

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

宮島 俊光^{*1}, 下地 崇^{*1}, 藤田 欣也^{*1}

Shared Virtual Space Communication System with Pseudo Awareness of Presence and Gaze

Toshimitsu Miyajima ^{*1}, Takashi Shimoji ^{*1} and Kinya Fujita ^{*1}

Abstract - A shared virtual space communication system was developed, with multi-user walk-through function using locomotion interface, and voice communication function by Voice over IP technology. An algorithm for the automatic control of the group of the voice communication was implemented using spatial relationship of the users. Visual and auditory awareness about the presence of the other users in the conversational area were also attained. The automatic group control and awareness representation were effective for natural and smooth voice communication in shared virtual space. A pseudo gaze awareness system using each user's voice volume has also been developed. It was experimentally demonstrated that the pseudo gaze awareness is effective to enhance the conversation.

Keywords : Shared Virtual Space, Communication, Awareness, Voice over IP

1. はじめに

ネットワークの普及に伴い、テキストを用いたチャット型や TV 会議型など様々な遠隔コミュニケーションシステムが開発され、3D グラフィックシステムの高速化を背景に仮想空間を用いたコミュニケーションシステムの試みも見られるようになってきた[1] [2]。

テキストチャット型はネットワークトラフィックが少なくシステム負荷が軽いことから、広く利用されているが、入力容易さや、声の抑揚などノンバーバル情報の伝達に関しては、音声を用いた直接会話に劣る。

実写画像と音声を用いる TV 会議型は、表情や声の抑揚などのノンバーバル情報を伝達可能である点で優れているが、カメラが必要であることとネットワーク帯域を多く消費することに加え、実写画像を使用するため背景画像によって意図しない個人情報の伝送が生じる可能性が懸念される。またマルチユーザ型システムでは、平面的にユーザ画像を配置した場合、全員の視線方向が同一に見え、各ユーザの興味対象に関する視線アウェアネスが得られない(以下、アウェアネスを AW と表記)。そのため、誰が誰に向かって話しているのかわからず会話に混乱をきたしやすい。そこで、注視対象ユーザの特定を可能とするための視線 AW 制御機能を提供することが望まれている。

前述の意図しない個人情報の伝送に関する問題を解決する手段として、実写映像ではなく 3DCG を用い、さらにノンバーバル情報を伝えるために、カメラ映像を認識して顔の特徴点を抽出し、相手に表示されるアバタの表情を制御するシステムが開発されている[3]。これによりノンバーバル情報の伝達と意図しない個人情報の伝送の問題は解決するが、注視対象ユーザを特定するための視線制御機能の実現法は別途考える必要がある。

また、TV 会議型のもう一つの問題である視線 AW の欠落を解決する試みとして、ユーザ数に合わせて特殊なスクリーンを立体的に配置して、スクリーン背後にプロジェクタとカメラの両者を配置することで視線の問題を解決したシステム MAJIC[4]や、赤外線 LED と CCD カメラを用いてユーザの視線方向を検出し、仮想空間のアバタに反映させる GAZE[5]のようなシステムもあるが、装置が大がかりでユーザ数が固定される点が難点である。

このほか、特殊な装置を使用しないものとして、9 方向の静止画像を用いて擬似的に視線の表示をおこなう BrowserMAJIC[6]もある。WWW ブラウザ上で動作するため誰でも利用できるという利点があるが、ユーザの視線位置をマウスポインタ位置で代用しているため、各ユーザはマウス操作によって自らの視線方向を伝える必要がある。

また、これらの研究とは異なり、ユーザがネットワーク上の共有仮想空間内を自由に移動し、出会った他のユーザと音声チャットを楽しむことを可能とする、仮想空間音声コミュニケーションシステムの

*1: 東京農工大学大学院

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

研究がおこなわれている。FreeWalk[7]は、仮想空間内のウォークスルーと位置関係を用いた会話対象の制御を実現し、さらに知的エージェントを配置したときのユーザ挙動の解析が試みられているが、会話が可能範囲内に他のユーザが近づいてきたことを自然に通知する機能や、発話状態にあわせたアバタの視線AW制御機能など、コミュニケーションを促進するためのAW機能は有していなかった。

筆者らも、複数の遠隔ユーザがネットワークを介して仮想空間を共有し、足踏式空間移動インタフェースで同時に歩行するシステムを開発して自然な同時歩行を実現したが[8]、同様にユーザ間のコミュニケーションを支援するAW機能は有していなかった。本研究では、共有仮想空間歩行システムをもとに、IPネットワークにおける音声通信(VoIP)を用いた音声会話機能と、仮想空間内の位置関係を利用した会話対象の自動制御機能を実装し、さらに自然なコミュニケーションの促進を目的に、会話可能範囲内に他のユーザが存在することをジェスチャや足音などを介して通知する会話圏内の存在AW提示機能、ならびに注視対象を会話圏内の他のユーザに限定し、発話情報を用いてアバタの視線方向を制御する擬似視線AW機能を実装した。さらに、仮想空間音声コミュニケーションシステムにおける、それらの有効性について実験的に検討したので報告する。

2. システム設計

2.1 システム概要

ネットワーク上の複数のユーザが仮想空間を共有して、その中を自由に移動して自然な会話をおこなうシステムを実現するためには、まず仮想空間内のユーザ位置を管理し、互いの位置を通知する機構、ならびに自己と他者の位置に応じて仮想空間を描画する機能、および仮想空間における位置関係を用いて適切に会話グループを構成しグループ間で音声通信をおこなう機能、などが必要である。さらに、他者の存在認知を促す存在AW通知機能や、発言対象にあわせて視線方向を擬似的に制御する視線AW通知機能なども、会話開始のきっかけや自然な会話を

