

アバター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

石井 亮*¹ 宮島 俊光*¹ 藤田 欣也*¹

Avatar's Gaze Control to Facilitate Conversation in Virtual-Space Multi-User Voice Chat System

Ryo Ishii*¹ Toshimitsu Miyajima*¹ and Kinya Fujita*¹

Abstract – In order to facilitate multi-party conversations in a shared-virtual-space voice chat environment, we propose a shared-virtual-space voice chat system with an automatic avatar's gaze direction control function using user utterance information, based on an avatar's gaze behavior model of turn-taking in multi-party conversation. In our gaze behavior model, a conversation was divided into three states; during-utterance, right-after-utterance, and silence. The avatar's gaze behaviors were controlled based on a probabilistic state transition model in each state. The developed system attained easy-to-use automatic gaze control without eye-tracking camera or manual operation.

Keywords : Gaze, Avatar, Voice Chat, Multi-Party Conversation, Utterance Information

1. はじめに

身体情報を含むノンバーバル情報は、対面会話において重要な役割を担っている^[1]。実写型ビデオ会議システムは、会話パートナーの実写画像を用いることで多くのノンバーバル情報を伝えることが可能であるが、意図しない個人情報を伝送することが懸念される。この個人情報伝送の可能性は、実写画像ではなく3次元CGアバタと仮想空間を用いることによって、回避可能である。しかし、3次元CGアバタは表情や視線、身振りなどの身体動作を中心としたノンバーバル情報が欠如するため、音声を用いて頷き動作を制御する^[2]など、ノンバーバル情報の補完が、仮想空間での円滑なコミュニケーションの促進に重要になるものと考えられる。

ノンバーバル情報の中でも、会話中の視線は会話の開始合図や発話権の授受、発話の促進などの重要な機能を有しており、仮想空間コミュニケーションシステムにおいても、適切なアバタの視線制御の実現が必要と考えられる。アバタの視線を制御する方法には、ユーザの視線を計測してアバタの視線に反映させるものと、装置を用いず人間の視線行動モデルに基づいた自律的に視線を制御する方法の二つが提案されている。ユーザの視線を直接アバタに反映させるシステムは、ユーザの視線を赤外線LEDとCCDカメラを用いて検出しアバタに反映させるGAZE^[3]などがあるが、アイトラッキングが前提となるため、利用環境が限定され

る。視線行動モデルに基づいた制御法は、二者対面会話における凝視・非凝視時間を2次多項式で表現し、凝視位置をランダムに決定して凝視動作をさせるものや^[4]、信念ネットワークにより、各時刻における直前の話者・非話者の凝視、動作、発話状態から、話者の次の凝視が発生する確率を推定する方法などが提案されている^[5]が、これらは、おもに擬人化エージェントの発話にあわせて視線行動を自律的に生成することを主目的としており、ユーザが発する発話に基づいて視線行動を生成するものではなかった。

これに対し、筆者らは複数人同時会話が可能なる共有仮想空間マルチパーティ音声チャットシステムを開発し、円滑なコミュニケーションの促進を目的に、ユーザの音声情報を利用したアバタの注視対象制御を試み、その有効性を実験的に確認した^[6]。このシステムでは、大掛かりな装置やユーザの直接入力が必要としないアバタの視線制御法として、各ユーザの発話情報のみを用いて注視対象者を算出する制御モデルを提案した。しかし、この制御法は持続的に注視対象者を注視するため、挙動が不自然であることと、ユーザに威圧感を与える可能性があり、適度な視線はずしが必要であった。また、発話時のみを考慮した制御であり、沈黙時の視線制御は考慮されていなかった。さらに、本来、人間の視線には、発話の委譲と要求を行う調整子機能がある^[7]が、他の研究を含め、音声情報を用いたアバタの視線制御において、発話権制御機能は考慮されていなかった。

そこで本研究では、共有仮想空間マルチパーティ音声チャットシステムにおける会話の促進を目的に、人間の視線行動モデルに基づき、発話情報を用いてアバ

*1: 東京農工大学大学院

*1: Graduate School of Tokyo University of Agriculture and Technology

タの注視制御を行うシステムを開発した。本稿では、開発したシステムの視線制御アルゴリズムならびに評価実験の結果に関して報告する。

2. システム設計

2.1 マルチパーティ音声チャットシステム

図1に、共有仮想空間マルチパーティ音声チャットシステムの構成の概略を示す。本システムは、サーバ/クライアント型の構成となっており、各ユーザの仮想空間での位置がサーバを介してクライアント間で共有され、各クライアントではそれぞれのユーザ位置にアバタが描画される。また、サーバはユーザの位置から会話可能な範囲にいる他のユーザを検出し、会話グループを形成する。各クライアントはサーバから通知されるグループ内の他クライアントに対して、P2P方式で音声を伝送すると同時に、音圧等の音声情報を伝送する^[6]。

