

視線情報を用いたアウェアネスの提示による遠隔対話の生起支援

田之頭 吾音^{*} 川口 一画[†] 志築 文太郎[†] 高橋 伸[†]

概要. 遠隔地間においては、対面状況と比較して対話が生起しにくい。対話の中でも、時間の決まっているフォーマルコミュニケーションよりも、偶発的に発生すると言われるインフォーマルコミュニケーションの生起は困難である。本手法では、社会学的知見に基づき、対面状況において対話が生起される際に交わされる動きを遠隔地間において再現することにより対話の生起を支援する。本手法を実現するために、遠隔地で対話を行う二人のユーザ間に、視線情報および音声情報を伝達する物理アバタをそれぞれ一台ずつ配置する。

1 はじめに

近年、Zoom[13] や Microsoft Teams[10] のような遠隔コミュニケーション支援のためのシステムが普及している。これらのシステムは主に、日時および議題があらかじめ決まっているフォーマルコミュニケーションのために用いられる。これに対し、日時および議題があらかじめ決まっておらず、偶発的に発生するコミュニケーションは、インフォーマルコミュニケーションと呼ばれる。仕事の合間の雑談等がその一例であり、Egido[5] によればインフォーマルコミュニケーションの場において重要な情報のやり取りが行われ、それにより実際の決定がなされることも多い。インフォーマルコミュニケーションのような何気ない対話が開始されるためには、「相手に話しかける」という行為の心理的負担を軽減する必要がある。そのための手段として、物理的な人形の動作を用いて相手に対話に応答可能かどうかを伝達するアウェアネス提示システムが提案されている[8]。アウェアネスとは Dourish ら [4] により「自分の活動や状況に関わる他人の活動や状況を示すもの」と定義されている。これらのシステムを用いることにより常に相手の応答可否を知ることができ、話しかける行為の心理的負担を軽減できる。

一方で、対面状況においては相手が応答可能かどうかだけでなく、互いの状態（作業にどの程度集中しているか等）をより細かく伝え合うことによりタイミングを計り、相手の作業を妨げずに対話を開始できる。そこで本研究では、既存システムで提示されていた相手の応答可否だけでなく、より細かな状態を相互に伝達可能なシステムを提案し、話しかける行為の心理的負担をより軽減することを目指す。

イニシエーショントークに関する社会学的知見によると、対話が開始する際には視線のやり取りが重要な役割を果たしており、特に相互注視は対話開始



図 1: 物理アバタ



図 2: 提案システムを注視する様子

の合図になる。この知見に基づき、本研究では遠隔地間において互いの視線情報を物理的なアバタを介して伝達し合い、相互注視が成立した場合に音声通話を開始するシステムを提案する（図 1, 2）。特に、対面状況においてオフィス内の隣接したデスクの人間同士が 1 対 1 の対話を始める場面を遠隔地間において再現することにより、テレワークにおいて 1 対 1 の対話の生起を支援することを目指す。

2 関連研究

本章では、まず人間同士のインタラクションにおける対話の開始に関する社会学的知見を説明する。その後、遠隔コミュニケーションにおける応答の可否のみを伝達するアウェアネス支援の関連研究、お

よび非言語情報を用いて細かな相手の状態を伝達するアウェアネス支援の関連研究を述べる。

2.1 対話の開始に関する社会学的知見

本研究では、遠隔地間において対話を開始するために、対面において対話が開始する際に起こる動きを再現する。そのために、人間同士のインタラクションにおける対話の開始場面であるイニシエーショントークに関する社会学的知見に着目した。

Salvadori[12] はオフィスにおいて人間が対話を開始する際の人と人とのインタラクションの観察を行った。人間同士が対話を開始する際、まず始めに相手の応答可否を知るために相手の状態を確認する。そして、自身が応答可能である場合、作業を中断する、体の向きを変えるなどの行為によりそれを示し、相手の注意を引こうとすることを明らかにした。また、それらの注意を引く行動に対し相手も応答可能である場合、視線を返す、体を向けるなどの行為により応答可能であることを示す。

他にも、対話が開始する際の非言語情報の役割を調査した研究がある。非言語情報とは、人間のコミュニケーションにおいて言葉の持つ意味情報以外の全ての情報のことを指す。Kendon[7] は人間同士が会話を開始する際に、互いの様子を伺いながら物理的な距離を縮め、最終的には相互注視を会話を開始する合図とすることを述べた。Heath ら [6] は人間同士が対話を開始する際の非言語情報の働きに焦点を当て、身振り、手振りや体の動きが対話の開始に関してどのような役割を果たすかを調査した。例として、医療相談の場面において、医師および患者間でのインタラクションが開始する際に、視線情報のやり取りが相手の行動を喚起および促進させ、インタラクションの開始および進行に影響を与えることを明らかにした。これらの研究によると、対話の開始には非言語情報は重要な役割を果たしており、特に対話の開始の合図として相互注視が行われる。