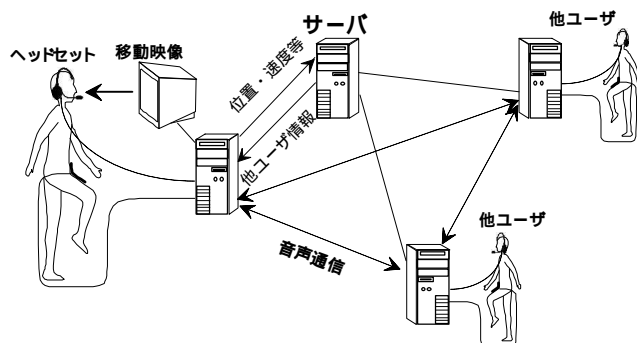


図1 遠隔共有仮想空間歩行・会話システムの構成

促進し、自然な会話コミュニケーションの実現に有効と考えられる。これらを下記の4項目にまとめ、本章および3,4章では、試作したシステムにおけるこれらの機能の実現方法を述べる。

- 1) 仮想空間位置管理とウォークスルー
- 2) 音声通信と会話対象の制御
- 3) 会話圏存在AW
- 4) 擬似視線AW

2.2 仮想空間位置管理とウォークスルー

物理的に離れたユーザが仮想空間を共有し、その中を自由に移動することで会話対象を選択し、コミュニケーションを図るためには、位置や速度などの情報を互いに送受信し共有する仮想空間システムが必要である。本研究では、これまでに開発したクライアント/サーバ式多人数共有仮想空間歩行システム[8]をもとに、後述の音声会話機能と会話圏における他のユーザの存在と、擬似的な視線のAW制御機能を実現した。このシステムは、図1のようにサーバで各ユーザの位置や速度を管理し、クライアントではサーバから通知される他者の位置、速度、曲率をもとに予測と補間をおこない、他者を3Dアバタとして表示するとともに、足踏み式歩行インタフェースを用いてウォークスルーをおこなうものである。1秒間に1回程度の低頻度の通信でも、予測機構によって他者の位置誤差が0.2m程度以下に維持される点が特徴である。図2に本研究で開発した共有仮想空間コミュニケーションシステムにおける、会話中の画像の例を示す。

2.3 音声通信と会話対象の制御

現実世界での会話対象は物理的距離が近いユーザに限定されること、また、仮想空間内に同時に多数のユーザが存在する場合の輻輳等を考慮すると、それぞれのユーザが会話可能なユーザ数を限定することが必要となる。仮想空間内のユーザを小グループに分割するためには、例えば仮想空間を複数の部屋に分け、会話グループを部屋ごとに構成する手法などが考えられる。しかし、この場合、仮想空間の構造が限定されるという問題が生じる。他方、中西らのFreeWalkでは、現実空間を模擬し距離に応じ

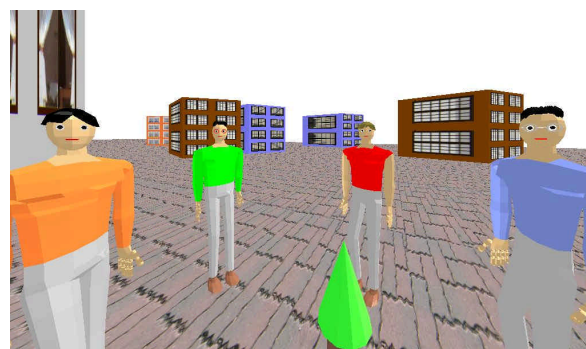


図2 評価実験中の会話風景

3. 会話圏存在 AW の実現

て音圧を小さくする制御をおこなうことで、会話対象を有限範囲内のユーザに限定している[7]。距離に応じて音量を制御する方法は、会話の対象を自然に制御可能である反面、マイク等からの混入雑音の影響を受けやすくなる。そこで、本研究では聞き取りやすさを優先し、以下の方法で会話対象を制御し、会話可能圏内では音圧が同一となる設計を採用した。

会話対象の検出には、Benford らが提唱するオーラ[9]を利用し、ユーザ間距離にもとづく検出法を用いた。会話対象の検出処理はユーザの位置情報を必要とするため、ユーザ位置管理サーバが兼ねる構成とした。サーバでは、ユーザごとに仮想空間に存在する他の全ユーザとの距離計算をおこない、図 3(a)のように距離がオーラ半径よりも短い場合に、両ユーザで会話グループを構成した。しかし、この処理のみでは、図 3(b)のように、ある程度の距離において 3 名以上のユーザが存在した場合、中央のユーザは両端のユーザと会話可能であるのに対して、両端のユーザ同士は会話が成立しないことになる。そこで、直接オーラ領域内に相手が存在しない場合でも、もう 1 人のユーザを介して鎖状にオーラが干渉する場合には、統合グループを構成する処理をおこなった。オーラ領域は、実験的に半径 8m の円形に設定した。また、評価実験では、ユーザがキーボードを押すことで会話を開始する手動会話制御機能も実装し、実験的に比較した。

音声通信機能は、遅延を最小限に押さえるため、互いに会話可能な圏内に存在するクライアント同士をサーバからの通知にもとづいて P2P 方式で接続する方式を採用した。プロトコルは通信速度を重視するため UDP を採用し、帯域を考慮して音声圧縮方法を選択可能にした。開発環境には Microsoft 社の DirectX を使用し、100BASE-T 環境内で同一グループ 6 名以上での同時通話が可能であることを確認した。