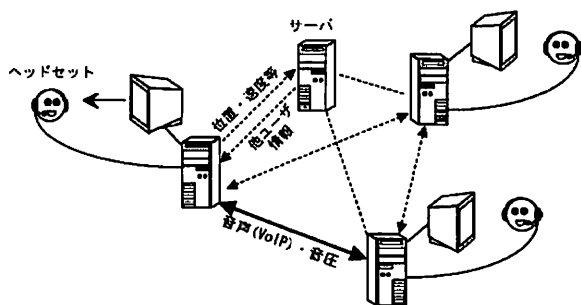


図1 システム構成図
Fig.1 System configuration

本研究ではこのシステムをもとに、伝送された各ユーザの音圧情報から発話の有無を検出し、発話情報を用いて、各クライアントにおける他ユーザのアバタの視線を注視行動モデルに基づいて制御する方式を提案する。発話検出に際しては、125ms周期で算出した発話音圧を各ユーザごとの最大音圧で正規化し、設定した閾値を越えた時を発話と判定した。発話終了判定条件は、日本語地図課題対話コーパスにおいて、400ms以上の無音区間によって発話単位を定義した先行研究^[6]を参考に、実験的に無音状態500msとした。図2に、3人で使用したときにユーザ側で表示される実行画面と、システム使用風景の例を示す。

2.2 従来の音声を用いた注視制御とその問題点

コミュニケーションにおいて、ノンバーバル情報は大きな役割を果たすため、リアルタイムに人間とコミュニケーションを行う擬人化エージェントやアバタチャットシステムにおいても、CGキャラクタがノンバーバル情報を表出することが求められる。ここで、一般に、表出されるノンバーバル情報は言語情報と相関すると期待されるため、ユーザの音声情報を用いて

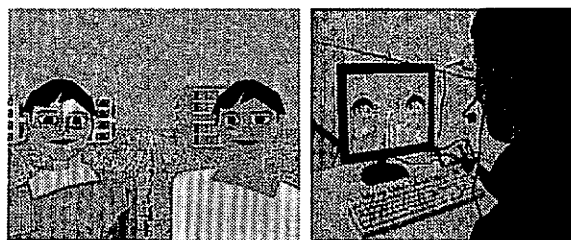


図2 実行画面(左)と使用風景(右)の例
Fig.2 Examples of the screenshot of the virtual space (left) and the snapshot of the user using the developed system (right).

ノンバーバル行動を制御する試みなどが行われている^[2]。

特に視線に関しては、擬人化エージェントの注視行動を制御する試みとして、ユーザの発話の非流暢さ(Disfluency)を検出して注視行動制御を行う、Listening Agentが開発されている。このシステムでは、カメラを用いてユーザの動作を計測することで、視線以外の身体動作も提示可能とした^[9]。しかし、このシステムはユーザとエージェント間の調和感覚の実現を目指すもので、発話情報に対して非常に簡単な視線フィードバックを行うのみで、注視対象の制御は行っていない。

また、実空間での多人数会話では、受話者は発話者に対して、話者は直前に発話していた受話者に対して視線を向ける傾向が強い。われわれは、これらの行動をモデル化し、各ユーザに対して、過去60秒間の音圧の重み付き積分として算出される話者注視効果と、発話開始から5秒間で消失する発言開始効果の2つからなる注目度(AP, Appeal Point)を定義し、最もAPが大きいユーザアバタに対して視線を向ける注視対象者制御法を考案し、共有仮想空間アバタ音声チャットシステムにおいて、その有効性を実験的に示した^[6]。しかし、この制御法では、1) 常に注視対象者を注視するため、挙動が不自然であり、ユーザに威圧感を与える可能性があった。さらに、2) 発話量のみを用いた制御であるため、沈黙時の視線制御機能は有していなかった。また、3) 複数人会話には会話権(floor)が存在し、視線行動を用いた発話促進など、会話権の授受が複雑に行われているが、この点も考慮されていなかった。

2.3 提案するアバタ注視制御モデル

実際の会話における人の視線行動は、発話中と無言状態で大きく異なる。発話中には、発話者は受話者に、受話者は発話者に視線を多く向ける。無言状態では、発話中に比較して他者を注視する時間が短く、周囲の様子を伺う。そこで本研究では、誰かが発話している

「発話中状態」と、誰も発話をしていない「無言状態」に会話を分割した。さらに、複数人会話では会話権が存在し、特に発話終了の前後に視線を用いて会話権を委譲する^[10]。そこで、視線を利用してユーザに発言を促して会話を活性化するための制御を導入するために、発話中状態から無言状態に移行する前に「発話終了直後状態」を定義し、発話終了直後状態を経て無言状態に移移するようにした。

