 定が完璧な精度ではないため、自分の思った通りの場所へ視線の伝達が可能ではないことが実験の ノイズとなることを防ぐためである。

3.1 システム概要

本実験で使用するマウスポインタを疑似的な視線とした通話システムを図 3.1 に示す. ユーザは

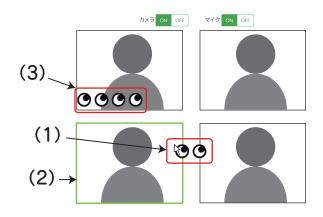


図 3.1 擬似的な視線を用いた通話システムの画面(左上のウィンドウに映っているユーザが操作している場合)

マウスポインタを操作し、疑似的な視線によってほかのユーザに注意を向ける。この時、マウスポインタには目のアイコンが追従し(図 1 の (1))、ポインタを置いているユーザのウィンドウに緑色の枠が表示される(図 1 の (2))。一方で、自らがほかのユーザから擬似的な視線を向けられている場合は、その人数分の目のアイコンがウィンドウの左下に表示される(図 1 の (3))。このウィンド

ウ内に表示されるアイコンは自分に向けられたものしか見ることができず, ほかのユーザがどの程 度見られているかという情報は提示されない.

3.2 目的

本実験の目的は、複数人のオンラインコミュニケーションにおいて、擬似的な視線を付与することにより適切な話者交代が行えることを明らかにすることである。具体的には、本研究で開発したシステムを用いて、話者交代に関する下記の3つの項目について考察する。

- 1. 衝突回数がどのように変化するのか明らかにする.
- 2. 沈黙時間がどのように変化するのか明らかにする.
- 3. ユーザの主観的評価がどのように変化するのか明らかにする.

3.3 実験の方法

実験は、第一著者を含めた大学生 4 人で行い、短時間で自然な会話を誘発するためにワードウルフというソーシャルコミュニケーションゲームを行った.ワードウルフは多数派と少数派のワードがそれぞれの参加者に割り当てられ、会話の中から質問などを通して少数派のワードを割り当てられた参加者を当てるゲームである.それぞれの参加者は自分がどちらのワードを割り当てられたのかわからないため、会話の中で他者の発言に反応したり質問することでそれを明らかにする必要がある.このゲームを,疑似的な視線情報を伝達できる通話システムおよびこのシステムから視線情報の伝達機能を削り、ビデオ通話のみが行える通話システムをそれぞれ用いて行った.参加者がタスクを行っている最中は、第一著者はカメラとマイクをオフにし、実験の様子を観察した.タスク終了後、カメラとマイクをオンにし、投票結果の集計および報告を行った.また、各グループにおいて1つのセッションを2分間とし、それぞれの通話システムで3セッションずつゲームを行った.ワードウルフに使用した単語群は、表3.1にある12個である[1]実験参加者はそれぞれの自宅から、PC上で動作するWebブラウザを通してシステムを利用した.実際の実験の様子を図3.2に示す.

表 3.1 ワードウルフに使用した単語

多数派	少数派
しゃぶしゃぶ	すき焼き
せんべい	クッキー
トマトジュース	オレンジジュース
公園	遊園地
金魚	カメ
天気予報	星座占い
	しゃぶしゃぶ せんべい トマトジュース 公園 金魚



図 3.2 実験の様子

実験の流れを図 3.3 に示す。表 3.1 にある単語のうち,前半 3 セッションで使用する飲食系に関する 6 単語を単語群 A,後半 3 セッションで使用する日常のお題に関する 6 単語を単語群 B としている.

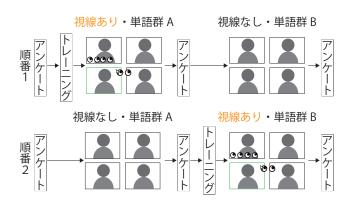


図 3.3 実験の流れ

参加者が視線付き通話システムを使う際には、視線情報がどのように伝達されるのかを説明した上で、「機能をうまく活用して、少数派を探し出して下さい」と、擬似的な視線の使用を促した.実験を始める前に、ゲームの流れを参加者に知ってもらうため、第一著者も参加した状態でのワードウルフを 1 回行った.また、擬似的な視線を参加者に慣れてもらうために、視線付き通話システムを用いた実験を行う前に、そのシステムを使用して、第一著者も参加したワードウルフを 1 回行った(図 3.3 のトレーニング).