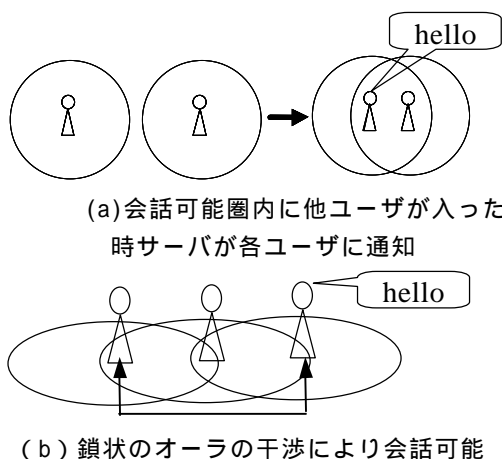


図 3 オーラを用いた会話対象の制御。(a)2 名の場合、(b)中央のユーザを介して両端のユーザも会話が可能になる場合

3.1 設計方針

オーラを用いた会話対象制御によって自動的な会話相手の制御が可能となるが、ユーザが会話可能であることを認識できなければ、円滑に会話を開始することができない。距離を用いた音圧制御による会話対象制御の場合も同様である。そこで、他のユーザが接近して会話が可能になったことに気づかせることが、自然なコミュニケーションの開始に重要な機能となる。本研究では、視覚と聴覚を介して、会話可能な圏内に他者が存在することに気づかせる存在 AW 情報提示機能（以下、会話圏存在 AW と表記）を実装し、その効果を実験的に検討した。

3.2 視覚的会話圏存在 AW

視覚的な会話圏存在 AW 情報は、拳手と足踏みの 2 つの動作によって提示した。拳手動作は、他のユーザが、新たに自分の所属するグループの会話可能圏内に入ったときに、会話可能状態になったことを視覚的に示す。また、会話圏内で静止状態にあるときには、後述の聴覚的存在 AW との整合性をとるために、片脚でリズムをとるような足踏み動作を提示した。拳手はコミュニケーションを図ろうとする意思を意味すると考えられ、物理的な接近とは意味が異なるが、仮想空間を自由に移動して不特定の対象と音声会話を楽しむような用途においては、自動制御によって会話のきっかけを与えることは会話促進の点ではむしろ好ましいこと、接近はコミュニケーションの意思がある時になされる場合が多いこと、またデスクトップ環境では小さな動作は認知しにくいこと、などから選択した。拳手動作の様子を図 4 に示す。

3.3 聴覚的会話圏存在 AW

聴覚的会話圏存在 AW は、視野の範囲外に存在する会話可能対象の存在を示すのに特に有効と期待される。また、音圧を適切に設定すれば、会話を妨げずに会話圏存在 AW を定常的に表現することが可能であり、会話可能圏内に入ったときのみの一時的な情報提示よりも大きな効果が期待できる。本研究では、歩行時と停止足踏み時に併せて、歩行時の足音と足踏み時の足音の 2 種類を切り替えて提示した。

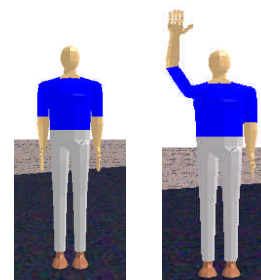


図 4 会話可能時の視覚的会話圏存在 AW

4. 視線 AW の擬似的実現

4.1 設計方針

現実空間における会話では、音声言語のバーバル情報だけでなく、ジェスチャや声の抑揚といったノンバーバル情報を利用することにより、様々な情報を相手に伝達している[10]。中でも、視線はコミュニケーションにおいて重要な機能を持っており、会話中に視線をそらせたり交差させたりする行為は、会話の開始合図や発話権の授受など会話の調整機能を有している[11]。そのため、視線交差の実現のための視線一致に関する研究がおこなわれており、MAJIC[4]や ClearBoard[12], Agora[13]は光学的に視線一致を可能にするものである。

また、視線には表現機能もあり[14]、特に凝視は大きな作用を持つ。凝視の自動制御に関する研究では、擬人化エージェントの凝視を運動モデルにもとづいて制御する試み[15]や、凝視の量や時間を制御することで相手に与える印象を制御する試み[16]などがある。

視線のもう一つの機能として、注視によって興味の対象を暗黙的に知らせる視線 AW がある。遠隔コミュニケーションシステムで視線 AW を実現するには、上述の視線一致システムに加えて、仮想空間を使用する場合にはアイトラックを用いてユーザの視線方向を検出しアバタ視線方向に反映させる[5]、あるいは音声認識によって会話対象を特定するなどの方向が考えられる。しかし、いずれも装置が大がかりになり、簡便に視線 AW を実現する方法はなかった。

ここで、日常的な多人数での雑談場面を想定すると、会話相手のいずれかを注視している時間が大半と考えられる。したがって、注視対象を他のユーザのアバタに限定しても、比較的問題はないと考えられる。さらに、会話中には話者を注視する傾向が強いと考えられるため、アバタに発話者を注視させることによって、擬似的に視線 AW(以下、擬似視線 AW)が実現されるものと期待される。

他方、会話中のノンバーバル動作を制御する方法には、渡辺らの音圧情報をもとに「うなずき」動作を提示するものがある[17]。会話の内容には関係なく音圧情報のみにもとづく制御であっても、会話促進に大きな効果が得られることが知られている。本研究では、注視対象を他のユーザのアバタに限定し、各ユーザの発話音圧情報と仮想空間における位置関係を利用してアバタの視線方向を制御する、擬似視線 AW の実現法を提案する。

4.2 視線方向の制御方法

現実空間における複数人数での会話を観察すると、聴衆は話者に対して、話者は以前会話をした相手、

または同時に会話をしている相手に対して、視線を向ける傾向がある。また、誰かが話している時にもう一人が話し始めると、後から話し始めた方が注目を集める傾向がある。本研究では、前者を話者注視効果、後者を発言開始効果と呼び、これらの効果を注目度 (Appeal Point 以下 AP) として定義し、音圧情報から算出される AP を用いてアバタの視線制御をおこなうことで、擬似視線 AW を実現した。各ユーザの AP を算出し、AP が最大となるユーザのアバタに視線を向けることで、話者注視が実現される。AP 最大となる話者のアバタは、2 番目に AP が大きなユーザを注視させる制御をおこなった。すなわち直前に発話していた、あるいは現在発話中の別のユーザを注視する設計とした。アバタの視線方向の表示は、遠方にあるアバタの視線方向の認識の容易さを考慮して、瞳の制御ではなく、頭部を回転して顔を注視対象のユーザに向ける方法でおこない、頭部の回転角が 90 度以上の場合には、自然な動作に見えるように体幹も回転させた。