図3に、本研究で用いた「発話中状態」、「無言状態」、「発話終了直後状態」の3状態からなる状態遷移モデルを示す。状態遷移の制御には、発話検出と発話終了判定を用いた。また、発話終了直後状態から無言状態への遷移は一定時間経過で行う。以下に、各状態での制御方法について詳しく述べる。

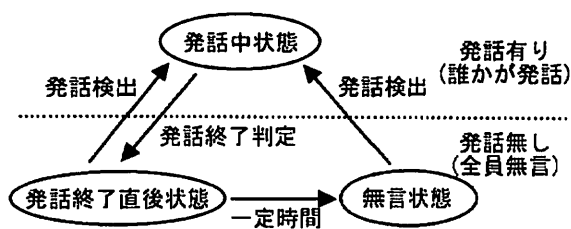


図3 会話の状態遷移モデル
Fig.3 State transition model of the conversation.

2.3.1 発話中状態制御

発話中状態においては、先行研究の発話者注視アルゴリズム^[6]を用いてAPを算出し、APが最大となるユーザを最優先注視対象者とし、さらにモデルに基づく各種の制御を行った。APのみを用いた擬似視線制御は、特定のユーザが長時間連続して発話すると、話者アバタを他のアバタが持続的に注視し、適度に話者以外を見る自然な挙動が欠如するという問題を有する。そこで本研究では、最優先注視対象者以外の「他者を見る」状態と、全アバタ以外の「外を向く」状態を定義し、低確率でこれらの挙動を行わせた。また、話者を注視する際に、視線交差の長時間の持続は威圧感を与える可能性がある^{[1], [12]}。われわれは、共有仮想空間における会話促進を目的としており、話者への威圧感には会話促進を妨げる恐れがあることから、過度に持続的な凝視は避ける必要がある。そのため、適度に視線をそらせる機能が求められる。他方、注視量が増加すると注視された話者はより多く発言する傾向があるため^[11]、発話促進の観点からは、完全に視線をそらせることは回避し、威圧感を与えない程度に話者に注視をすることが望まれる。そこで、最優先注視対象者に対して単純に視線交差させるだけでなく、ユーザが視線交差とは認知せず、かつ、注視されていると認知するあいまいな注視位置を実験的に決定し、この位置

を見る注視行動を、「あいまい注視」と定義した。このあいまい注視と視線交差を、適度に反復する視線行動を行うことで、注視効果を損なわずに威圧感の軽減を期待する。なお、擬人化エージェントの非凝視時の視線位置(目への凝視以外に見る位置)が顔よりも低い位置であったとき、「強い」、「自信のある」という印象評定値が最も低くなったという報告^[12]から、あいまい注視位置は、調整法を用いた予備実験を被験者5名に対して行い、視線交差位置から5°下方とした。実験では、ディスプレイから80cmの位置に被験者を着座させ、アバタの虹彩位置をキーボードを用いて上下方向に操作させて、アバタに視線交差されたと認知する位置および下方向に視線交差が外れたと認知する位置をそれぞれ3回ずつ回答させ、平均値を求めた。図4に、視線交差とあいまい注視を行うアバタ、ならびに両者の眼球周辺の拡大図を示す。

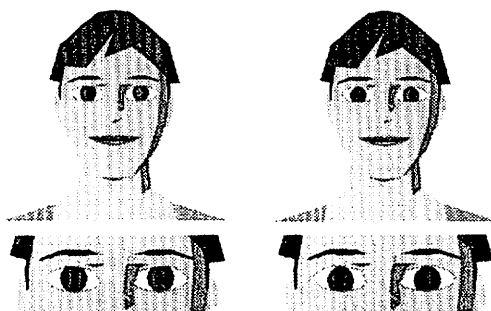


図4 視線交差(左)、あいまい注視(右)
Fig.4 Eye contact (left) and vague gaze (right).

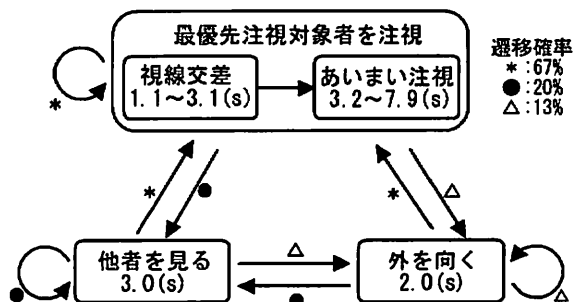


図5 発話中状態制御の状態遷移図
Fig.5 State transition model of during-utterance state.