事前アンケートでは、参加者は、「氏名」「性別」「学科」「オンラインツールの使用頻度」につい

て回答する. また、各通話システムでの実験終了後、参加者は以下の設問に 5 段階のリッカート尺度で回答する(1 が「全く当てはまらない」,5 が「非常に当てはまる」に対応する). なお、これらの質問は西村らの研究 [8] を参考にし、本研究に合う項目を利用した.

- Q1 無理なく発言することができた
- Q2 タイミングの悪い発言が多かった
- Q3 ぎこちないやりとりが多かった
- O4 不自然な沈黙が多かった
- Q5 他の人が自分に注意を払っているのがわかった

さらに、アンケートの最後に「その他お気づきの点がありましたらご記入ください」という自由 記述欄を設けた.

視線付きの通話システムを用いた実験においては追加項目として「疑似的な視線があってよかった点は何ですか?」「疑似的な視線があって気になった点は何ですか?」という自由記述欄を設けた.参加者はこの自由記述欄と以下の追加設問に5段階のリッカート尺度で回答する(1が「全く当てはまらない」、5が「非常に当てはまる」に対応する).

- 1. 視線(マウスカーソル)を意識して動かしましたか?
- 2. 視線が集まっていることは意識しましたか?

3.4 実験参加者

参加者は、小樽商科大学の商学部 1 年生 6 名および 2 年生 6 名(社会情報学科 2 名、企業法学科 3 名、経済学科 1 名)の計 12 名(男性 9 名、女性 3 名)、平均 19.6±1.0(標準偏差)歳であった. 実験においては、コミュニケーションがある程度円滑に行えるよう、学年を揃え、男性 2 名、女性 1 名となるようにグループを構成した.その結果、4 グループが作成され、この内の 1 グループは 男性のみのグループであった.参加者のオンラインツール利用頻度は、「日常的に使用する」が 6 名、「時折使用する」が 6 名であり「ほとんど使ったことがない」と回答した参加者はいなかった. また、2 グループが図 3.3 に示す順番 1 の流れで実験を行い、2 グループが順番 2 の流れで実験を行った.

3.5 結果

衝突回数について、視線ありの場合には発話時に合計 196 回のうち 17 回の衝突が発生し、視線なしの場合には発話時に合計 180 回のうち 17 回の衝突が発生した。この衝突回数についてフィッシャーの正確確率検定を行ったところ有意な差は確認されなかった($p=.858, \phi=.004$)。さらに、各参加者の合計沈黙時間について視線あり(平均:313 秒、標準偏差:27 秒)と視線なし(平均:316 秒、標準偏差:25 秒)の条件で対応のある t 検定を行ったところ、有意な差は確認されな

かった (t(11) = 0.406, p = .693, d = .110).

 $Q1\sim Q5$ のアンケート結果を表 3.2 に,そのグラフを図 3.4 示す.それぞれのアンケート項目について,対応のある t 検定を行った.その結果,Q1「無理なく発言することができた」 (t(11)=0.609,p=.555,d=.281),Q2「タイミングの悪い発言が多かった」(t(11)=0.00,p=1.00,d=.000),Q3「ぎこちないやりとりが多かった」(t(11)=1.173,p=.266,d=.500),Q4「不自然な沈黙が多かった」(t(11)=1.465,p=.171,d=.472) については有意な差は確認されなかった.一方で,Q5「他の人が自分に注意を払っているのがわかった」(t(11)=4.330,p=.001,d=1.91) については有意な差が確認された.「視線(マウスカーソル)を意識して動かしましたか?」については,平均 3.83,標準偏差 0.937 であった.また,「視線が集まっていることは意識しましたか?』については,平均 4.33,標準偏差 0.651 であった.