4.2.1 話者注視効果

最も単純な話者注視の実現方法としては、各ユーザ音声のオン/オフを検出し、現在発話中のユーザのアバタに視線を集中させる方法が考えられる。ここで、複数のユーザが発言した場合を考えると、音圧の大きいユーザを検出する方法が、より自然である。また、過去によく発言したユーザは、その後も発言する可能性が高いため、話者が注目を集める効果は、現在だけでなく過去の発言の多寡にも依存する。これは、瞬時値に代えて、過去一定期間の発言総量を用いることで実現可能であるが、単純に適用すると、現在の発話中のユーザに視線が向かない可能性が生じる。そこで、話者注視効果 AP の算出にあたっては、現在の発話状況を優先して反映する必要がある。

4.2.2 発言開始効果

発言開始の判定は、単純には、音圧の過去値が 0 であり、現在値が 0 より十分に大きければ、発言開始と判定できる。しかし、実際の会話音声には断続が生じるため、発言開始検出後の一定時間は発言開始の検出をおこなわないなど、発言開始と断続を区別する必要がある。また、発言開始時に注目を集める効果は、話者注視効果と比較して瞬時的な性格をもつため、数秒以内に効果がなくなるのが適当と考えられる。

4.3 注目度の計算

本項では、前項までに述べた効果の具体的な計算方法を述べる。まず、話者注視効果による注目度 AP_c は、(1)式のように過去 60 秒間の音圧の積分値を用いることで、他に発言者がいないときには、過去の発話量の多い者が注目を集める効果を実現した。現在値を優先するための重み係数は、指数関数状に減少する時間関数とし、時定数 t は実験的に 1.5 秒と

した．計算に用いる音圧は，人間の知覚特性を考慮して対数変換をおこなった．時刻 t における音量を $v(t)$ とするとき，話者注視効果による注目度 AP_C は，次式を用いて算出した後，0~1 に正規化した．

$$AP_C = \int_{-60}^0 \log(v(t)) \exp\left(\frac{t}{60}\right) dt \quad \dots (1)$$

発言開始効果による注目度 AP_S については，発言開始と音声の断続を区別するため，過去 5 秒以内に発言がある場合には，発言開始効果が生じないものとした．また，発言開始効果は瞬時的性格のものであるため，5 秒後に効果が 0 となるよう線形に減少させた．発言開始時刻を t_s とすると，時刻 t における発言開始効果による注目度 AP_S は(2)式で定義される．

$$AP_S = \frac{5 - (t - t_s)}{5} \quad \dots (2)$$

($t_s \leq t \leq t_s + 5$ ，ただし発言開始検出後 5 秒間は発言開始検出をおこなわない)

最終的な AP は，これら話者注視効果による注目度 AP_C と発言開始効果による注目度 AP_S を用いて，以下のように定義した．

$$AP = aAP_C + bAP_S \quad \dots (3)$$

a, b の値は実験的に調整し，1:2 とした．話者注視効果と発言開始効果によって，ユーザは，受話時には他のユーザのアバタの視線から発話者をより明快に認知することが可能になり，発話時には聴衆の受話が強調される印象を持つものと期待される．したがって，発話終了の認知の補助や話者の意図を表現する機能は持たないが，話者や受話感の認知補助や増強によって，発話権の授受になんらかの心理効果を持つ可能性が期待される．

また，最も注目度の高いユーザのアバタを注視する処理をおこなった場合，話者交代が起きなければ，そのユーザを注視し続ける．しかし，長時間の持続的な注視は支配感により悪い印象を与える可能性があるため[16]，30 秒に 1 回程度の頻度でランダムに他のユーザのアバタに視線を変更する処理を加えた条件も設定し，実験的に比較した．

4.4 自己優先処理

一般に，話者は聴き手に注視されると，相手が聞いていると認知するものと推察される．また，発話終了を検出して適切な遅延でアバタをうなずかせることで，話者にアバタが話を聞いている印象を与えることも可能である[17]．そこで遠隔共有仮想空間においても，計算機を使用しているローカルユーザが発言している時に，アバタのローカルユーザ注視確率を高めることで，より「相手が聞いている感覚」

が増強されるものと期待される．

また，前節までは，一般的な話者注視の実現方法を検討したが，遠隔共有仮想空間システムにおいては，各ユーザに提示される仮想空間における視線方向が矛盾していても問題は生じにくい．例えば，A，B，C の 3 ユーザがいる場合，ユーザ A の画面内ではユーザ B はユーザ A を，ユーザ C の画面内ではユーザ B はユーザ C を見ている大きな問題は生じない．

そこで，この点を利用して，話者注視効果による注目度をローカルユーザのみ 2 倍にし，他のユーザの発言開始効果発生条件を，過去 30 秒間の沈黙とする処理をおこなった．これらの処理により，他のユーザが発言中でもローカルユーザが発言すると優先的に注視され，逆にローカルユーザが発言中には，他のユーザに注視が移動しにくくなる．以後，これらの処理を自己優先処理と呼ぶ．

5 . 評価実験

5.1 会話圏存在 AW と会話制御の評価

5.1.1 会話圏存在 AW と会話制御の評価

会話開始制御と会話圏存在 AW の，コミュニケーションの自然さへの効果を検討することを目的に，2 件法の一対比較法を用いて主観評価をおこなった．被験者は，本システムを初めて使用する 22~28 歳の情報系男子学生 6 名とし，図 5 のように，色分けした目印を配置した仮想環境内の異なる 6 地点に 1 名づつランダムにユーザを初期配置した．被験者には，初期位置から，2 人づつの組が 3 組できるように設定した地点への移動を目標位置の色を用いて指示し，目標位置の近傍にいる相手に対し会話が可能であると判断したら，相手の名前をできるだけ早く聞き出す課題を与えた．表 1 に実験条件を示す．課題は，実験条件ごとに初期位置や会話対象を変えてそれぞれ 6 回連続しておこない，これを 1 セットとして一対比較をおこなった．主観評価は各比較の終了時とし，どちらがより自然に感じるか，好ましい方を回答させた．なお，本実験は名前を聞き出す課題であるため，アバタの固有情報(髪型，服装など)は同一にし，移動の指示は会話の相手が続けて同一