図5に、発話中状態のAPによる最優先注視対象者選定後の制御モデルを示す。各状態はそれぞれ持続時間経過後に確率モデルに従って次の状態に移移する。なお、最優先注視対象者を注視する際は、まず視線交差することが自然であるため、必ず最初に視線交差を行い、適度な時間が経過してからあいまい注視状態へ遷移させた。また、最優先注視対象者が変更された場

合には、その直後に必ず、最優先注視対象者を注視する状態に移るものとした。各継続時間・確率のパラメータは文献[4], [11]を参考に、著者および共著者が動作確認を行って特に強い違和感がないことを確認し決定した。

2.3.2 無言状態制御

実際の複数人会話において、発話が終了し無言状態が続くと、人は視線を下げ、ときどき誰かが話し始めないか様子を伺うように視線を他者に向ける。われわれは、無言状態において、このような自然な挙動を行う「観察モード」と、注視によるユーザの発話の促進効果を期待する「期待モード」の2状態を定義した。

図6のように、無言状態制御は観察モードから開始し、「誰かを見る」、「外を見る」の2状態を、確率にしたがって遷移する。このとき、「誰かを見る」状態は、過去の発話量が多いユーザを注視することが自然と考えられるため、発話量を正規化した確率モデルを用いた。また、期待モードは、ユーザの発話意欲を促進するため、ユーザに視線交差とあいまい注視を交互に行う。各状態の継続時間と観察モードでの状態遷移確率は、先行研究を参考に、著者らが簡単な予備実験を行って違和感がないことを確認し決定した。

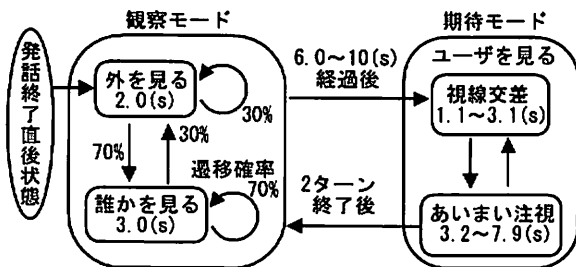


図6 無言状態制御の状態遷移図
Fig. 6 State transition model of silence state.

2.3.3 発話終了直後状態制御

発話終了前後の話者交替における視線行動に関しては、様々な報告がなされている。特に、3者会話において、話者交替時には発話終了前後1.0秒(終了前0.6秒、後0.4秒の間)の区間で以下のような視線行動の報告がある[10]。

- i) 現話者は次話者を見続けている。
 - ii) 非話者は現話者から次話者に視線を移す。
- また、話者交替が行われないときには、
- iii) 非話者は現話者を見続ける。

本研究では、このような発話権移譲を意味する視線行動を発話終了直後にアバタに行わせることで、ユーザの発話権取得が促進され、会話が活性化されることを期待し、以下のような制御を考案した。

図7に示すように、当該端末利用ユーザが最終発話者である場合、上記のiii)にしたがって、発話終了後直

ちにユーザに向けて視線交差させた。他方、他のユーザが最終発話者であった場合には、視線計算対象アバタが最終発話か受話者かによって制御方法が異なる。最終発話者の場合には、i)から、最終発話者アバタは、直ちにユーザに視線交差させた。最終発話者の視線交差は会話権の委譲を意味するため、ユーザの発話意欲を促進するものと期待される。受話者の場合には、ii)にしたがって、最終発話者を注視した後にユーザを注視させた。すなわち、受話者アバタは、最終発話者アバタよりも遅れてユーザに視線交差させた。

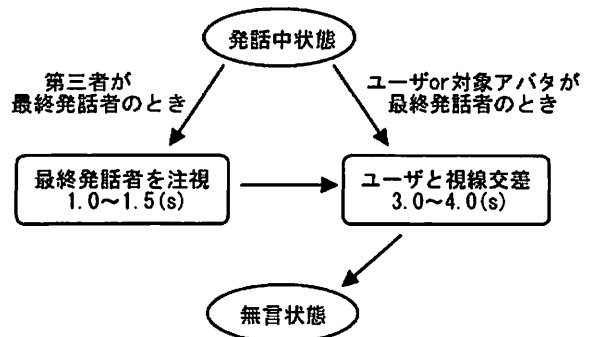


図7 発話終了直後状態制御の状態遷移図
Fig. 7 State transition model of right-after-utterance state.