	視線あり		視	線なし
	平均	標準偏差	平均	標準偏差
Q1	3.67	0.78	3.42	1.00
Q2	2.42	0.90	2.42	1.31
Q3	2.75	1.36	3.42	1.31
Q4	2.75	1.06	3.33	1.37
Q5	4.17	0.94	2.33	0.98

表 3.2 アンケート結果

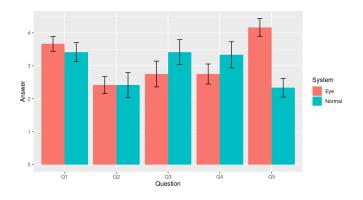


図 3.4 アンケート結果のグラフ (エラーバーは標準誤差を表している)

表 3.3 に「疑似的な視線があってよかった点」についての結果を示す。表 3.3 より、「発言に注意した」という回答が最も多く、3 人から回答があった。

表 3.4 に「疑似的な視線があって気になった点」についての結果を示す.表 3.4 より,「落ち着かなかった」「誰に見られているか分からない」という回答がそれぞれ 3 件ずつあった.

実験の最後に行ったアンケートにて、今回の実験に使ったシステムについて、視線なしの通常の 通話システムと、視線付きの通話システムどちらが良かったのかを聞いたところ、「視線付きの通

表 3.3 「疑似的な視線があってよかった点」のまとめ

アンケートの回答	回答数
発言に注意した	3
発言を促せた	2
圧をかけられた	2
使える情報が増えた	1
話しやすかった	1
疑われているのが分かった	1
会話を誰に投げかけているか分かる	1
見られていると積極的に話せた	1

表 3.4 「疑似的な視線があって気になった点」のまとめ

アンケートの回答	回答数
落ち着かなかった	3
誰に見られているか分からない	3
対面で話すときより視線を意識した	1
意識的に見なければならない	1
相手が発言するのではと気になった	1
視線を動かすのを忘れた	1
視線が集まると話さなければ	1
いけないと感じた	1
特になし	1

話システムの方が良かった」が 8 件,「通常の通話システムの方が良かった」が 1 件,「どちらも変わらなかった」が 3 件という結果になった.

実験 2: 視線推定を用いたシステム

疑似的な視線を用いた実験では、主観的な評価において疑似的であっても視線情報があった方がよい可能性が示された. しかし、視線を動かした位置にマウスポインタを動かさなくてはならないという行動が障害となってしまうことがわかった. そこで、マウスポインタを動かさずとも視線情報が伝達できるように、単一の Web カメラから視線を推定した結果を伝達可能なシステムを用いた実験を行う.

4.1 システム概要

本実験で使用する、視線情報を伝達可能な通話システムを図 4.1 に示す.

本システムには、単一の Web カメラからの映像のみを使用する視線トラッキングライブラリである WebGazer.js を用いた. 推定されたユーザの視線位置は赤い点で表現され、視線の先にいるユーザの映像の周りには緑色の枠が表示される. 実験 1 のシステムと同じく、自らがほかのユーザから視線を向けられている場合は、その人数分の目のアイコンがウィンドウの左下に表示されるが、誰が見ているかという情報は提示されない.



図 4.1 視線推定を用いた通話システムの画面(左上のウィンドウに映っているユーザが操作している場合)

4.2 目的

本実験の目的は、現状の完璧ではない視線認識を使った場合でも、マウスポインタを操作するという障害を排除したらユーザの主観的評価は上がるのかを明らかにすることである。そのため、以下の項目について評価を行う.

- 1. 衝突回数がどのように変化するのか明らかにする.
- 2. 沈黙時間がどのように変化するのか明らかにする.
- 3. ユーザの主観的評価がどのように変化するのか明らかにする.