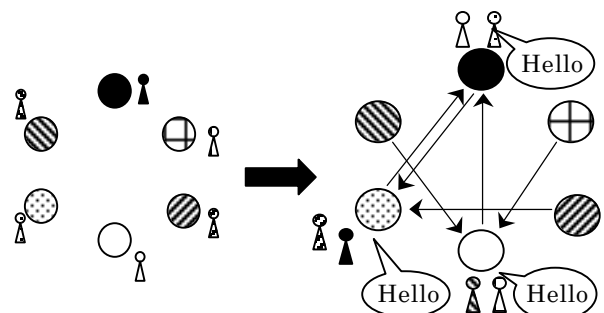


図 5 存在 AW 評価実験のモデル図

にならないように設定した。なお、手動会話制御と自動会話制御のそれぞれにつき、会話圏存在 AW を提示する条件で 5 分間の操作習熟の後に、実験をおこなった。AP を用いた擬似視線 AW 制御は、本実験ではおこなわなかった。

また、以降すべての実験において、仮想空間内の移動は、各ユーザの移動速度を統一する目的で、歩行インタフェースではなくキーボード操作でおこなった。操作は、左右のカーソルキーが仮想空間での左右の回転に、前後が加速・減速に相当し、速度上限を設定することで、ユーザの速度を統一した。実験の順番は、それぞれの条件の提示順が偏らないように、あらかじめ設定した。また、被験者をそれぞれ別の部屋に配置することで、直接音声が届かないようにした。

表 1 会話制御法と会話圏存在 AW 比較実験の条件

会話開始制御	会話圏存在 AW
手動	無
自動	無
手動	AW 有 (視覚 + 聴覚)
自動	AW 有 (視覚 + 聴覚)
常時会話	-

5.1.2 視覚と聴覚の存在 AW の比較

次に、視覚的存在 AW と聴覚的存在 AW の会話開始への効果を比較する目的で、前項と同様の 2 件法の一対比較法による実験をおこなった。実験条件は、視覚 + 聴覚、視覚、聴覚の計 3 種類とした。被験者は会話制御の評価実験と同一の 6 名とし、各条件のコミュニケーションの自然さに関する主観評価をおこなった。なお、前章の実験結果から自動会話制御を用いた。

5.2 擬似視線 AW 制御の評価実験

5.2.1 擬似視線 AW 機能の評価実験

話者注視効果、発言開始効果、およびランダム性の、自然な会話への効果を比較することを目的に、コミュニケーションの自然さに関する主観評価をおこなった。被験者には、より自然なコミュニケーションが可能と感じた条件を、実験の最後に好ましい順に回答させ、順序尺度を求めた。また、提案する擬似視線 AW 機能は、遠隔ユーザの視線方向を意味しないため本来の会話調整機能は有さないが、話者の認知を補助することになるため、話者交代になんらかの会話調整機能と類似した効果が生じる可能性も考えられる。そこで、課題遂行中の発話が終了してから次に誰かが発話を開始するまでの時間(以後、沈黙時間)を、計測した。被験者は 22~26 歳の情報系男子学生 10 名とし、5 名ずつの 2 群に分割し、各 5 名に対して次の課題を課した。

課題では、仮想環境内で他の 4 人が見えるように各ユーザを配置し、実験者から指示されたテーマに

従って 5 分間会話することを指示した。テーマは、あらかじめ日常生活やスポーツなどに関して 50 テーマを選定し、調査紙を用いて被験者が会話や議論が困難と考えるテーマを削除した中から選択した。

表 2 に実験条件を示す。実験は順序効果を考慮し、ランダムとした。仮想空間内の移動は、会話圏存在 AW 評価実験と同様にキーボード操作でおこない、実験前に 5 分間の習熟期間を設けた。なお、会話圏存在 AW に関しては、先の実験結果においてもっとも自然だと評価された条件(自動制御で会話圏存在 AW ありの条件)で実験をおこなった。

表 2 擬似視線 AW 制御評価実験の条件

視線制御なし
ランダム(30 秒に 1 回の頻度で他ユーザに視線移動)
話者注視
話者注視 + 発言開始
話者注視 + ランダム
話者注視 + 発言開始 + ランダム

5.2.2 自己優先処理の評価実験

前項の擬似視線 AW 機能の評価実験においては、話者注視効果を用いた視線制御に加えて、ランダム性や発言開始効果を加えたときの効果を比較した。後述のように、話者注視効果と発言開始効果を組合せた条件で、自然な印象が得られる傾向が認められたので、さらに、ローカルユーザが発言を開始した場合に他ユーザからの注目を得やすく、それを奪われにくくするような自己優先処理を付加した場合と、ローカルユーザと遠隔ユーザの発言開始効果を同一とした場合の比較実験を、前項と同様の手順でおこなった。

表 3 に実験条件を示す。自己優先処理でローカルユーザ以外の発言開始効果生起条件を 30 秒間の沈黙としたものにあわせて、発言開始効果生起条件を一樣に 5 秒間の沈黙としたものと 30 秒としたものの 2 種類、および発言開始効果なしの合計 4 条件に対して比較をおこなった。

表 3 自己優先処理の評価実験条件

発言開始なし
発言開始[5 秒沈黙]
発言開始[30 秒沈黙]
自己優先[発言開始(自 5 秒・他 30 秒沈黙) + 自己 AP2 倍]