3. 評価実験

3.1 発話中状態制御の評価

発話中状態制御における、あいまい注視、視線はずしの効果を検証し、発話中制御モデルの有用性を示すことを目的に、リッカート法による5段階の主観評価を行った。被験者には、両端に反対語(感じる/感じない、思う/思わない)を記した線分を用いて、直接、数値で回答させた。評価項目は、表1のように、アバタの挙動の自然さ、アバタからの威圧感のなさ(非威圧感)、アバタに発話を期待されていると感じたか、アバタに話しかけたいと思ったか(発話意欲)を問う4項目とした。被験者は24名とし、3名ずつの8組に分け、図2のように仮想環境内の相手アバタとそれぞれ対面するように配置して、実験者から指示されたテーマに従って5分間会話する課題を課した。会話内容は、あらかじめ被験者から会話可能な議題を聞き取り調査し、共通する内容に関して会話をさせた。なお、実験中は、相手アバタを注視するように指示した。主観評価は、設定した実験条件ごとに、それぞれ会話直後に質問紙を用いて実施した。また順序効果を低減するために各グループで条件の試行順序を変えた。表2に実験条件を示す。

主観評価値の平均値を図8に示す。全体に、視線はずしやあいまい注視を行わない条件に比較して、これ

表 1 主観評価に用いた設問

Table 1 Question sentences used for subjective evaluation.

設問 1	アバタの挙動は自然でしたか
設問 2	アバタから威圧感を感じましたか
設問 3	アバタが発言を期待していると感じましたか
設問 4	アバタに話しかけたいと思いましたか

表 2 発話中状態制御の実験条件

Table 2 Experimental conditions for comparing during-utterance state control algorithms.

条件 A	話者注視制御のみ
条件 B	視線はずし有り
条件 C	あいまい注視有り
条件 D	視線はずし・あいまい注視有り

らを行う条件の評価が高く、さらに、両方の注視制御をおこなう条件で最も高い評価値を示した。Friedman 検定の後に Scheffe の方法で多重比較おこなった結果、自然さでは視線はずしやあいまい注視をおこなう全条件、非威圧感ではあいまい注視をおこなう 2 条件が、話者注視制御のみの条件に対し有意に高かった。また、話者から視線をはずす行為は、本来は、興味の対象が他にあることを意味するため、視線はずしはユーザの発話意欲を損なう可能性が懸念されたが、持続的注視よりも若干高い値を示した。これは、設定した遷移確率が 20% と低かったことに加え、自然さの向上が発話意欲につながった可能性が考えられる。あいまい注視と視線はずしの両方を行う条件で発話意欲が有意に高かったのは、上記に加えて、あいまい注視位置から視線交差への動きによって視線交差が強調され、発言への期待感が高くなった可能性が考えられる。視線はずしとあいまい注視を組み合わせた条件は、話者注視制御のみに比較してすべての評価値が有意に高く、提案する発話中状態制御の有効性が確認された。

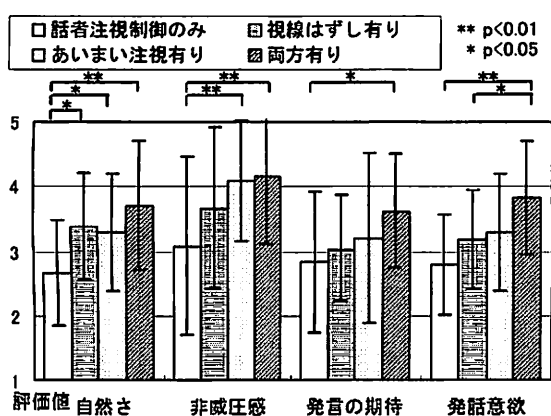


図 8 リッカート法を用いた発話中状態制御の主観評価結果

Fig. 8 Subjective evaluation of during-utterance state control algorithms.

3.2 無言状態制御の評価

次に、無言状態制御における、観察・期待モードの効果を検証し、無言状態制御モデルの有用性を示すことを目的に、発話中状態の実験と同様の主観評価を行った。評価項目と実験方法および検定方法は、先の実験と同一とし、先の実験とは異なる被験者 24 名で行った。表 3 に実験条件を示す。3 条件とも、発話中状態制御を動作させる条件のもとで、無言状態における視線制御方法を変更して実験を実施した。

表 3 無言状態制御の実験条件

Table 3 Experimental conditions for comparing silent state control algorithms.