これらの知見によると、対話を開始する際に、まず相手の応答可否を観察する必要がある。そして話し手は相手の方を見る、体を動かす、作業を中断するなどの行為により相手の注意を引こうとする。話し手の注意を引こうとする行為に対し、受け手は応答可能であることを示すために、視線を返す、体を向けるなどの行為を行う。そして相互注視が行われることにより対話が開始する。

本研究では、視線情報をアウェアネスとして遠隔地間において伝達し合い、視線情報のやり取りによって対話の開始に至るインタラクションを再現することを目指す。また、強く応答を求める場合に限り、視線情報に加えて音声情報を伝達することにより、応答を求める強さの幅を広げる。

2.2 応答の可否のみを伝達するアウェアネス支援

遠隔地においてインフォーマルコミュニケーションのような何気ない対話を開始するためには、あらかじめ相手の状態を知る必要がある。そこで、アウェアネスの提示により相手の状態を伝える手法がある。アウェアネスの提示により、相手が応答可能かどうかを判断できるようになり、対話の生起を支援できる。Kuzuoka ら [8] は、人形を用いて、応答可能な状態であればユーザの方に人形を振り向かせ、不可能な状態であれば人形の背を向けることによりアウェアネスを提示するシステムを提案した。

このようなアウェアネス提示により、相手の状態を把握できるようになり、対話の生起を支援できる。しかし、2.1 節で述べたように、対面状況においてはこれらシステムのように応答可否を伝えるのみでなく、互いの状態をより細かく伝えあい、対話を開始するタイミングを調整している。すなわち、対面状況のように対話を開始するためには、非言語情報を含む、より細かな相手の状態を伝え合う必要がある。そのため本研究では、視線情報を伝え合うことによりアウェアネスを提示する。

2.3 非言語情報を用いて細かな相手の状態を伝達するアウェアネス支援

2.2 節で述べたように、対面状況のように対話を開始するためには、非言語情報を含むアウェアネスを提示する必要がある。

児玉ら [14] は、テレワークにおいてインフォーマルコミュニケーションを誘発するために在宅作業をする人の映像とメインオフィスの映像を伝達し合うシステムを提案した。このシステムでは、メインオフィスにおいてディスプレイの前を人が通ると、その人の方向にディスプレイの向きを変えることにより注意を引き、相互注視を行い通話を接続する。Roussel ら [11] は常時接続型のビデオメディアスペースを配置することによりアウェアネスを提示し続けるシステムを提案した。Dou ら [3] は部屋全体を映すことによりアイコンタクトの問題を解消するインフォーマルテレプレゼンスシステムを提案した。

これらのシステムは相手の応答可否だけでなく、対面状態において対話が開始する際に必要となる細かなアウェアネスを提示することを可能にする。しかし、ビデオを用いてアウェアネスを提示する必要があり、プライバシーの問題および使用場面の制限といった課題がある。これに対し本研究では、物理アバタを用いて視線情報を伝達することにより、ビデオを用いずに応答可否および細かなアウェアネスを提示する。

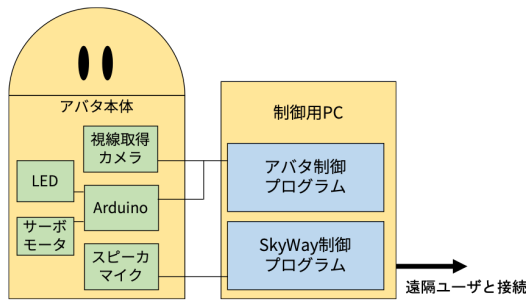


図 3: システム構成

3 システム設計

本章では、まず本研究の設計指針およびシステム構成を説明する。その後、本システムにおけるインタラクションデザインを説明する。

3.1 設計指針

本研究の設計指針として、2.1 節において述べたイニシエーショントークに関する知見に基づき、対面状況において対話を開始する際に交わされる動きを遠隔地間において再現する。特に本研究では、対話の開始に至る視線のやり取りに着目し、遠隔地間にそれぞれ一台ずつ配置し、アウェアネスとして視線情報を伝達し合う。まず視線情報によって、話し手は相手の注意を引く行動としてアバタの方に視線を送る。ここで、話し手は自分の話したい度合いに応じてアバタへの視線の送り方を変える。例として、強く応答を求める場合には注視を行うが、弱く応答を求める場合は短くアバタの方に視線を送る程度の動きを行う。また、最も強く応答を求める場合には、注視と同時にアバタに音声による呼びかけを行い、強く応答を求めていることを示す。そして遠隔地に配置したアバタによりこれらの行動を再現する。それに対し受け手は、応答可能であれば視線を返し、応答不可能であれば視線を返さないことにより応答可否を示す。そして双方がアバタの方に視線を送り、アバタを介した相互注視が行われた場合、そのまま音声通話を接続する。以上の流れに従い、「話したい度合いおよび応答可否に応じて視線および音声情報のやり取りを行い、物理アバタを介して相互注視を検出した場合、音声通話を行う」といった設計指針を設定した。本設計により、話し手は受け手に応答可否の判断を委ねることができ、話しかけることへの心理的負担が低下が見込まれる。