6. 実験結果

6.1 会話圏存在 AW と会話制御の実験結果

6.1.1 会話開始制御と会話圏存在 AW の効果

評価実験の結果をもとに、サーストンの方法で間隔尺度を求めたものを図 6 に示す。まず、会話圏存在 AW に関しては、AW 提示条件が自然な会話と認知し好まれる結果となった。これは会話圏存在 AW

により、他ユーザが会話圏内に入ったことの認知が容易になったことが一因と考えられる。また、会話圏存在 AW の有無にかかわらず、自動会話制御が手動制御より自然な印象を与える結果となった。手動制御では意図的に会話対象を制御する必要があるため、現実空間での会話開始との相違が違和感を与えた可能性が考えられる。

常時会話条件で評価が低かったのは、ユーザ数が増加すると他のユーザと発話が重なる確率が高くなることや、遠方のユーザの声が聞こえることが不自然に感じた原因と考えられ、このことは、被験者の内観報告からも裏付けられた。

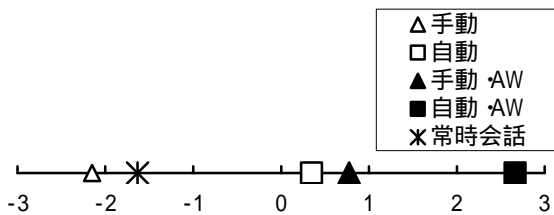


図6 会話開始制御法の一対比較結果

6.1.2 視覚と聴覚の存在 AW の比較

実験結果をもとに、サーストンの方で間隔尺度を求めたものを図7に示す。視覚と聴覚の単独提示を比較すると、聴覚より視覚の方がコミュニケーションの自然さの点で好まれる結果となった。各種感覚情報を矛盾して提示した場合に、視覚を介して提示された情報が優位になる現象が知られているが、会話対象の存在認知においても、視覚情報を介した提示が自然さの点でも優位な結果となった。また、視覚と聴覚の両方の会話圏存在 AW 情報を提示した場合、単独での提示よりも主観評価が高い結果となった。これは、視覚と聴覚のマルチモーダルな情報提示によって、より現実空間に近い情報提示がおこなわれたことに加え、視野角外に他ユーザが存在するため視覚情報のみでは会話圏存在 AW 情報が提示できない場合にも、聴覚情報により提示が可能になったことも一因と考えられる。

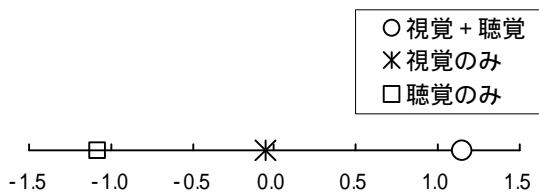


図7 会話圏存在アウェアネスの一対比較結果

6.2 擬似視線 AW の実験結果

6.2.1 擬似視線 AW の効果

順序尺度を用いた自然さの主観評価結果を図8に、平均沈黙時間を図9に示す。主観評価において、視線制御をおこなわない条件は、すべての被験者が最

も不自然と感じ、次いでランダムのみが不自然な印象を与える結果となった。この2条件では、ランダムの方が自然な印象を与えたことから、擬似視線 AW の4条件においても、ランダムを付加した2条件が、より自然な印象を与えるものと予想されたが、逆の傾向が見られた。複数のユーザがそれぞれ発話したため、視線制御をおこなう4条件においてランダム制御の有効性が期待される長時間の凝視が生じる場面が少なく、逆にアバタのランダムな視線移動が、話者注視効果や発言開始効果による擬似視線 AW を阻害したためと考えられる。ただし、ランダム制御の評価が被験者によって大きくばらついているように、中にはランダム性を好むユーザも存在し、より自然な持続的凝視の回避の検討が必要である。

また、発言開始効果を付加した2条件は、発言開始効果のない2条件よりもやや評価が高い傾向が見られた。発言開始に対して小さな遅延でアバタの視線方向が変化し、発言開始を認識しやすくなったためと考えられる。顕著な効果が得られなかった原因は、発言開始効果はローカルユーザへの注視応答を改善すると同時に、遠隔ユーザへの応答も同様に強化するためと考えられた。そこで、次節では、ローカルユーザの発言開始効果を強化した場合に関して検討する。

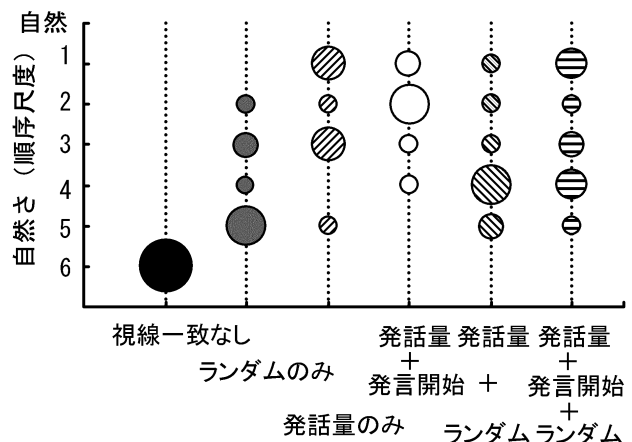


図8 各擬似視線 AW 制御条件における主観評価 (円の大きさは回答頻度を意味する)

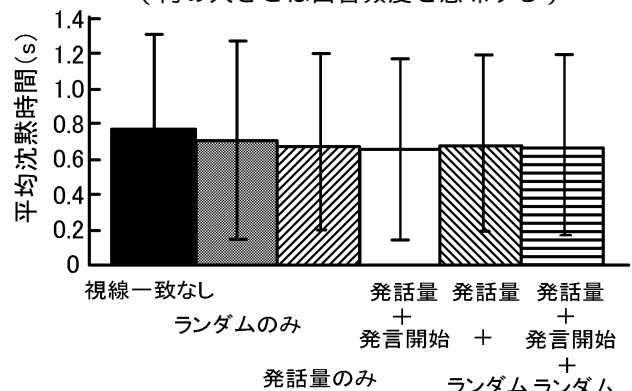


図9 各擬似視線 AW 制御条件における平均沈黙時間

平均沈黙時間は、視線制御なしと他の5条件の間、およびランダムと話者注視+発言開始の間においてのみ、t検定で危険率5%の有意差が見られた。有意差がみられたのは一部であるが、主観的により自然な会話と認知された条件において短くなる傾向を示した。