条件 A	無言状態制御無し
条件 B	観察モード有り
条件 C	観察・期待モード有り

図 9 に主観評価値の平均値を示す。無言状態制御が無い条件に対し、観察モード有りの条件は、非威圧感、発話意欲の評価値が有意に高く、自然さも若干高い値を示した。ユーザや他のアバタ以外にも視線を移すことによって威圧感が軽減され、間接的に発話意欲が促進された可能性が考えられる。観察モードに加えて期待モードを有する条件は、発言の期待感をはじめ、全ての評価値が無言状態制御無しに対して有意に高い結果となった。期待モードは低頻度でユーザを注視する制御方式であることから、すなわち、無言状態における過度にならない程度のユーザへの注視によって、自然な印象を損なうことなくユーザの発話意欲を促進できる可能性が示唆された。

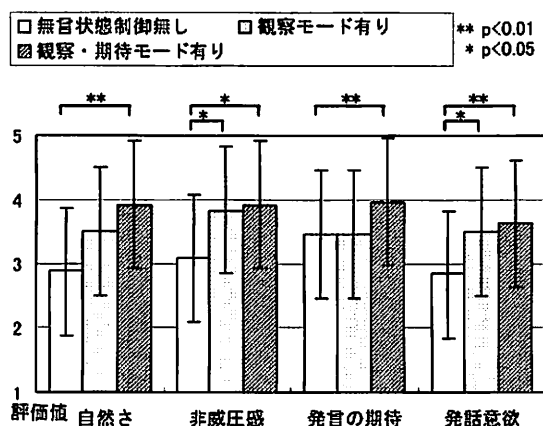


図 9 リッカート法を用いた無言状態制御の主観評価結果

Fig. 9 Subjective evaluation of silent state control algorithms.

3.3 発話終了直後状態制御の評価

発話終了直後状態制御における、ユーザへの注視効果を検証し、発話終了直後状態制御モデルの有用性を示すことを目的に、先の 2 つの実験と同様のリッカ

ト法を用いた主観評価を、先の2実験と異なる被験者15名に対して行った。発話終了直後制御では、会話パートナーの意思に拘わらずユーザへの注視を行う。しかし、この制御は会話パートナーの会話権委譲の意志に基づいてアバタが注視している印象を、ユーザに擬似的に与えることを期待するものである。そこで、その効果を問う設問「アバタは本人の意思を反映していると感じましたか」を追加した。表4に実験条件を示す。2条件とも、発話中状態制御および無言状態制御を動作させる条件のもとで、発話終了直後制御の有無を切り替えて実験を実施した。発話終了直後制御無しの場合（条件A）では、発話終了を検出した段階で無言状態制御を開始した。

表4 発話終了直後状態制御の実験条件
Table 4 Experimental conditions for comparing right-after-utterance state control algorithms.

条件 A	発話中・無言状態制御のみ
条件 B	発話終了直後状態制御有り

得られた主観評価値の平均値を図10に示す。発話中・無言状態制御のみの条件に対し、発話終了直後制御有りの条件は、発言の期待感に関する評価値がWilcoxonの符号順位検定で有意に高く、発話意欲の評価値も高い値を示したことから、期待した発話促進効果を示唆する結果となった。また、ユーザへの注視は発話意欲促進効果が期待できる反面、威圧感を与える可能性が懸念されたが、有意な威圧感の増強は確認されなかった。さらに、興味深い結果として、発話直後の注視によって意思反映感が有意に増強されるという結果が見られた。本研究で実装した制御は、会話パートナーが実際にユーザのアバタを注視しているか否かに拘わらずアバタにユーザを注視させるものであることから、自律的に動作する擬似制御であっても、人間が行う視線行動と類似した視線行動をアバタに行わせれば、ユーザは会話パートナーの意図を反映しているかのごとく感じることを示唆するものと言える。すなわち、アバタ音声チャットシステムにおいて、発話終了直後のユーザに対する注視行動は、ユーザに対する自然で効果的な発話促進法となる可能性が示された。

4. 考察

評価実験において、発話中制御ではAPのみを用いた持続的発話者注視アルゴリズムと比較して、視線はずしやあいまい注視を行う制御の主観評価値が、全項目で高くなる傾向を示した。より人間らしい自然な挙動である視線はずしは、自然さの向上に有効であった。また、注視効果を損なわずに威圧感を軽減するために導入した、ユーザの目の少し下を見るあいまい注視は、

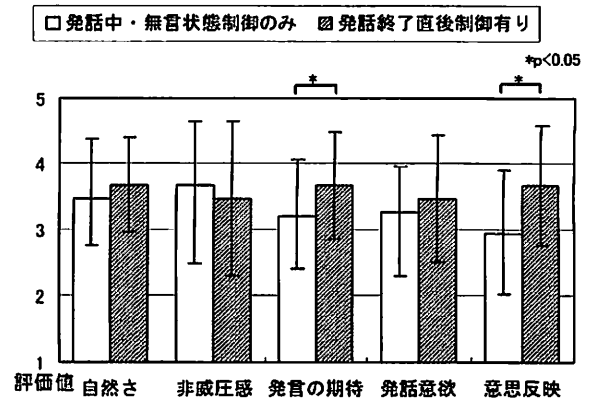


図10 リッカート法を用いた発話終了直後状態制御の主観評価結果
Fig.10 Subjective evaluation of right-after-utterance state control algorithms.