3.2 ハードウェア設計

本研究のシステム構成を図 3 に示す。視線情報を伝達する物理アバタを制御用 PC に接続し、データの入出力を行う。

利用者の視線検出には、OMRON 社の HVC-P2 を用いた [2]。HVC-P2 は基板に接続されたカメラ

を用いて、顔、視線の検出および角度推定等の様々な機能をデバイス単体により実行できる。HVC-P2 は制御用 PC によって制御され、検出された顔方向および視線方向の値のみを数値情報として物理アバタへ送信する。

視線情報の提示には二つのサーボモータおよびフルカラー LED を用い、Arduino UNO により制御を行った。サーボモータにより、pan-tilt 各 1 自由度の動作が可能である。

また、アバタ本体を介してそのまま音声通話を行うために、小型スピーカを搭載している。小型スピーカには MAG-LAB 社の VS-E100 を用いた [9]。

アバタ筐体は 3D プリントによって製作し、全高は約 180 mm とした。

3.3 ソフトウェア構成

本研究ではアバタを制御するために制御用 PC を用いる。制御用 PC では、音声通話、遠隔地間におけるアウェアネスの送受信を行うための SkyWay 制御プログラムおよびアバタ制御用のプログラムの二つのプログラムを使用している。

SkyWay 制御プログラムでは、音声通話を実現するために WebRTC のプラットフォームである SkyWay を使用している [1]。SkyWay では、音声通話のみでなく、遠隔地間におけるデータの送受信を行うことが可能であり、視線情報の送受信を行っている。また、アバタが取得したユーザの視線情報を遠隔地のアバタによって表現するためにはあらかじめ遠隔地間のアバタ同士を紐付けておく必要がある。本システムでは、SKyWay で用いられる、相手を一意に識別する PeerID により接続先を定めている。そしてアバタ間の紐付けおよびデータの送受信のために、視線検出を開始する前に Web インタフェース上においてあらかじめ相手の PeerID を指定している。

アバタ制御プログラムでは、HVC-P2 の制御および Arduino とのシリアル通信を行っている。アバタ制御プログラムにより、視線情報の取得および表現を行うことが可能である。

これら二つのプログラム間の通信により、視線情報を取得して遠隔地に送信し、送信された値に応じてアバタの視線提示を制御することが可能である。なおこれらのプログラム間の通信には WebSocket 通信を使用している。

3.4 インタラクションデザイン

続いて本研究におけるインタラクションデザインを説明する。本研究では、対話の開始に関するアウェアネスを細かく提示し合うため、互いの視線情報を相互に伝達する。なお今回実装したシステムでは、視線方向検出では精度が不安定になり、信頼性が不十分であったため、ユーザの視線情報を検出するた

表 1: 話し手の行動に応じた話し手の意図

話し手の行動	話し手の意図
一瞬のみ視線を送る	相手の様子を知りたい（積極的に話しかけようとはしない）
短く視線を送る	相手を取り込み中でなければ話しかけたい
注視する	相手を取り込み中である場合、もし相手が手を離せるタイミングであれば話しかけたい
注視する＋呼びかけ	相手を取り込み中である場合、相手に作業を中断させてでも話しかけたい



図 4: インタラクションの流れ

めに顔方向認識機能を用いた。顔方向の角度について、アバタ正面を0度として ± 15 度以内である場合、ユーザがシステムの方に視線を送っていると判断した。

具体的なインタラクションの内容として、まず話し手は視線の送り方により、どの程度相手に強く応答を求めているかを伝達し、さらに応答を強く求める際には視線を送ると同時に呼びかけを行う。例として、話し手が相手を取り込み中でなければ話しかけたい場合、アバタの方に短く視線を送る。これに対し、相手を取り込み中である場合に、もし相手が手を離せるタイミングであれば話しかけたい場合、アバタの方を注視する。さらに、相手を取り込み中であり、相手の作業を中断させてでも話したいような場合、アバタの方を注視し、同時に呼びかけを行う。これらのような視線の送り方を、表1にまとめる。

これらの話し手の視線の送り方に対し、受け手は視線を返すか否かにより応答可否を示す。その際、受け手はアバタの提示する情報から話し手の話したい度合いを認識することができ、その度合いに応じ