4.3 方法

実権は第一著者を含めた大学生 4 人で行い、自然な会話を誘発するために、あらかじめ用意したテーマを提示し、そのテーマを起点とした会話を行ってもらった。視線情報を伝達できる通話システムおよび本システムから視線情報の伝達機能を削り、ビデオ通話のみが行える通話システムを用いて行った。実験参加者は、各システムでそれぞれ1つのテーマについて 10 分ずつ会話を行った。参加者がタスクを行っている最中は、第一著者はカメラとマイクをオフにし、実験の様子を観察した。会話に使用したテーマは以下の2つである。

- A オンライン講義とオフライン講義のどちらがいいか
- B 卒業後, 道内に居たいか道外に出たいか

また、視線情報を伝達可能なシステムを用いる実験の前には、視線の動きを推定するためにキャリブレーションを行った。視線推定の精度に差が出ないように、Web カメラの性能及びモニタのサイズは全参加者で統一させた。また、視線推定の精度は80%以上になったことを確認して実験を行った。

実験の流れを図 4.2 に示す.

実験を始める前に、ウォーミングアップを兼ねて3分間の会話を行った.

事前アンケートでは,参加者は,「氏名」「性別」「年齢」「学科」「オンラインツールの使用頻度」について回答する.また,各通話システムでの実験終了後,参加者は以下の設問に 5 段階のリッカート尺度で回答する(1 が「全く当てはまらない」,5 が「非常に当てはまる」に対応する).なお,これらの質問は西村らの研究 [8] 及び Zhenyi らの研究 [3] を参考にし,本研究に合う項目を利用した.

- O1 無理なく発言することができた
- Q2 タイミングの悪い発言が多かった
- Q3 ぎこちないやりとりが多かった

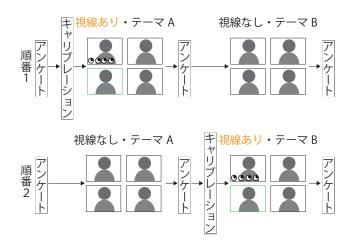


図 4.2 実験の流れ

- Q4 不自然な沈黙が多かった
- O5 他の人が自分に注意を払っているのがわかった
- Q6 会話が活発に行われた
- Q7 会話がスムーズに進んだ

さらに、各アンケートの最後に「その他お気づきの点がありましたらご記入ください」という自由記述欄を設けた.

視線付きの通話システムを用いた実験においては追加項目として「視線があってよかった点は何ですか?」「視線があって気になった点は何ですか?」という自由記述欄を設けた.参加者はこの自由記述欄と以下の追加設問に5段階のリッカート尺度で回答する(1が「全く当てはまらない」、5が「非常に当てはまる」に対応する).

- 1. 視線を意識して動かしましたか?
- 2. 視線が集まっていることは意識しましたか?

実験の最後には事後アンケートを行い, 視線情報が伝達可能な通話システムと, 通常のビデオ通話システムのどちらが良いか, およびその理由の記述欄を設けた.

4.4 実験参加者

参加者は、小樽商科大学の商学部 1 年生 1 名および 2 年生 6 名(社会情報学科 2 名、企業法学科 3 名、経済学科 1 名)、3 年生 4 名、4 年生 1 名の計 12 名(男性 5 名、女性 7 名),平均 20 ± 1.1 (標準偏差)歳であった.参加者のオンラインツール利用頻度は、「日常的に使用する」が 5 名、「時折使用する」が 7 名であり「ほとんど使ったことがない」と回答した参加者はいなかった.また、2 グループが図 4.2 に示す順番 1 の流れで実験を行い、2 グループが順番 2 の流れで実験を行った.

4.5 結果

衝突回数および沈黙時間の計測には、10 分間の会話のうち、前後 1 分間を除いた 8 分間を対象とした。衝突回数について、視線ありの場合には発話時に合計 214 回のうち 15 回の衝突が発生し、視線なしの場合には発話時に合計 200 回のうち 29 回の衝突が発生した。この衝突回数についてフィッシャーの正確確率検定を行ったところ、視線ありの場合で有意に衝突回数が低いことがわかった。 $(p=.03827,\phi=2.065)$ 。また、各参加者の合計沈黙時間について視線あり(平均:64 秒、標準偏差:10 秒)と視線なし(平均:48 秒、標準偏差:25 秒)の条件で対応のある 1 検定を行ったところ、有意な差は確認されなかった (t(3)=-1.657,p=.1961,d=-15.79).