6.2.2 自己優先処理の効果

順序尺度を用いた自然さの主観評価結果を図10に、平均沈黙時間を図11に示す。主観評価に関しては、発言開始効果を、ローカルユーザと遠隔ユーザで同一とした2条件は、発言開始よりもやや自然な印象を与える傾向を示したが、特に顕著な差は見られなかった。

これに対し、他者の発言開始効果生起条件のみを30秒間沈黙とし、ローカルユーザのAPを2倍とする自己優先条件は、5秒と30秒のいずれよりも自然に感じるという結果となった。これは、自己が発言を開始した時には注視を集めやすく、他のユーザが発言を開始しても視線の移動が生じにくくなったことが、好まれたものと考えられる。すなわち、ローカルユーザの発言に対する視線集中を強調する制御は、自然で好ましい印象を与えたと言える。平均沈黙時間は分散も大きく、各条件間での有意差を見出すには至らなかった。

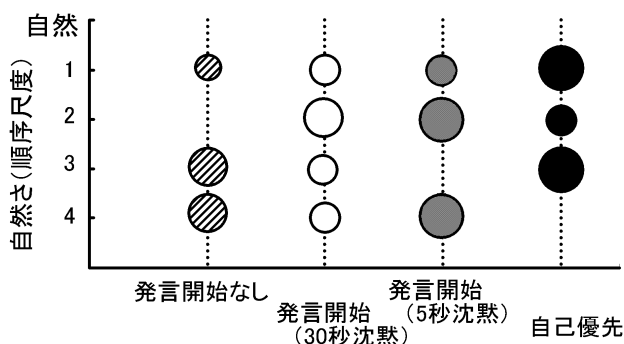


図10 自己優先処理の主観評価 (円の大きさは回答頻度を意味する)

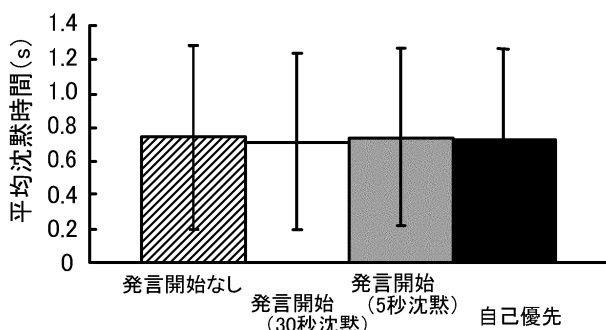


図11 自己優先処理と平均沈黙時間

7. 考察

実験において、会話開始制御は、自発的意図による手動制御よりも自動会話制御が、より自然である

という評価結果となった。これは、現実空間の会話開始が発声の一動作で済むため、同じように発声だけで済む自動制御が自然に感じられたと解釈できる。しかし、手動制御は発声に加えてキーボードの押下動作が必要であり、手動制御の操作性が低かったため、相対的に自動制御が自然に感じられたとも解釈できる。自然に感じた原因が現実空間との同一性が操作性か、現段階では特定できず、検討が必要である。また、自動制御を用いた場合には、会話を意図しない場合にも接近によって会話が可能となる可能性が存在するが、自動制御はユーザに会話のきっかけを与え、活発なコミュニケーションの促進につながることを期待される。会話を意図しない場合は、発話をおこなわなければよいだけであり、問題は生じない。したがって、仮想空間での位置を利用した会話対象の自動制御は、会話開始制御として有効といえる。

他者が会話可能圏内に存在することを通知する存在AWを、アバタの挙手動作と足踏み音を用いて視覚と聴覚を介して提示した結果、会話開始制御法にかかわらず、ユーザの接近の認知に会話圏存在AWが有効であることが示唆された。しかし、3.2で述べたように、コミュニケーションにおいて挙手は接近と異なる意味を持つため、会話の意図をもたないユーザが互いの会話圏に入ったときに、自動制御による挙手は、厳密にはユーザの意図と矛盾することになる。そのため、挙手ではなく、会釈など積極的な会話の意図がない場合にも社会的儀礼として交わされることが多い動作を会話圏存在AWとして利用することなども考えられる。ユーザの状態を反映した動作種類の変更の可能性なども含め、さらに検討が必要である。

擬似視線AW制御の主観評価では、視線制御なしよりもランダムが、さらに擬似視線AWの4条件が、より自然なコミュニケーションが可能という主観評価結果となった。擬似視線AWの4条件は、アバタの視線によって、受話時には発話者を明確にし、発話時には聴衆の受話を強調する。したがって、発話終了の認知の補助や話者の意図を表現する機能は持たないが、発話権の授受になんらかの心理効果を持つ可能性が考えられる。他方、評価実験におけるランダム条件は、視線制御なしと比較してアバタの視線が自律的に変化する点で異なる。また、APを用いた擬似視線AW条件では、他者の発話に応じて視線が変化することにより自律性が、また自己の発話によって視線が変化することによりインタラクティブ性がもたらされる。アバタが自律性やインタラクティブ性を持つことは、アバタに人間らしい印象を与え、これがコミュニケーションの自然さにつながった可能性も考えられる。

ここで、擬似視線 AW 制御の沈黙時間に注目すると、視線制御なしよりも、ランダムおよび擬似視線 AW の 4 条件での沈黙時間が危険率 5% で有意に短かった。これは、アバタ視線運動の自律性が自然さをもたらした、という上記の仮説を支持する。しかし、有意な差ではないものの、擬似視線 AW の 4 条件すべてがランダムよりも沈黙時間が短いことから、発話権授受になんらかの効果をもたらした可能性も否定には至らない。さらに今後の検討が必要である。