期待どおり威圧感軽減に有効であった。威圧感の軽減に関する今回の実験結果は、目への注視量と支配的因子が正の相関を持つという報告とも合致する[12]。また、軽微な視線はずしであるあいまい注視は、視線をはずすことによる発話促進への負の効果が懸念されたが、むしろ発言の期待感や発話意欲の促進効果が見られた。あいまい注視位置から視線交差位置への動きが生じることによって、視線交差が強調されたためと考えられる。これは、ユーザは他者に注視されることで発話促進されるという先行研究の知見[11]とも合致する。ユーザの目の少し下を見るあいまい注視は実装が容易であるため、アバタチャットシステムにおける有効な威圧感軽減法であり、また適度にユーザと視線交差させることで発話促進法としても利用できる可能性がある。

無言状態制御の評価実験では、APが高いアバタ以外にもランダムに視線を向ける観察モードにおいて、非威圧感の評価値が有意に高くなり、自然さも高くなる傾向が見られた。さらに、発話意欲向上にも効果があった。低い頻度でユーザやアバタ以外に視線をむける行為が、自然な威圧感軽減、ひいては発話意欲促進につながったと考えられる。また、意図的にユーザに対してアバタが視線を向ける期待モードは、発言の期待感や自然さに効果が見られた。これは凝視量が増えることによって、発話促進[11]や印象評価値が向上する[13]といった心理学的知見とも合致する。以上の知見からは、本研究が目的とする視線を用いた発話促進による会話の盛り上げに加えて、特定のユーザの発話促進など、アバタの凝視頻度と量の適切な制御による、会話制御の可能性なども期待される。

発話終了直後状態の評価実験では、発話終了時にアバタがユーザに視線を向ける行為が、ユーザへの発話

の期待と会話パートナーの意思反映感の向上に効果があった。まず、ユーザへ発話の期待感を増強したことは、発話終了時に視線を向けることが会話権委譲シグナルとなるという、人間の会話分析結果^[10]と合致する。すなわち、人工物であるアバタが人間と同様な挙動を行うことで、同等のシグナルを与えたことになる。これに関しては、エージェントと人条件で視線に対する印象に類似性があるといった報告^[14]とも合致する。また、ユーザにアバタが視線を向ける行為が、会話パートナーの意思反映感の向上に繋がったことから、ユーザが持つ会話パートナーの動作モデルとアバタの動作が一致することで、アバタがパートナーの意思を反映しているとユーザに感じさせることができた、と考えられる。以上のことから、このようなアバタを通じた仮想世界でのコミュニケーション環境では、人はアバタを会話パートナーの存在のみを示す代理物では無く、それ以上のパートナー自身を反映した化身として認識している可能性も考えられる。

また、本研究のシステムでは、情報源として会話パートナーの音声を利用してアバタの視線を制御しているため、視線制御によってみかけのコミュニケーションチャンネルは増加するが、会話パートナーに関する情報量は変化しない。これに対して、主観評価実験は、擬似的にアバタの視線を制御しユーザに提示することは、ユーザの発話意欲促進に効果があるという興味深い結果となった。同様に、音声情報のみを用いて、ディスプレイ上のキャラクタである InterActor に頷きなどの身体動作・反応を生成したシステムが開発されており、引き込み効果や話しやすさ向上などコミュニケーション支援となる結果が示されている^[2]。これらの結果の解釈としては、音声情報に加えて視覚情報が付加されたことで、ユーザが受け取るコミュニケーションチャンネルが増え、補完的に機能した可能性が考えられる。すなわち、音声情報を利用した擬似的なコミュニケーションチャンネルの増加は、仮想空間コミュニケーションに有用である可能性が高いものと考えられる。

さらに、アバタの視線行動によってユーザの印象改善や発話意欲の促進が可能であったことから、カジュアルな会話を楽しむような用途においては、会話パートナーの状態を正確に反映することを目指すのではなく、むしろノンバーバル情報が持つ機能を積極的に利用したコミュニケーション促進などの利用形態が考えられる。このコミュニケーション促進の観点からは、本研究のようなアバタ音声チャットシステムだけでなく、会話エージェントにおけるノンバーバル情報提示にも適用可能であり、人間の行動モデルをそのまま反映させるのではなく、期待される効果に基づいてノンバーバル情報を生成するなどの方法が考えられる。