 $Q1\sim Q5$ のアンケート結果を表 4.1 に,そのグラフを図 4.3 示す.それぞれのアンケート項目について,対応のある t 検定を行った.その結果,Q1 「無理なく発言することができた」 (t(11)=0.00,p=1,d=0),Q4「不自然な沈黙が多かった」(t(11)=1.8202,p=.096,d=..833),Q5 「他の人が自分に注意を払っているのがわかった」 (t(11)=-0.748,p=.470,d-0.417) についてはについては有意な差は確認されなかった.一方で,Q2 「タイミングの悪い発言が多かった」 (t(11)=-2.399,p=.0353,d=1.0833),Q3 「ぎこちないやりとりが多かった」 (t(11)=-2.777,p=.0180,d=-1.583) については有意差が確認された.「視線を意識して動かしましたか?」については,平均 4.083,標準偏差 1.240 であった.また,「視線が集まっていることは意識しましたか?」については,平均 3.917,標準偏差 1.084 であった.

	視線あり		視線なし	
	平均	標準偏差	平均	標準偏差
Q1	4.08	1.24	4.08	1.00
Q2	3.92	1.08	2.83	0.94
Q3	4.17	1.03	2.58	1.38
Q4	2.08	0.67	2.92	1.24
Q5	2.83	1.34	2.42	1.31
Q6	3.42	1.24	3.92	0.90
Q7	3.33	1.15	3.67	0.99

表 4.1 アンケート結果

表 4.2 に「視線があってよかった点」についての結果を示す.表 4.2 より,「話に集中してくれているのがわかった」という回答が最も多く,3人から回答があった.

表 4.3 に「視線があって気になった点」についての結果を示す.

実験の最後に行ったアンケートにて、今回の実験に使ったシステムについて、視線なしの通常の 通話システムと、視線付きの通話システムどちらが良かったのかを聞いたところ、「視線付きの通

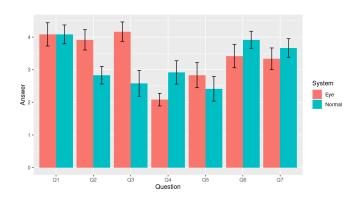


図 4.3 アンケート結果のグラフ (エラーバーは標準誤差を表している)

表 4.2 「視線があってよかった点」のまとめ

アンケートの回答	回答数
話に集中してくれているのがわかった	6
発言を促せた	1
普段より人の顔を見て話を聞けた	2
注目されているのかどうかがわかった	3

表 4.3 「視線があって気になった点」のまとめ

アンケートの回答	回答数
視線の推定位置との差が気になった	3
自分が注目されていると思い緊張した	2
視線を動かすことを意識しなければならないと感じた	3

話システムの方が良かった」が4件,「通常の通話システムの方が良かった」が8件という結果になった.

考察

5.1 主観的評価

実験 1 では,通常の通話システムと視線付き通話システムにおいて,衝突回数と沈黙時間,アンケートの $Q1\sim Q4$ に関しては統計的に有意な差は見られなかった.しかし,アンケートの結果から,実験参加者のうち,3 分の 2 が通常の通話システムよりも疑似的な視線付き通話システムの方を好んでいた.これは,アンケートの自由記述において,「使える情報が増えた」「疑われているのが分かった」「会話を誰に投げかけているか分かる」という回答から,映像と音声に加えて疑似的な視線情報が与えられることで,自分が他の参加者からどの程度意識を向けられているかという状況把握が容易になり,話しやすいと感じた参加者が多かったためであると考えられる.

視線推定システムを使用した実験2では、アンケートの結果から「タイミングの悪い発言が多かった」「ぎこちないやりとりが多かった」が視線ありのシステムで有意に多かった。また、3分の2の実験参加者が視線ありのシステムよりも視線なしのシステムを好んでいた。その理由として、「視線をうまく動かすことに意識が向いてしまった」「実際に見ているところとの差が気になってしまった」というものがあげられた。しかし、自由記述においては最初の実験と同じく、視線ありの場合のほうが使える情報が増えて良かったと答えた参加者は多かったため、視線情報を受け取るうえでは有用ではあったと考えられる。

どちらの実験においても、自然に完璧な視線情報を相手に伝えるということはできない状況であった.しかし実験2の結果から、視線情報が自分の意志とは異なった形で伝わることのほうがユーザにとってはストレスに感じるということが考えられる.