以上のように、本来の視線 AW を実現するためには、なんらかの方法でユーザの視線方向を計測する必要があるが、本研究では、音圧を利用した発話者への視線集中によって、興味の対象を会話圏内のユーザに限定し、擬似的な視線 AW 制御をおこなった結果、その有効性が示された。今回は制御をおこなっていない発話終了の認知を補助するなど、さらにユーザの認知を補助することによって、カメラなどのデバイスを使用せずに自然なコミュニケーションを支援する可能性の検討が、今後の課題である。

また、仮想空間での同時多人数コミュニケーションを実現するためには、一度に多くのアバタを表示する必要が生じる。しかし、小型ディスプレイの限られた視野角では表示可能な人数が限定されるという問題がある。映像に歪みが生じるが、計算上の視野角を大きくし、併せて音源定位を用いて発話者の特定を補助する方法などが考えられる。

本研究では LAN 内で評価実験をおこなったため、会話の障害となるほどの音声遅延は発生しなかった。しかし、より一般的にインターネットの利用を想定した場合、音声の遅延が想定され、そのような環境において会話を促す機能を検討する必要がある。例えば、遠隔ユーザからの返答がなくても、発話終了時に遠隔ユーザのアバタを自動的にうなずかせるなどの方法が考えられる。仮想空間における自然な多人数コミュニケーションの実現に向けて、発話にあわせた口唇の制御や表情の制御など、より豊かなノンバーバル情報の提示とその自動制御機構の検討が望まれる。

8. まとめ

本研究では、会話圏における他者の存在と、対象を会話相手に限定し発話情報によって制御する擬似視線 AW を提示する遠隔共有仮想空間内会話システムを開発し、その効果を実験的に検討した。その結果、会話対象の自動制御と会話圏存在 AW 提示が、自然なコミュニケーションに有効である可能性が示された。また、擬似視線 AW の提示は、多人数音声会話において、より自然な印象につながることを示された。より一層自然な仮想空間コミュニケーション

のためのノンバーバル情報の伝送と提示が今後の課題である。

謝辞

本研究の一部は総務省戦略的情報通信研究開発推進制度によるものである、ここに記して感謝する。本研究への助言をいただいた、北陸先端科学技術大学院大学國藤進教授、和歌山大学宗森純教授、吉野孝助教授、メディア教育開発センター大澤範高教授、および、本学中川正樹教授に感謝する。

参考文献

- [1] 松下温, 岡田謙一, " コラボレーションとコミュニケーション ", 共立出版(1995)
- [2] 垂水浩幸, " グループウェアとその応用, ソフトウェアテクノロジーシリーズ第 12 巻 ", 共立出版(2000)
- [3] 沖電気, " FaceCommunicator ", <http://www.oki.com/jp/FSC/vc/fcbbe/about.html>
- [4] 前田典彦, J.Giseok, 市川裕介, 松下温, 岡田謙一, "MAJIC:場の雰囲気重視したTV会議", 情報処理学会論文誌, 36(3), 775-783(1995)
- [5] R. Vertegaal, " The GAZE GroupWare System: Mediating Joint Attention in Multiparty Communication and Collaboration ", CHI'99, 294-301(1999)
- [6] 岡田謙一, 松下温, " 静止画像を用いた狭帯域ネットワーク用多地点会議システム ", 情報処理学会論文誌, 39(10), 2762-2769 (1998)
- [7] 中西英之, 吉田力, 西村俊和, 石田亨, " FreeWalk: 3次元仮想空間を用いた非形式的なコミュニケーションの支援 ", 情報処理学会論文誌, 39(5), 1356-1364(1998)
- [8] 下地崇, 藤田欣也, " 足踏式移動インタフェース WARP を用いた多人数共有仮想空間歩行システムの試作 ", 日本バーチャルリアリティ学会論文誌, 8(1), 11-18(2003)
- [9] S.Benford and L.Fahlen, " A Spatial Model of Interaction in Virtual Environments ", ECSCW'93, 109-124(1993)
- [10] 黒川隆夫, " ノンバーバルインタフェース, ヒューマンコミュニケーション工学シリーズ ", オーム社(1994)
- [11] S.D.Duncan, Jr., and D.W.Fiske, " Face-to-face interaction " Research, Methods, and Theory, Hillsdale, N.J.:Lawrence Erlbaum(1977)
- [12] M. Kobayashi and H. Ishii, " ClearBoard: A Novel Shared Drawing Medium that Supports Gaze Awareness in Remote Collaboration ",

Trans. IEICE , E76-B(6) , 609-617(1993)

- [13] H. Kuzuoka, J. Yamashita, K. Yamazaki and A. Yamazaki , "Agora: A Remote Collaboration System that Enables Mutual Monitoring " , CHI'99 , 190-191(1999)
- [14] A.Kendon, " Some functions of gaze-direction in social interaction " , Acta Psychologica, 26, 22-63(1967)
- [15] 中沢正幸, 西本卓也, 嵯峨山茂樹, " 擬人化音声対話エージェントにおける視線制御モデルの提案 " , 人工知能学会 SIG-SLUD-A303 , 21-26(2004)
- [16] 深山篤, 大野健彦, 武川直樹, 澤木美奈子, 萩田紀博, " 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 , 43(12) , 3596-3606(2002)
- [17] 渡辺富夫, 大久保雅史, 中茂睦裕, 檀原龍正, " 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 , 2(2) , 21-29(2000)

(2004 年 6 月 9 日 受付)

[著 者 紹 介]

宮島 俊光 (学生会員)



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コミュニケ
ーション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,コミュニケーション支援の研究

に興味を持つ .

下地 崇 (学生会員)



2002 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 . 同
大学工学研究科情報コミュニケー
ーション工学専攻卒業 . コミュニケー
ーション支援の研究に興味を持つ .

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
仮想および空間力触覚や歩行感覚
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) .