今後の課題としては、まず、先行研究を参考に実験者らが定めた、遷移確率や状態持続時間の最適化が挙げられる。さらに、より自然で円滑なコミュニケーションシステムの実現のために、表情や身体動作といった、視線以外のノンバーバル情報の提示を検討する必要がある。また、現在のアバタはユーザに応じて外観や動作を変更することが困難であるため、容易に使用できるアバタパーソナライズシステムが必要である。アバタのパーソナライズによって、ユーザの個人識別が容易になり、よりパートナーと話している感覚が増して、会話への引き込み効果も向上するものと期待される。

5. まとめ

共有仮想空間マルチパーティ音声チャットシステムにおける会話促進を目的に、視線行動モデルに基づいて、発話情報のみを用いてアバタの注視行動を制御するシステムを開発した。評価実験から、発話意欲促進などにおける提案手法の効果が確認された。自然で会話がはずむアバタ音声チャットシステムの実現に向け、視線以外のノンバーバル情報の制御や、簡便なアバタパーソナライズシステムなどが望まれる。

謝辞

本研究の一部は、文部科学省科学研究費補助金および特別教育研究費共生情報工学研究推進経費によるものである。ここに記して感謝する。また、本研究の結果に対して、多くの有益な議論をいただいた中野有紀子特任准教授に、この場を借りて感謝する。

参考文献

- [1] 松尾：コミュニケーションの心理学，ナカニシヤ出版（1999）
- [2] 渡辺，大久保，中茂，榎原：InterActor を用いた発話音声に基づく身体的インタラクションシステム，ヒューマンインタフェース学会論文誌，Vol.2, No.2, pp.21-29（2000）
- [3] Vertegaal, R. : The GAZE GroupWare System: Mediating Joint Attention in Multiparty Communication and Collaboration, ACM CHI, pp.294-301（1999）
- [4] Lee, S.P., Badler, J.B. and Badler, N.I. : Eyes Alive, ACM SIGGRAPH, pp.637-644（2002）
- [5] Bilvi, M. and Pelachaud, C. : Communicative and Statistical Eye Gaze Predictions, AAMAS,（2003）
- [6] 宮島，下地，藤田：視線と存在の擬似アウェアネス機能を有する共有仮想空間コミュニケーションシステム，日本バーチャルリアリティ学会論文誌，Vol.10, No.1, pp.71-80（2005）
- [7] Kendon, A. : Some Functions of Gaze-direction in social interaction, Acta Psychologica, Vol.26, pp.22-63（1967）
- [8] 堀内，中野，小磯，石崎，鈴木，岡田，仲，土屋，市川：日本語地図課題対話コーパスの設計と特徴，人工知能学会誌，Vol.14, No.2, pp.261-272（1999）

- [9] Maatman, R.M., Gratch, J. and Marsella, S. : Natural Behavior of a Listening Agent, IVA, pp.25-36 (2005)
- [10] 榎本, 伝 : 3人会話における参与役割の交替に関わる非言語行動の分析, 人工知能学会研究会資料, SIG-SLUD-A301, pp.25-30 (2003)
- [11] Vertegaal, R. and Ding, Y. : Explaining Effects of Eye Gaze on Mediated Group Conversations: Amount or Synchronization?, ACM CSCW, pp.41-48 (2002)
- [12] 深山, 大野, 武川, 澤木, 萩田 : 擬人化エージェントの印象操作のための視線制御方法, 情報処理学会論文誌, Vol.43, No.12, pp.3596-3606 (2002)
- [13] 磯, 木村, 桜木, 大坊 : 視線行動が印象形成に及ぼす影響 -3者間会話場面における非言語的行動の果たす役割-, 対人社会心理学研究, No.4, pp.83-91 (2004)
- [14] 深山, 澤木, 大野, 村瀬, 萩田 : ユーザとのコミュニケーションにおける擬人化エージェントの視線移動制御, 電子情報通信学会技術報告, HIP2000-46, pp.9-14 (2001)

(2007年3月9日受付, 7月27日再受付)

著者紹介

石井 亮



2006年東京農工大学工学部情報コミュニケーション工学科卒業。現在,東京農工大学大学院工学府情報工学専攻在学中。マルチユーザ音声会話時の視線行動のモデル化と共有仮想空間におけるアバタ視線制御に関する研究に従事。

宮島 俊光



1997年早稲田大学教育学部理学科数学専修卒業。現在,東京農工大学大学院工学教育部情報工学専攻技術職員として,バーチャルリアリティ技術を利用したコミュニケーション支援の研究に従事。

藤田 欣也 (正会員)



1988年慶應義塾大学大学院理工学研究科修了。相模工業大学,東北大学医学部,岩手大学を経て,現在東京農工大学大学院教授。遠隔共有仮想空間および力触覚や歩行感覚の提示,ならびに医用福祉工学に関する研究に従事。工学博士。