5.2 発話権の衝突回数

本研究では視線情報を伝達するシステムを用いて、円滑な話者交代を実現することを目的とし、発話権の衝突回数を比較した。その結果、実験2では視線ありのシステムのほうが衝突回数が有意に少なくなることがわかった。実験1では有意差は確認されなかったことから、これは実際の視線を推定していることの特徴であると考えられる。実験1では、マウスポインタをあまり頻繁には動

かさないユーザがいた,しかし,実験2では実際の視線を推定し,ユーザの意思とは無関係に視線情報を伝達する.そのため,視線情報の伝達量が増え,発話すべきタイミングがつかみやすかったのではないかと考えられる.

5.3 視線の推定

実験2で使用したシステムについて、アンケートにおいて「視線の推定位置との差が気になった」という回答が得られた.実験中の会話においても、視線の精度について言及している場面が見られた.本実験では、視線推定の精度が80%以上であることを確認して実験を行い、実験開始前にそれぞれの参加者に正しく視線情報が伝達されることも確認した.しかし、オンラインコミュニケーションにおいては、ビデオ映像を使用する場合、身振り手振りを同時に行ったり、顔の向きや位置を頻繁に変更するという場面が多く見られた.そのため、実験中に徐々に視線の推定位置がずれてきてしまい、差が大きくなっていったと考えられる.今後は推定位置がずれないような実験条件を設定することが必要である.

5.4 場の共有

実験1では、疑似的な視線情報を付与することにより、場の共有を作り出すことを目指した.その結果、アンケートにおいて「視線を感じて落ち着かなかった」「見られていることが明白で発言に注意した」「視線を送ることで圧をかけられた」といった、視線を強く意識した回答がみられた.したがって、視線付き通話システムは通常の通話システムよりもお互いの存在をより意識して会話を行っており、場の共有感をより強く感じられたのではないかと考えられる.また、アンケートより、「会話に使える情報が増えた」「見られていることを意識した」という意見が見られたため、参加者の主観的には、マウスポインタを用いた疑似的な視線であっても、実際の視線情報と同じような情報として捉えられていたのではないかと考えられる.

視線推定を行った実験 2 においては、視線情報を実際の視線からとっていたため、マウスポインタを使用した際よりも情報量が増え、客観的な評価においても有意差が確認された。また、アンケートで実験 1 と同じように「話を聞いてもらえている実感があった」「話してほしい相手に合図を送ることができた」という回答が見られたため、情報の受け取り手としては、場の共有感が高まったと感じられたのではないかと考えられる。しかし、「視線の認識の差が気になった」や、「視線の位置が気になり、話すことに集中できなかった」という回答から、情報を発信する側としては、あまり場の共有感を高めることができなかったと考えられる。

また、誰が誰を見ているのかという情報は本研究のシステムではわからなかったため、場の共有という点においては、全員の視線情報を公開し、視線の重なりや意図的な注視を共有できるようにする必要があると考えられる。さらに、実験2において、ユーザが視線を画面外へ向けるという場面が多く見られた。その際、本実験のシステムでは、一番近い位置のユーザへ視線情報が伝達されてしまい、視線を外している状態が再現されない。場の共有においては、視線を外して何かを考え

ていることや、その視線の先に何があるかといったことが重要になってくると考えられるため、それらの情報も伝達可能にする必要があると考えられる.

おわりに

本稿では、オンラインのコミュニケーションにおいて、一般的なユーザが利用可能な機材を用いて視線情報を付与することにより、「適切な話者交代」を行うことで、場の共有をつくりだすことを目的とした。具体的には、マウスポインタを操作することによって擬似的な視線を再現するシステム、および単一の Web カメラから視線を推定するシステムを用い、その視線の有無による比較実験を行い、結果について述べた。本実験の結果から、疑似的な視線であっても、視線を強く意識して会話を行っていることが分かった。また、視線情報をユーザの意思とは無関係に伝達することは、情報を受け取る側としては有用であるが、発信する側としては、わずかな差でも気になってしまい、主観的評価に大きな影響を与える可能性が示された。

今後は、ユーザがより自然に意図した場所へ視線情報を伝達できるシステム、またはより高い精度の視線推定システムを用い、場の共有感をより高められるシステムを提案していく.