て応答の判断ができる。

図4にこれらのインタラクションの流れをまとめる。まず話し手は自分の話したい度合いに応じてアバタの方に視線を送ることにより相手の注意を引く(図4a)。これに対し受け手は応答可否を示すために、話し手の話したい度合いを考慮し応答可否を判断し、応答可能であればアバタに視線を返し、そうでなければ視線を返さない(図4b, 4c)。そして受け手が視線を返しアバタを介して相互注視が行われた場合、対話に移行する(図4d, 4e)。

本システムでは、相手の視線情報に応じて目の色を変化させる(図5)。相手のみが自分の方を向いているときには目の色を黄色に点灯させ、相互注視が成立した場合、緑に変化させる。それ以外の場合、目の色は赤色にしておく。また、話し手が強く応答を求めるために呼びかけを行った場合、点滅させることによりそれを示す。視線方向の制御については、システム方向を0度とし、相手の視線方向がアバタ正面を0度として ± 15 度以内であればアバタの視



図 5: 視線情報による目の色の变化

線を 0 度の方向に向け、それ以外においてはシステム本体から 30 度ごとにそれぞれ ± 15 度以内の場合、その角度を向く（15 度～45 度の場合、30 度を向き、45 度～60 度の場合、60 度を向く等）。

4 議論

本システムでは視線情報によるアウェアネスを提示することにより、相手に応答の判断を委ねることができる。今後は本システムにより話しかける行為の心理的負担が低下するかどうかを評価するための実験を行う予定である。

また、現在本システムは遠隔地間における 1 対 1 での対話の生起支援を目的としており、複数人での対話の生起支援には対応できていない。一方、対面状況においては、複数人により対話を開始する場面や、既に対話が行われているところに新たに他の人が加わり複数人により対話を行う場面がある。また、Kendon[7]によれば、複数人により対話をする際には対話の参加者は互いに体を向けあい、互いの間に一定の空間を維持しようとする。そこで将来的には、遠隔地間において複数台ずつのアバタを配置し、対話を行っている人同士のアバタの体を向け合うことにより複数人での対話の生起を支援することを検討している。

5 終わりに

アウェアネスとして視線情報を提示することにより遠隔地間において対話の生起を支援するシステムを提案した。今後は評価実験を行い、本システムが実際に対話の生起を支援することができるかどうかを調査する。

参考文献

- [1] N. C. Corporation. SkyWay. <https://webrtc.ecl.ntt.com/>.
- [2] O. Corporation. ヒューマンビジョンコンポ (HVC) シリーズ. <https://plus-sensing.omron.co.jp/product/hvc-p2.html>.
- [3] M. Dou, Y. Shi, J.-M. Frahm, H. Fuchs, B. Mauchly, and M. Marathe. Room-sized informal telepresence system. In *2012 IEEE Virtual Reality Workshops (VRW)*, pp. 15–18. IEEE, 2012.

- [4] P. Dourish and S. Bly. Portholes: Supporting awareness in a distributed work group. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 541–547, 1992.
- [5] C. Egido. Video Conferencing as a Technology to Support Group Work: A Review of Its Failures. In *Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work, CSCW '88*, p. 13–24, New York, NY, USA, 1988. Association for Computing Machinery.
- [6] C. Heath and H. Christian. *Body movement and speech in medical interaction*. Cambridge University Press, 1986.
- [7] A. Kendon. *Conducting interaction: Patterns of behavior in focused encounters*, Vol. 7. CUP Archive, 1990.
- [8] H. Kuzuoka and S. Greenberg. Mediating awareness and communication through digital but physical surrogates. In *CHI'99 Extended Abstracts on Human Factors in Computing Systems*, pp. 11–12, 1999.
- [9] MAG-LAB. VS-E100. <https://www.amazon.co.jp/MAG-LAB-mimo-speaker->
- [10] Microsoft. Microsoft Teams. <https://www.microsoft.com/ja-jp/microsoft-365/microsoft-teams/free>.
- [11] N. Roussel, H. Evans, and H. Hansen. MirrorSpace: using proximity as an interface to video-mediated communication. In *International Conference on Pervasive Computing*, pp. 345–350. Springer, 2004.
- [12] F. A. Salvadori. Open office interaction : initiating talk at work(Doctoral dissertation). In *King's College London*, 2016.
- [13] I. Zoom Video Communications. Zoom. <https://zoom.us/>.
- [14] 児玉裕輝, 葛岡英明, 徐建鋒, 明堂絵美, 原田悦子, 大澤博隆. 雑談を誘発するテレプレゼンスロボットシステム. 研究報告グループウェアとネットワークサービス (GN), 2018(20):1–7, 2018.