参考文献

- [1] ワードウルフのお題 200 問! 恋愛・スポーツ・食べ物系のお題も. https://kyoheiomi.com/board-game/recommendation/wordwolf, jun 2021.
- [2] Martin Groen, Marian Ursu, Spiros Michalakopoulos, Manolis Falelakis, and Epameinondas Gasparis. Improving video-mediated communication with orchestration. Computers in Human Behavior, 28(5):1575 1579, 2012.
- [3] Zhenyi He, Keru Wang, Brandon Yushan Feng, Ruofei Du, and Ken Perlin. <u>GazeChat: Enhancing Virtual Conferences with Gaze-Aware 3D Photos</u>, page 769–782. Association for Computing Machinery, New York, NY, USA, 2021.
- [4] Adam Kendon. Some functions of gaze-direction in social interaction. <u>Acta Psychologica</u>, 26:22 63, 1967.
- [5] Alexandra Papoutsaki, Patsorn Sangkloy, James Laskey, Nediyana Daskalova, Jeff Huang, and James Hays. Webgazer: Scalable webcam eye tracking using user interactions. In <u>Proceedings</u> of the 25th International Joint Conference on Artificial Intelligence (IJCAI), pages 3839–3845. AAAI, 2016.
- [6] Abigail J. Sellen. Speech Patterns in Video-Mediated Conversations. In <u>Proceedings of the SIGCHI Conference on Human Factors in Computing Systems</u>, CHI '92, page 49–59, New York, NY, USA, 1992. Association for Computing Machinery.
- [7] 小磯 花絵 and 伝 康晴. 円滑な話者交替はいかにして成立するか. 認知科学, 7(1):93-106, 2000.
- [8] 西村 圭亮, 上野 晃嗣, 坪井 創吾, and 下郡 信宏. テレビ会議において視線の伝達が話者交替に及ぼす影響の分析. Technical Report 33(2009-GN-71), 株式会社東芝研究開発センター, 株式会社東芝研究開発センター, 株式会社東芝研究開発センター, 株式会社東芝研究開発センター, mar 2009.
- [9] 玉木 秀和, 東野 豪, 小林 稔, 井原 雅行, and 岡田 謙一. 遠隔会議における発話衝突低減手法. 情報処理学会論文誌, 53(7):1797–1806, jul 2012.
- [10] 福田 崇, 松崎 勝彦, and 山名 早人. 安価な web カメラを用いた model-based 視線推定. 電子情報通信学会技術研究報告. HIP, ヒューマン情報処理, 109(471):113–118, mar 2010.
- [11] 富夫 渡辺 and 雅史 大久保. コミュニケーションにおける引き込み現象の生理的側面からの分析評価. 情報処理学会論文誌, 39(5):1225–1231, may 1998.

- [12] 富夫 渡辺 and 雅史 大久保. 身体的コミュニケーション解析のためのバーチャルコミュニケーションシステム. 情報処理学会論文誌, 40(2):670–676, feb 1999.
- [13] 徳 差 雄 太, 松 谷 健 史, 空 閑 洋 平, and 村 井 純. 低 遅 延 に よ り 自 然 な 遠 隔 コ ミ ュ ニ ケ ー シ ョ ン を 実 現 す る 映 像 配 信 シ ス テ ム の 提 案. In マルチメディア, 分散協調とモバイルシンポジウム 2013 論文集, volume 2013, pages 911–917, jul 2013.
- [14] 榎本 美香 and 伝 康晴. 話し手の視線の向け先は次話者になるか<特集>相互作用のマルチモーダル分析). 社会言語科学, 14(1):97–109, 2011.
- [15] 石井 亮, 宮島 俊光, and 藤田 欣也. アバタ音声チャットシステムにおける会話促進のための注 視制御. ヒューマンインタフェース学会論文誌, 10(1):87–94, 2008.